

Telektronikk 1.94

Prognostisering



Innhold

TEMA:

- Introduksjon, *Kjell Stordahl* **3**
- Prognosemetoder – en oversikt, *Kjell Stordahl* **5**
- Bruk av ekspertundersøkelser til prognoser for nye tele-tjenester i privatmarkedet, *Kjell Stordahl* **53**
- Underlag for prognoser for telefonabonnement: Data og analyse, *Carlo Hjelkrem og Johannes Bøe* **65**
- Glattingsmodeller,
Johannes Bøe og Carlo Hjelkrem **76**
- Regresjonsmodeller,
Kjell Stordahl, Carlo Hjelkrem og Johannes Bøe **87**
- Vekstmodeller, *Johannes Bøe og Carlo Hjelkrem* **103**
- Box-Jenkins metode,
Carlo Hjelkrem og Johannes Bøe **110**
- Sammenlikning av prognosemodeller fra ulike metoder,
Johannes Bøe og Carlo Hjelkrem **125**
- Prognoser for abonnementssetterspørsel,
Carlo Hjelkrem og Johannes Bøe **134**
- Tellerskritt – grunnlagsdata og prognoser,
Johannes Bøe og Carlo Hjelkrem **145**
- Prognoser for nye tjenester hvor data er mangelfulle,
Johannes Bøe og Carlo Hjelkrem **156**
- Prognoser for pris på nettkomponenter basert på utvidelse av lærekurvmodellen,
Borgar Tørre Olsen og Kjell Stordahl **166**

SPESIAL:

- Scenarios for introduction of broadband services in telecommunication networks,
Rolf B Haugen and Bente Mannsåker **173**

STATUS:

- International research and standardisation activities in telecommunication: Introduction, *Endre Skolt* **183**
- Signal processing, *Gisle Bjøntegård* **184**
- Data networks and open system communications, *Berit Svendsen and Mette Røhne* **186**
- Teletraffic and dimensioning, *Harald Pettersen* **188**
- Terminal equipment and user aspects,
Trond Ulseth **191**
- Telecommunications Management Network,
Ståle Wolland **194**
- A presentation of EURESCOM, *Endre Skolt* **196**

Telelektronikk

90. årgang Nr. 1 - 1994
ISSN 0085-7130

Redaktør:
Ola Espvik
Tel. + 47 63 80 98 83

Redaktør, Status:
Endre Skolt
Tel. + 47 63 80 92 11

Redaksjonssekretariat:
Gunhild Luke
Tel. + 47 63 80 91 52

Adresse:
Telelektronikk
Televerkets Forskningsinstitutt
(Norwegian Telecom Research)
P.O. Box 83
N-2007 Kjeller, Norway

Redaksjonsråd:
Viseadministrerende direktør Ole P Håkonsen
Forskningsdirektør Karl Klingsheim
Strategidirektør Bjørn Løken

Design:
Design Consult AS

Utforming og illustrasjoner:
Gunhild Luke, Britt Kjus, Åse Aardal
Televerkets Forskningsinstitutt



Introduksjon

AV KJELL STORDAHL

På samme måte som Televerket har utviklet seg har også behovet for prognoser og prognosearbeid utviklet seg.

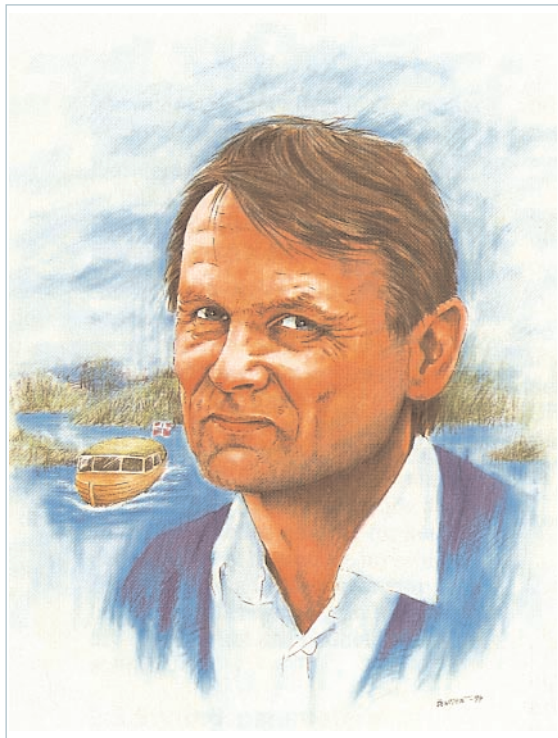
Tidligere var det få tjenester og få nett. Går vi en del år tilbake leverte Televerket hovedsakelig tjenestene telefoni og teleks. I dag er det en rekke tjenester, og i den infrastruktur som Televerket nå har, tilbys også store mengder tilleggstjenester og også skreddersydde kommunikasjonsløsninger. Dette stiller krav til håndtering av et stort volum med ulike prognoser.

For under ti år siden ble de siste ventelistene på telefonabonnement avskaffet. I perioden før var ventelistene til dels svært lange, og det var ikke noen særlig god ide å reklamere for hovedabonnement. Markedsaktivitetene i Televerket var derfor svært begrenset. På samme måte var også prognoseaktivitetene begrenset.

Tidligere var også utbyggingen av telenettet mer stasjonært. Det var færre behov å ta hensyn til ved planlegging og utbygging av nettet. Samtidig var det sterke begrensninger på tilgang til investeringsmidler. Dette gjorde blant annet at prognosene ikke i samme grad som nå var underlag for den reelle nettutbyggingen. Andre faktorer, spesielt distriktpolitiske ut fra Televerkets samfunnsansvar, ble brukt til prioritering av fordelte investeringsmidler.

Televerket ble heller ikke, på samme måte som i dag, betraktet som en forretningsenhet. Nå stilles det sterkere krav til oppfølging av inntekter og kostnader. I den sammenheng er det også viktig å få kontroll på inntektsutviklingen ved å lage gode prognosemodeller og å se på effekten av ulike prisstrategier. På kostnadssiden er det svært viktig å foreta en optimal bruk av investeringsmidlene. Trafikk- og abonnementsprognosene gir et svært viktig underlag her.

Ser vi bort fra den tiden det var separate private telefonanlegg i Norge, har Televerket ikke opplevd særlig konkurranse på sine basistjenester. Denne utviklingen er nå i ferd med å snu. Ut fra internasjonal utvikling og ikke minst EUs direktiver går Televerket mot full konkurranse på alle sine tjenester og etter hvert også sin basis infrastruktur. Det er konkurranse på mobiltelefon og datatjenestene. Fra 1 januar 1993 ble det også tillatt for andre nettoperatører å selge kapasitet som er leid i Televerkets infrastruktur. Det vil sannsynligvis ikke gå mange år før Televerkets siste eneretter: telefon og basis infrastruktur inkludert satellitt – vil bli avskaffet. Dette betyr at Televerket allerede nå må innrette seg etter en kommende fri konkurranse. I motsetning til tidligere må det derfor utarbeides strategier for organisasjon, bemanningsutvikling, inntekter/kostnader, prisstrategier, etc, med sikte på den kommende konkurransesituasjonen. Prognoser er et svært viktig underlag i disse strategiene.



Erfaringer fra prognosearbeid har vist at bearbeiding og tilrettelegging av underlagsdataene er den mest tidkrevende og også viktigste del av prosessen. Nærmere 80 % av arbeidet gjøres her. Det er svært viktig for de som lager prognoser å ha god kjennskap til underlagsdataene. Det er først når en får god innsikt i underlagsdataene og i hvorledes de endrer seg over tid, at en får forståelse for selve prognosemodellen. Et eksempel her er at da Dagbladet fjernet sine to sider med annonser på sex-samtaler så forårsaket det en markant nedgang i veksten på utenlandstrafikk. Ved å gå inn i underlagsdataene var det lett å identifisere at nedgangen i veksten på utenlandstrafikk i den gitte perioden ikke var generell, men kom i spesielle takstklasser der det gikk sex-samtaler til blant annet Hong Kong og Nederlandske Antiller.

Det må understrekes at det er svært viktig å få etablert systemer, målinger og underlagsdata/statistikk slik at en har en database som viser en historisk utvikling. Først når dette er på plass er det mulig å lage gode prognoser samtidig som det selvsagt er viktig å ha et referansenivå ut fra den historiske utviklingen.

En prognosemodell skal på best mulig måte beskrive en fremtidig utvikling. Sammenlikning mellom prognosene og faktisk observerte verdier viser hvor gode prognosene og prognosemodellen er. Ved å følge opp utviklingen – eksempelvis måned for måned – er det på et tidlig tidspunkt mulig å detektere og justere for spesielle hendelser som påvirker utviklingen.

Televerket har inntil nylig vært i en enerettsituasjon på områder hvor hovedinntektene er hentet inn. Etter hvert vil andre nettoperatører komme inn og sakte men sikkert ta markedsandeler. Dette vil gjøre prognosearbeidet vanskeligere, men samtidig mer utfordrende.

Det er å håpe at dette nummer av Teletronikk blir lest av mange. Den første artikkelen gir en generell innføring i prognosemetoder. De øvrige artiklene gir mer detaljert innføring i ulike prognosemodeller, samtidig som det vises hvorledes prognosemodellene anvendes på ulike problemstillinger.

1 Utvikling av prognosearbeidet i Televerket

Starten til den mer profesjonelle håndtering av prognosefunksjonen skjedde på begynnelsen av 1970 tallet. I tilknytning til den store omorganiseringen ble det opprettet et prognoseutvalg som kartla 130 ulike prognosebehov. Etter anbefaling av utvalget ble det opprettet et sentralt prognosekontor ØSP på Økonomiavdelingen i Teledirektoratet som fikk sentralt prognoseansvar.

Midt på 1980-tallet ble det foretatt en justering. Prognoseansvaret ble for de enkelte tjenester overført til Tjenesteavdelingen i Teledirektoratet, men det tekniske ansvaret for utarbeiding av prognosene og også regional koordinering av prognosene var fremdeles tillagt ØSP, som så ble kalt ØSU – Utredningskontoret.

I tilknytning til ny organisasjon 1 januar 1993 er prognoseansvaret fordelt på de enkelte divisjoner.

2 Prognosearbeid nå

2.1 Innledning

Prognoser er blitt et svært viktig element innen strategisk planlegging/ langtidspanlegging, markedsplanlegging, nettutbygging og ikke minst som underlag i budsjettprosessen i Televerket.

Det vil bli sett nærmere på hvorledes prognosene anvendes i budsjettarbeidet – både i tilknytning til inntektsbudsjettet og investeringsbudsjettet – og i strategisk planlegging. Prognosene er også et viktig fundament for planlegging av framtidig aktivitets- og bemanningsvolum innen ulike sektorer av Telegruppens virksomhet.

Prognoser anvendes også på en rekke felter i tilknytning til utbygging av telenettets infrastruktur.

2.2 Bedriftsplan/strategiplan

Televerket laget tidligere hvert år en bedriftsplan på et meget overordnet nivå. Basert på denne lager så den enkelte region en strategiplan som også er forholdsvis overordnet. Deretter lages det så handlingsplaner ut fra de rammer og strategier som er trukket opp på overordnet nivå.

Bedriftsplan og strategiplan har en planperiode på 5 år. Handlingsplanene har samme perspektiv, men de er mer konkret spesielt for nærmeste år.

I strategiplan er det foretatt analyser av framtidige inntekter og kostnader. Dette gir dermed rammer for den framtidige virksomheten.

Anslag for de forventede inntektene er basert på forutsetninger om priser i årene framover og prognoser for trafikk- og abonnementsinntekter for ulike tjenester. Dette er som en forstår, svært viktige prognoser for Televerket. Inntektsprognosene angir hvilke muligheter og frihetsgrader Televerket vil ha de kommende årene. Ut fra disse settes rammer for kostnadene. Det vil si hvor store investeringer som kan tillates, hvor stor bemanning Televerket kan ha, hvor mye Televerket kan nedbetale på sine statslån, etc.

2.3 Styringsparametere

I enhver bedrift vil det være spesielle parametere som er svært viktige for bedriftens utvikling. Det trenger ikke være det samme sett med styringsparametere hvert år. Ulike tjenester og produkter kan variere fra år til år. Det samme gjelder betydningen av servicekvalitet.

Den enkelte divisjon og den enkelte region blir nå mer og mer betraktet som en selvstendig resultatenhed. Det samme gjelder selvsagt for Televerket. Markedet er atskillig mer turbulent enn tidligere, og Televerket er i reell konkurranse på en rekke felter. Derfor stilles det sterke krav til kontroll, ikke minst over den økonomiske utviklingen. Viktige styringsparametere er:

- trafikkutvikling ulike tjenester
- abonnementsutvikling ulike tjenester.

For Televerket basis vil abonnements- og trafikkinntekter for spesielt telefontjenesten og abonnementsinntekter for digital/utleide samband være viktige. For Tele-mobil og TBKnett vil henholdsvis de analoge inntekter på mobile tjenester og datatjenester være viktige.

For å få oversikt over utviklingen av disse styringsparametrene er det helt nødvendig å utarbeide gode prognosemodeller.

For et par år siden ble serviceparametere som kundemeldte feil, leveringstid på abonnement og framkommeligheten på trafikken i nettet prioritert som de viktigste styringsparametrene i Televerket. Televerket har nå gjennom organisasjonsutvikling og modernisering av telenettet god kontroll over utviklingen av disse parametrene.

Det som har meget stor betydning nå, er å følge opp den bedriftsøkonomiske utviklingen. I utgangspunktet lages det årlige prognoser for budsjettet. Disse periodiseres. For de fleste styringsparametrene vil det bli laget sesongbetonte prognoser pr måned eller kvartal som så kan følges opp kontinuerlig av ledelsen.

2.4 Konkurransesvurderinger og prisstrategier

I en strategisk plan er det viktig å ta utgangspunkt i en forventet utvikling. En av de mest fundamentale elementene i en bedrifts utvikling er de framtidige inntektene. Inntektene er avhengig av det volumet som selges og den pris en får på produktet/ tjenesten. I mange tilfeller er det lettere å predikere framtidig prisutvikling. I den situasjon som Televerket er, med kommende konkurranse på en rekke felt, må prisene på de ulike tjenestene tilpasses konkurransesituasjonen. Det betyr at de fleste priser må senkes i forhold til hva de har vært i en enerettssituasjon.

Ut fra forutsetninger om en stadig nedsettelse av prisene, må det så settes opp anslag over forventet etterspørsel etter tjenestene. Her kommer det inn to faktorer: For det første vil etterspørselen etter en tjeneste over en periode følge en utvikling som kan prognostiseres. I tillegg til dette vil også etterspørselen være avhengig av prisen på tjenesten. Dersom prisen blir lavere, blir etterspørselen høyere og omvendt.

Til modellering av hvorledes etterspørselen endrer seg med prisen, benyttes i økonomien begrepet *priselastisitet*. Den er definert som den relative endring i etterspørsel dividert med den relative endring i pris.

Dersom etterspørselen øker med 1 % når prisen går ned med 1 %, er priselastisiteten -1, hvilket vil si at vi får de samme inntekter. Dersom vi ser på prisen på de ulike takstklasser for telefon (lokal-takst, nærtakst, fjernetakst og internasjonale takster), så ligger her priselastisiteten mellom 0 og -1. Dette betyr at etterspørselen øker med mindre enn 1 % når prisen reduseres med 1 %. Dermed taper Televerket på å redusere prisene. Like fullt er det viktig å gjøre dette for å ha priser som er tilpasset den framtidige konkurransen.

Ut fra prognostisert vekst for ulike tjenester kombinert med anslag for ekstra generert trafikk-abonnementsvekst ved prisreduksjoner, utarbeides det strategier

for prissetting. Disse strategiene er basert på klare forutsetninger om at Televerket skal kunne møte den framtidige konkurransesituasjon med lave priser. Det forutsettes at prisene spesielt settes ned på de markedssegmenter der Televerket først møter konkurranse.

2.5 Inntektsprognoser

Televerket basis og andre forretningsenheter lager ved faste perioder inntektsprognoser basert på prognoser for tjenestene:

- Telefon
 - tellerskrittinntekter
 - inntekter samtaleautomater
 - abonnementsinntekter
 - øvrige telefoninntekter
- Mobiltelefon
- Datex
- Datapak
- Digital
- Leide samband
- Teleks
- Alarmtjenesten
- Øvrige tjenester og andre inntekter.

Enkelte tjenester, som telefontjenesten, er klart viktigere enn andre. 75–80 % av Televerkets totale inntekter kommer fra telefontjenesten. Derfor er det også viktig å bruke forholdsvis mer ressurser på å lage gode prognoser for denne tjenesten.

Som tidligere nevnt, lages det i tillegg til de nevnte prognoser også årsprognoser for kommende budsjettperiode for de samme tjenestene.

2.6 Markedsplaner

For Markedsdivisjonene er det svært viktig å ha oversikt over utviklingen på tjenestesiden. Dette kommer også til uttrykk i markedsplaner som lages av divisjonene. En viktig del av markedsplanene er abonnementsprognoser totalt og pr markedssegment for de ulike tjenester. Samtidig lages det også inntektsprognoser for tjenestene basert på både trafikk- og abonnementsinntekter. I markedsplanene ligger det også vurderinger av ikke bare primærtjenestene, men også mindre tjenester som eksempelvis tilleggstjenester på telefon.

2.7 Bemanning/ dimensjoneringsplaner

Et prognostisert aktivitetsnivå for ulike enheter vil være dimensjonerende for bemanningsstyrke. Analogt er det også mulig å lage prognoser for inntekter og anslag for kostnader for å finne fram til en maksimal bemanning. Det var denne framgangsmåten som ble benyttet for å gi anslag for Televerkets totale bemanning i dimensjoneringsprosessen. Et annet eksempel på hvorledes prognoser brukes til bemanningsplanlegging er Opplysningstjenesten. Prognoser for antall tellerskritt (samtaler kombinert med samtalelengde) mot opplysningstjenesten er dimensjonerende for den bemanningen tjenesten kan ha. For Televerkets planfunksjon har nettets størelse kombinert med utviklingen av investeringsnivået vært bestemmende for bemanningsutviklingen. Her er investeringsnivået spesielt avhengig av totalt antall tilknytninger og prognostisert vekst i antall tilknytninger.

2.8 Investeringsbudsjettet

I budsjettprosessen lages det oversikt over den kommende budsjettperiodes investeringsprosjekter. Den største del av disse prosjektene kommer fra nettutbygging. Dette er behov som oppstår i tilknytning til modernisering av nettet, til utvidelse av nettet med eksempelvis mer kapasitet og ved utbygging av nettet for å dekke nye behov / nye abonnenter. Underlag for å finne fram til investeringsprosjektene vil blant annet være prognoser.

2.9 Strategisk planlegging av nettutbyggingen

Det er en rekke faktorer som har betydning for den strategiske planlegging av nettutbyggingen. I denne sammenheng vil vi konsentrere oss om hvorledes prognoser brukes for å støtte opp denne planleggingen.

Først nevnes imidlertid en del vesentlige faktorer som har betydning:

- Innføring av ny teknologi
- Innføring av nettkomponenter som har stor kapasitet og er billige sammenliknet med tidligere
- Innføring av nettløsninger som gir lave driftsutgifter og lave administrative kostnader
- Modernisering av nettet gjennom utskifting av gamle sentraler med nye digitale sentraler

- Oppgradering av tjenestetilbudet i nettet.

Hovedinvesteringene i Televerket har vært:

- Investeringer i nye sentraler
- Investeringer i nettet mellom sentraler (nærnett/fjernnett)
- Investeringer i nettet under sentraler (aksessnett).

I tilknytning til strategiske planer for nettutbygging er dette viktige elementer. Et overordnet mål for de planer som utarbeides for nettutbyggingen, er at de skal være de mest lønnsomme blant ulike alternativer.

Som nevnt skjer det en digitalisering av sentralene i nettet. Det samme gjør det på overføringsveiene i nettet. På høyere nivå i nettet er det nå mye digitalt utstyr. Men på lavt nivå i nettet er det i stor grad analog overføring og dermed analogt utstyr. Det vil etter hvert bli foretatt en modernisering av nærnett, deler av aksessnett, og fjernnett ved innføring av SDH-utstyr, som er neste generasjons transmisjonsutstyr.

Televerket får i 1994 et nytt tjenesteintegreert digitalt nett – ISDN. Her er det en forutsetning at det brukes digital overføring mellom alle abonnenter. Hver abonnent vil få tilgang til en rekke nye tjenester og i tillegg vil en abonnent kunne ha opptil 8 aksesspunkter for overføring av flere tjenester. Dette kalles grunnaksess (2B+D). I tillegg vil det bli tilbudt en utvidet aksess (30B+D) som er beregnet for hussentraler.

De som blir abonnenter på ISDN, vil ikke trenge vanlig telefonabonnement, da dette er en del av de tjenester en får på ISDN. Dette betyr at etter hvert som ISDN tar av, vil vi få en tilsvarende reduksjon av telefonabonnement. Det vil nå bli vist at dette resonnementet har meget stor betydning for de framtidige investeringer.

Som et eksempel ser vi nå på følgende prognoser som ble laget i 1991:

- Prognoser for ISDN. Denne er delt på:
 - prognoser for private abonnement (grunnaksess, 2B+D)
 - prognoser for forretningsabonnement (grunnaksess, 2B+D)
 - prognoser for forretningsabonnement (utvidet aksess, 30B+D)

- Prognoser for vanlig telefonabonnement.

I tillegg ser vi på totalprognosen for telefonabonnement og ISDN. Denne kalles prognosen for telefoni-ekvivalenter. Det gjøres oppmerksom på at de offisielle prognosene har endret seg en del fra 1991 til 1993 – blant annet fordi vi har hatt store forsinkelser med kommersiell introduksjon av ISDN.

Figur 2.1 – 2.3 viser de nevnte prognosene. Av figurene ser vi at totalt antall telefonabonnement etter hvert reduseres etter å ha nådd et toppunkt på midten av 1990-tallet. Dette er under forutsetning av at antall ISDN-abonnement øker slik prognosene angir.

Sammensetningen av disse prognosene gir grunnlag for en rekke strategiske beslutninger:

- Etter at det totale antall telefonabonnement har nådd toppunktet, vil det ikke være nødvendig å investere mer i spesielt sentralutstyr (analogt) for telefonabonnement. Dermed bør en være varsom med å investere i slikt utstyr nå.
- Når en installerer en ny sentral eller utvider kapasiteten på en sentral, skal den dimensjoneres, det vil si bestykkes med nummer, på en slik måte at det ikke er nødvendig å utvide sentralen i løpet av denne planperioden. Denne har vært på 4 år, men er nå i Region Oslo nedkortet til 2 år.
- I stedet for å kjøpe nytt sentralutstyr nå, som kanskje bare kan brukes noen få år, flyttes det på sentralutstyr fra sentraler der det er rikelig med nummer til sentraler der det mangler nummer. Dette øker nummerutnyttelsen ved sentralene.

Disse strategiene er fullstendig basert på prognoser for telefonabonnement og ISDN, og brukes nå av Region Oslo for å minimalisere investeringene og risikoen for overinvestering.

2.10 Nett og nettutvikling

Det har den senere tid vært en rivende utvikling på nettsiden. Vi har nå en rekke nett som er inne i en rask teknisk utvikling. Ser vi på den historiske utvikling, så hadde vi tidligere telegrafnettet og telefonnettet. Nå har vi:

- telenettet (telefoni-nettet)
- teleksnett
- datapaknett

- datexnett
- alarmnett
- mobiltelefonnett
- personsøkernett
- satellittnett.

Televerket har også ansvaret for den landsomfattende kringkastingstjenesten. Samtidig foregår det forsøk med:

- ISDN
- bredbåndnett.

Bredbåndnett er nett med høy overføringskapasitet. Televerket er nå i en fase der ulike nett utprøves.

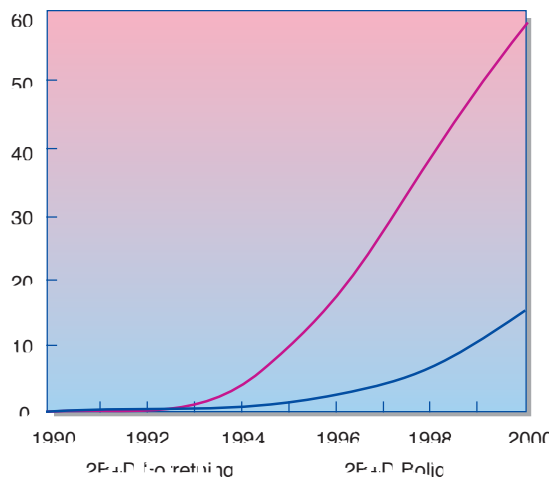
Det som imidlertid er klart, er at på svært lang sikt vil vi ha et landsomfattende bredbåndnett både til bedrifter og til private husstander. Dette vil bety en radikal økning i kapasitet i telenettet og en endring i teknologien. Bredbåndnett er også kalt B-ISDN. Det er for tidlig å fastslå hvilke standarder og løsninger som vi vil få for det framtidige bredbåndnettet. Vi vet imidlertid at de første standarder vil være basert på ATM (Asynchronous Transfer Mode).

Hvorledes vil så dette nettet utvikles og når? Det vil være markedet som styrer denne utviklingen sammen med teknologien. Det vi vet er at det stadig blir billigere å implementere høy kapasitet i nettet; dette reduserer kostnadsdifferansen mellom det nettet vi nå har og et bredbåndnett. Prognosene for utvikling av et bredbåndnett er altså avhengig av den teknologiske utvikling av nettkomponenter og hva de vil koste. Likeledes er det avhengig av de applikasjoner og de tjenester som et slikt nett kan tilby.

Situasjonen nå er at mange private abonnenter har kabelfjernsynsnett som er bredbåndoverføring. Dette er enveisnett – det vil si at det sendes ut kommunikasjon i én retning. I tillegg er nettene stort sett basert på analog teknikk, mens det framtidige bredbåndnettet vil være basert på digital teknikk.

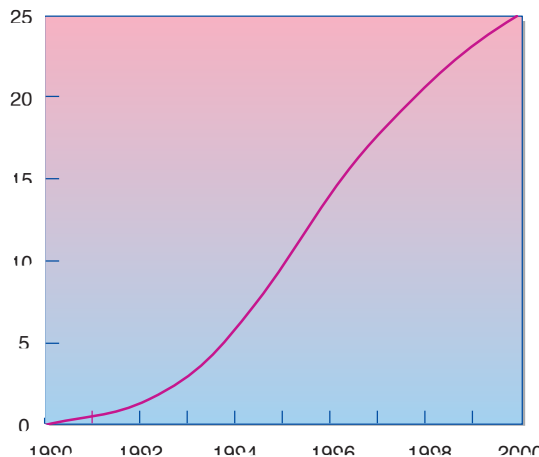
For bedrifter vil det være andre behov som gir etterspørsel etter bredbåndstjenester. Tidligere antok en at overføring av levende bilder – eksempelvis konferansefjernsyn – ville være en viktig faktor for etterspørselen. Det som imidlertid har vist seg å være den klart viktigste faktoren for bredbåndoverføring for bedrifter, er høyhastighets dataoverføring. Det er et stort behov for overføring av

Antall grunnaksesser (x 1000)



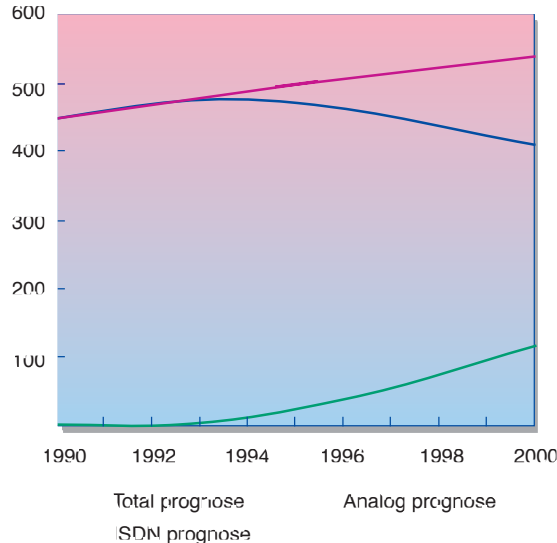
Figur 2.1 Prognose (fra 1991) for antall grunnaksesser ISDN

Antall 30E+D aksesser (x 100)



Figur 2.2 Prognose (fra 1991) for antall utvidede aksesser ISDN

Analoge ekvivalenter (x 1000)



Figur 2.3 Prognose (fra 1991) for totalt analoge tilknytninger, totalt antall digitale tilknytninger samt totalsum telefonabonnement/ISDN

data mellom lokale datanett. Dette er LAN-til-LAN kommunikasjon. Det blir da telenettet som må overføre dataene mellom de lokale datanettene.

Det er svært viktig for Televerket, så tidlig som mulig, å få forståelse for det framtidige bredbåndsbehovet. Dette behovet vil selvsagt være avhengig av pris. Det vil si hva det koster å tilby bredbåndstjenester til abonnenter, hvilke tjenester som tilbys og hva abonnentene er villige til å betale. De prognosene som lages, må ta hensyn til disse faktorene. Siden dette er nye tjenester og det ikke eksisterer historiske data, må det blant annet brukes kvalitative prognosemetoder. I tillegg må prognosene være svært langsiktige.

Denne type prognoser vil gi et underlag for den framtidige nettstrukturen på telenettet. Vi snakker her om telenettet etter år 2000. Det er klart at det er mange usikre faktorer som prognosene vil være basert på; dermed blir prognosene også meget usikre. Etter hvert som vi får ny informasjon, vil imidlertid prognosene bli sikrere.

2.11 Nettstruktur

Det innføres nå en ny teknikk i telenettene. Hovedelementer er digitale nettkomponenter og digitalt svitsjeutstyr. Uten å gå nærmere inn på ulike typer utstyr, kan det slås fast at det digitale utstyret er billigere samtidig som det er atskillig mer driftssikkert. Dette fører da til mindre investeringer ved anskaffelse av utstyr og mindre driftskostnader ved drift av utstyret i forhold til tidligere.

For større prosjekter utarbeides det i Televerket en økonomisk analyse for å finne fram til det beste utbyggingsalternativet. I en slik analyse kan det være mange kostnadsfaktorer som har betydning. Følgende kan nevnes:

- Investeringkostnader
- Driftskostnader
- Administrative kostnader.

Et fundament for disse analysene er prognosene som beskriver den framtidige trafikktutviklingen og abonnementsutviklingen.

I tillegg til kostnadene, må det i den økonomiske analysen også ses på nye muligheter ved en nettutbygging. Den kan muliggjøre et større tjenestetilbud på deler av

nettet eller hele nettet og den kan forårsake en økning i trafikken.

Som vi har vært inne på, skjer det en overgang fra analog til digital overføring i telenettet. Dette betyr at deler av telenettet som nå er digitalt, kan overføre digitale tjenester.

Televerket er nå inne i en periode med stor utskifting av sentraler. De sentraler som nå settes inn i telenettet, er digitale. Samtidig blir transmisjonssystemene som brukes i overføring mellom sentralene også digitale. Dette er imidlertid en utbygging som på landsbasis vil ta mange år.

Telenettet er et hierarkisk nett. Det vil si at det er basert på en trestruktur hvor det er sentraler i grenpunktene. Sentralene svitsjer trafikken i riktig retning avhengig av adressen til anropet.

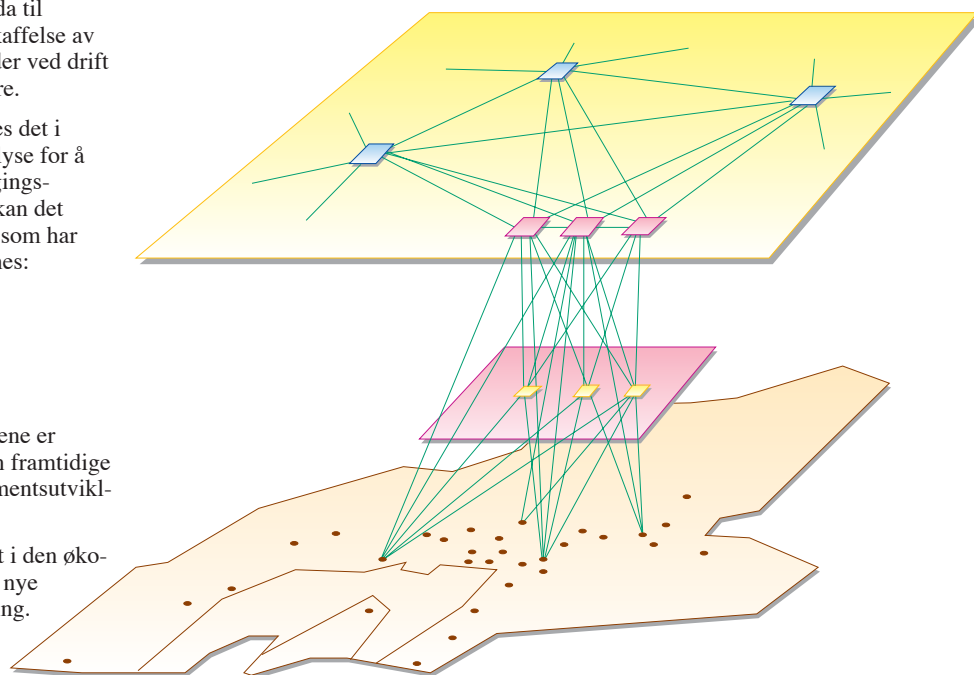
Figur 2.4 viser den planlagte strukturen for Oslo-nettet. Dette er en målnettstruktur som er retningsgivende for den videre utbygging av nettet i Oslo. Figuren viser et hierarki med flere nivåer. På laveste nivå er det et kart over Oslo-regionen hvor endesentralene er angitt som punkter. På neste nivå er det tre gruppesentraler og på nivået over er det fjernsentraler.

For trafikk mellom abonnenter i sentralområdet svitsjes trafikken i ende-

sentralen. For trafikk som går mellom abonnenter i Oslo-regionen som ikke hører til samme sentralområde, går trafikken først opp til endesentralen, deretter til gruppesentralen og så ned igjen. For trafikk som skal ut av Oslo-regionen, må trafikken opp på fjernsentralen før den føres ned igjen.

Den nettstrukturen som er beskrevet, går under navnet sisteveisnettet. Det fins imidlertid, der det er stort trafikkgrunnlag, også tverrveier mellom sentralene. Det betyr eksempelvis at dersom det er stort trafikkgrunnlag mellom endesentral A og endesentral B, så rutes trafikken på en tverrvei direkte mellom A og B i stedet for å bli rutet opp til en gruppesentral og ned igjen. Dersom det er ikke er ledig kapasitet på tverrveien, vil trafikken bli rutet på sisteveisnettet, det vil si fra endesentral opp til gruppesentral og ned igjen.

Av figur 2.4 framgår det at hver endesentral har forbindelse til alle tre gruppesentraler. Dette er gjort ut fra pålitelighetsmessige årsaker. Det er stilt krav om at hovedsakelig all trafikk (80 % av trafikken på den travleste tiden av døgnet og selvsagt desto mer på øvrige tider) skal kunne overføres, selv om en av gruppesentralene ikke fungerer.



Figur 2.4 Planlagt nettstruktur for telenettet i Oslo-regionen

De økonomiske beregninger som ligger til grunn for denne nettstrukturen, er basert på

- det nettet som på forhånd var etablert
- kostnader på ulike nettkomponenter
- driftskostnader på ulike nettkomponenter
- kostnader ved administrasjon av nettet
- prognoser for abonnementsutvikling
- prognoser for trafikkutvikling.

I den økonomiske analysen dimensjoneres også selve nettstrukturen. I tilknytning til dimensjoneringen vurderes:

- antall gruppesentraler
- nødvendig kapasitet på sambandsbunter mellom endesentral og gruppesentral
- antall tverrbunter mellom endesentraler.

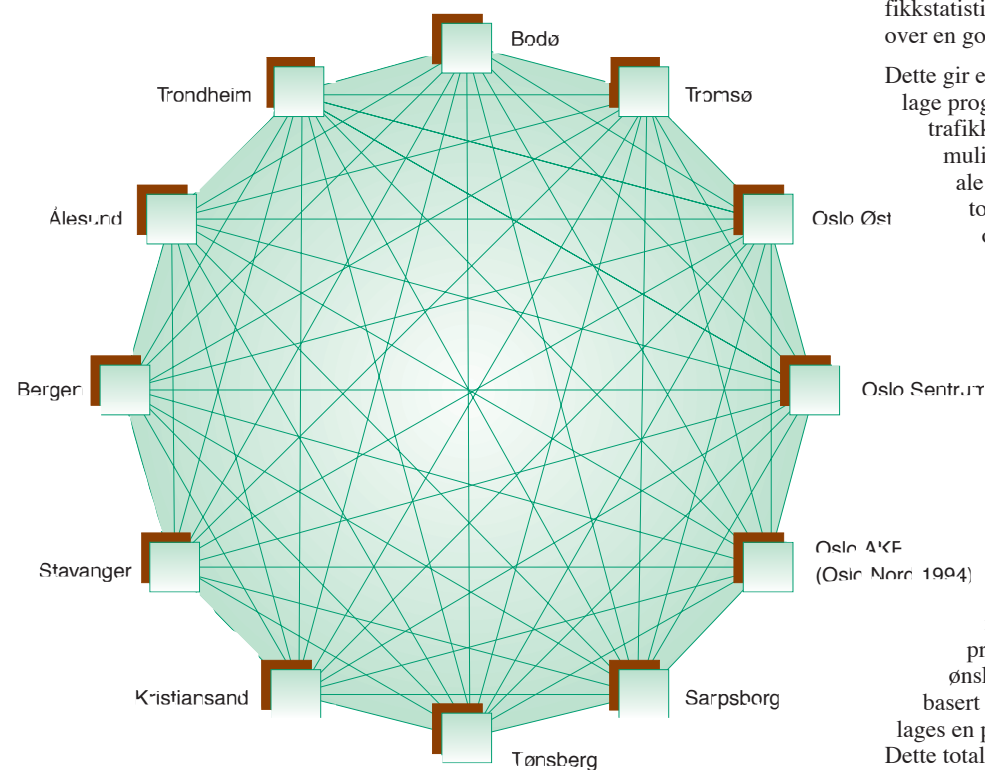
I slike økonomiske analyser ses det på ulike alternativer og de kostnader og eventuelle inntekter som skiller alternativene i løpet av en periode som vanligvis er 10 år. Det foretas en såkalt nåverdianalyse for hvert enkelt utbyggingsalternativ som så viser hvilket alternativ som er mest lønnsomt.

For å kunne foreta denne analysen over en periode på 10 år, er det nødvendig å kjenne til hvorledes nettet må bygges ut i løpet av denne tiden på grunn av økning i antall abonnenter og økning av trafikken. Dette legges inn i analysene gjennom bruk av prognosemodeller og prognoser for henholdsvis abonnement og trafikk.

2.12 Fjernnettet, nærtettet og aksessnettet

Fjernnettet

Fjernnettet er det nettet som går mellom fjernsentralene. Dette er et fullt masknett, hvilket betyr at enhver fjernsentral har sambandsbunter mot enhver annen



† Utvidelser

1992: Bergen

1993: Oslo I1 Oslo I12

Figur 2.5 Struktur for fjernnettet

fjernsentral. På sikt vil det være mulig å overføre trafikken mellom fjernsentralene ved en avansert dirigeringsform som er kalt TAR. Dersom et anrop fra sentral A til sentral B ikke kan overføres direkte på grunn av mangel på ledig kapasitet, vil en forsøke å overføre anropet via sentral C. Dersom dette ikke går, prøves overføring via sentral D, etc. I tillegg skal det legges inn en funksjon som kalles sambandsreservering. Denne funksjonen vil prioritere anrop som er forholdsvis langt i oppsettingsfasen. Dette betyr at anrop som er kommet opp til fjernsentralen og skal rutes ned igjen, blir prioritert framfor anrop som ikke er kommet så langt. Slike funksjoner som dette fører til en mer effektiv trafikkflyt og de skal brukes både i fjernnettet og eksempelvis i nærtettet i Oslo.

I figur 2.5 er det vist hvorledes strukturen på fjernnettet er.

Trafikken som går i fjernnettet, registreres i en trafikkmatrise der inngående og utgående trafikk til de ulike fjernsentralene angis. Disse trafikkmålingene foretas hvert kvartal slik at det foreligger en trafikkstatistikk med 4 trafikkmatriser i året over en god del år.

Dette gir et meget godt grunnlag for å lage prognoser for den framtidige trafikken i fjernnettet. Her er det mulig å lage prognoser for den totale trafikken i fjernnettet ut fra totaltrafikken i trafikkmatrisen, og likeledes den totale trafikken ut fra de ulike sentraler ved å summere radene i matrisen og den totale trafikken inn til de ulike sentraler ved å summere kolonnene i matrisen. I tillegg lages det prognoser for trafikken mellom de enkelte sentraler ved å ta utgangspunkt i matrise-elementene i trafikkmatrisen.

I et slikt system med trafikkprognoser kan det også foretas mange korrigeringer av prognosene dersom dette er ønskelig. Eksempelvis kan det, basert på totaltallet for fjerntrafikken, lages en prognose for totaltrafikken. Dette totaltallet kan være forskjellig fra summen av prognosene for enkelttrafikken i trafikkmatrisen. Ut fra dette kan det foretas justeringer av prognosene for enkelttrafikken ved å veie de enkelte prognosene ut fra hvor usikre de er. Dermed fås et sett med korrigerte prognoser.

De trafikkprognosene som lages, gir grunnlag for

- dimensjonering av fjernsentralene
- dimensjonering av kapasiteten i nettet
- planlegging av nettstrukturen.

Dimensjonering av kapasiteten i nettet gir underlag for bestilling av transmisjonsutstyr og legging av kabler i nettet. Det lages prognoser for det transmisjonsutstyret som skal bestilles de nærmeste 5 årene. Dette kalles disponeringsplan.

Nærnett

Prinsippene for utbygging av nærnettet er i hovedtrekk de samme som for utbygging av fjernnettet. Det er imidlertid mulig å gå mer detaljert fram.

Vi har tidligere vist at Televerket har en rekke forskjellige tjenestenett. Det som er svært viktig å være klar over, er at når Televerket bygger ut transmisjonsveier i nærnettet, så skal disse ha kapasitet til å dekke trafikken i de ulike tjenestenettene. For å tydeliggjøre dette, innføres nå et nytt begrep – *transportnettet* – istedenfor nærnettet. Som underlag for å dimensjonere dette nettet, må det lages *prognoser* for

- trafikk i telefonnettet
- trafikk i datapaknettet
- trafikk i datexnettet
- trafikk i mobiltelefonnettet (NMT 450 og NMT 900)
- trafikk i det nye mobiltelefonnettet (GSM) for Tele-mobil og NetCom
- trafikk i personsøkernettet
- trafikk i ISDN
- trafikk med bredbåndsoverføring
- antall digitale abonnement
- antall samband til NRK
- diverse entreprisere spørsmål, eksempelvis for kabel-TV-selskaper.

De prognosene som lages her, er alle på enkelte strekninger i transportnettet. Anta at det er behov for bredbånd- eller høyhastighetsoverføring mellom to bedrifter A og B i Oslo-området for overføring mellom de lokale datanett i bedriftene. For å få overført denne informasjonen er det nødvendig å bygge ut en relativt høy kapasitet på ett strekninger for å komme fra bedrift A til bedrift B. Det må altså lages prognoser for denne type trafikk på alle delstrekninger i transport-

nettet. På samme måte må det lages trafikkprognoser for de øvrige tjenestenettene.

På grunnlag av disse trafikkprognosene, kan så kapasiteten på de enkelte strekninger i transportnettet beregnes. Dermed beregnes også strukturen på nettet. Strukturen vil i hovedsak være basert på strukturen i telenettet, da den dominerende delen av trafikken går i dette nettet. På grunnlag av trafikkprognosene bestemmes videre antall optiske fibrer som legges ned på de enkelte strekninger. Nedlegging av fibrer vil alltid være basert på svært langsiktige prognoser. Som for fjernnettet, planlegges så det transmisjonsutstyret som skal brukes de neste 5 år. Det er dette som kalles disponeringsplan.

I strukturen i transportnettet er det også innebygd en sikkerhet. I kapittel 2.11 er det vist hvorledes nettstrukturen i Oslo-regionen er. Der framgår det at det er forbindelse mellom hver endesentral og hver gruppesentral. Dette skulle en tro betyr svært stor sikkerhet i nærnettet. Det er imidlertid en kjensgjerning at svært mange av disse overføringsveiene går i de samme fysiske kanalene. Dette betyr at dersom en kanal kuttes eksempelvis på grunn av graving, vil alle overføringsveier i denne kanalen lammes. Derfor er det laget en reserveveisplan i Oslo-regionen. Denne planen, som nå er gjennomført, ivaretar påliteligheten i nettet ut fra kravet om at enhver overføringsvei mellom to endesentraler i nærnettet alltid skal ha to uavhengige fysiske forbindelser (kanaler). Nettet er også dimensjonert slik at det har like stor kapasitet på reserveveiene som på de ordinære veiene.

Aksessnettet

Den nedre delen av nettet kalles aksessnettet eller lokalnettet. Dette nettet har til nå blitt bygd ut som en trestruktur. I den senere tid har det blitt utbygd utskutte abonnenttrinn. Dette betyr populært sagt at en del av sentralen er plassert desentralt, slik at svitsjing og konsentrasjon av trafikken kan gjøres lengre ute i nettet. Med svitsjing menes mulighet for å sende trafikken eller samtalen ulike veier avhengig av adressen på trafikken (samtalen). Med konsentrasjon av trafikken menes et prinsipp som er gjeldende under dimensjoneringen av hele nettet: Nettet dimensjoneres ut fra forventet antall for samtaler som går der, med en minimal mulighet for sperr. Når de enkelte linjer fra abonnentene (abonnentlinjer) kommer opp til en konsentrator

eller et utskutt abonnenttrinn, vil disse nettkomponentene videre oppover i nettet ha en lavere kapasitet enn summen av kapasiteten av abonnentlinjene, fordi det ikke kan regnes med at alle abonnentene ringer samtidig.

Lengre ned i nettet vil vi ikke ha mer konsentrasjon eller svitsjing av trafikken. Det vil imidlertid fortsatt være en trestruktur. Det har vist seg naturlig å legge mange av abonnentkablene i samme grøfter eller kanaler på vei mot den enkelte abonnent. Istedenfor å bruke en stor mengde enkeltkabler brukes en koaksialkabel eller en parkabel som legges ned i grøften eller kanalen. På spesielle punkter der det er hensiktsmessig deles så kablet opp ved bruk av en hovedfordeler til mindre kabler som går i ulike retninger. Det samme gjøres så lengre ned i nettet ved bruk av en endefordeler.

Det er i ferd med å bli lagt en ny infrastruktur i aksessnettet i tettbygde områder. Istedenfor trestruktur fra endesentral og nedover i nettet, vil det i mange tettbygde områder bli lagt fiberkabler i ringstruktur med endesentral, utskutte abonnenttrinn og transmisjonsutstyr i enkelte punkter i ringen. Fra hvert av disse punktene, som er kalt tjenestetilknytningspunkt, går det så en trestruktur ut mot abonnentene. Fordelen med en fiberring er mange. For det første føres det større kapasitet lengre ned i nettet, med tilhørende muligheter for desentral svitsjing og trafikk-konsentrasjon. Dermed blir det større pålitelighet fordi trafikken kan føres to uavhengige veier til endesentralen. Ny teknikk benyttes lengre ned i nettet, noe som fører til at det er lettere å tilby overføring av bredbåndstrafikk til og fra større kunder.

Vi har allerede omtalt trafikk-konsentrasjon. Uten å gå detaljert inn på dimensjoneringen, så dimensjoneres de enkelte strekninger etter statistiske prinsipper. Input til dimensjoneringen er hvor mange abonnenter som vil være lokalisert i de enkelte områdene og de behov de enkelte abonnentene har. Siden lokalnettet ikke utvides hver gang det skjer tilvekst av kunder, må det lages prognoser for både abonnenttilveksten og trafikktilveksten i løpet av en periode. Det er disse prognosene som brukes som underlag for aksessnettutbyggingen.

For hvert enkelt sentralområde lages det prognoser for abonnenttilveksten. Her skilles det mellom privat abonnement og forretningsabonnement, som er abonne-

ment fra bedrifter. I tillegg er det viktig å skille ulike forretningsabonnement – spesielt der det er hussentraler med ulike antall linjer.

Det er den senere tid blitt vanskeligere å lage abonnementsprognoser. Dette skyldes flere faktorer. Spesielt kan nevnes at det på grunn av den økonomiske situasjonen skjer svært mange oppsigelser av abonnement. Dette er oppsigelser både fra abonnenten selv og oppsigelser av Televerket fordi abonnenten ikke har betalt. Dette gjør det vanskeligere å lage prognoser. Det må derfor lages prognoser for brutto tilgang på abonnement og på avgang av abonnement for å finne fram til prognoser for netto abonnementsstilgang. For mindre sentralområder vil bevegelsen i abonnement også ha betydning – ikke bare netto-tilveksten. Med andre ord er det viktig å kjenne til hvilke abonnement som sies opp og hvilke som kommer som tilvekst.

Abonentene har også ulikt trafikk-mønster. Store hussentraler vil pr linje generere mer trafikk enn private abonnenter. Dette må det tas hensyn til i trafikkprognosene. I tillegg må det lages prognoser for spesiell høyhastighets-trafikk fra ulike forretningsabonnenter. Stort sett lages 2 Mbit/s prognoser for denne trafikken. Summen av prognosene for 2 Mbit/s og for den ordinære trafikken gir da underlag for dimensjonering av aksessnettet. I tillegg vil det bli tatt med spesielle bredbåndsbehov; dette gjelder spesielt i fiberringene.

2.13 EUs innkjøpsdirektiver

Inngåelse av EØS-avtalen vil innebære endring i de offentlige innkjøpsregler. Dette vil føre til større konkurranse og lavere priser for Televerket på større anbud. EUs direktiver vil åpne proteksjonistiske markeder i Europa og gi norsk industri nye muligheter. På den annen side vil hjemmemarkedet åpnes for en mer omfattende konkurranse, hvilket kan være en ulempe.

For telekommunikasjonsutstyr er det en terskelverdi på 600 000 ECU (ca 5 mill kr). For anleggssektoren er terskelverdien ca 40 mill kr. Alle offentlige innkjøp med pris over de gitte terskelverdier skal legges ut på anbud i databasen Tender Electronic Daily.

Det stilles også krav til forhåndskunngjøring i Europa om kommende anskaffelser. Det skal lages 12 måneders prognoser på volumet av utstyr som skal anskaffes. Dette betinger et utvidet prog-

noseopplegg fra fagenheter, regioner og divisjoner i Telegruppen.

3 Kvalitative og kvantitative prognosemodeller

3.1 Bruk av kvalitative og kvantitative prognosemodeller

I denne artikkelen legges hovedvekten på de *kvantitative prognosemodeller*. Det er prognosemodeller som tar utgangspunkt i historiske data i form av observasjoner som er observert over et gitt tidsrom. Ut fra disse observasjonene kan det så ved matematiske/statistiske modeller (kvantitative modeller) lages framskrivninger for utviklingen. Dette blir da prognosene.

Det finnes en rekke ulike kvantitative prognosemodeller som kan benyttes. Av de mest kjente modeller nevnes:

- Regresjonsmodeller
- Glattingsmodeller (Brown, Holt og Holt-Winters modeller)
- Tidsrekkemodeller
- Kalmanfiltermodeller
- Metningsmodeller.

I flere av modellene kan det tas med forklaringsvariable. Disse modeller er nærmere beskrevet senere i artikkelen.

I en del tilfeller mangler en data eller observasjoner, noe som gjør at det ikke er mulig å bruke kvantitative prognosemodeller. Det kan eksempelvis være i de tilfeller at en ny tjeneste introduseres. Da vil det ikke være mulig å benytte kvantitative prognosemodeller, men isteden kan mer subjektive metoder benyttes.

Prognosemodeller som hovedsakelig er basert på subjektive vurderinger kalles *kvalitative prognosemodeller*. Det finnes flere typer kvalitative prognosemodeller. Metodene har et analytisk utgangspunkt, selv om datagrunnlaget er spinkelt. De mest kjente kvalitative prognosemodeller er:

- Ekspertundersøkelser
- Delphimetoden
- Scenariometoden
- Analogimetoden
- Markedsundersøkelser.

Det vil i det følgende bli gitt en kort beskrivelse av disse metodene.

3.2 Ekspertundersøkelser

Når det ikke eksisterer data, utviklingstrender eller annen kvantitativ informasjon som underlag for prognosene, vil en prognoseansvarlig selv ha ansvaret for å foreta en subjektiv vurdering for å få fram resultater. I en slik situasjon vil vedkommende i mange tilfeller ønske bistand fra andre eksperter for å få deres synspunkter. Dette kan gjøres mer eller mindre tilfeldig ved at det tas kontakt i de nærmeste omgivelser, eller det kan gjøres mer systematisk.

Dersom en prognoseansvarlig plukker ut et sett med eksperter og han systematisk framlegger problemstillingen på samme måte, gjerne med et sett gjennomarbeidede spørreskjemaer, så kalles dette en ekspertundersøkelse. Ekspertenes svar vil dermed være selve prognosegrunnlaget.

3.3 Delphimetoden

Delphimetoden er en raffinering av ekspertundersøkelse. Navnet Delphi kommer fra oraklet i Delphi som er kjent fra oldtidens Hellas, der oraklet ble oppsøkt og hvor det gav spådommer som var mer eller mindre vel funderte.

I Delphimetoden velges det ut et større sett med eksperter. En del av ekspertene skal være spesialister innenfor det aktuelle fagområdet, men mange eksperter kan være spesialister innen fagområder som ikke representerer det området som skal analyseres. I undersøkelser som Televerket har gjennomført for å kartlegge etterspørselen etter nye teletjenester og nye applikasjoner, har en del av ekspertene blitt valgt fra Televerket, men ikke en ubetydelig del av ekspertene har kommet fra andre områder.

Delphimetoden går ut på å gjøre et grundig forarbeid før kartleggingen foretas. Det må klargjøres hvilke forutsetninger som er til stede. Skal det lages prognoser om utviklingen av en spesiell teletjeneste, bør denne beskrives på best mulig måte, både med hensyn til teknisk kvalitet, med hensyn til utforming og design og med hensyn til mulige applikasjoner. Om mulig bør det også gjøres vurderinger med hensyn til prisintervall.

Deretter kan det utformes et spørreskjema til bruk i Delphiundersøkelsen. Dette spørreskjemaet bør være godt gjennomarbeidet. Aktuelle spørsmål i en Delphiundersøkelse er forventet introduksjonstidspunkt for en ny tjeneste og forventet utvikling av etterspørselen etter tjenesten. I figur 3.1 er det som eksempel

vist en illustrasjon av fjernarbeid fra et spørreskjema i en undersøkelse som Televerket gjennomførte i 1976 [1]. I tillegg er det gitt en verbal beskrivelse av tjenesten samt eksempler på anvendelser og tekniske forutsetninger. I figur 3.2 er det vist en illustrasjon av nøyaktig den samme tjenesten fra et spørreskjema i en Delphiundersøkelse som nå (1993–1994) gjennomføres i RACE-prosjektet TITAN i 10 ulike europeiske land. Fjernarbeid er kun en av en rekke tjenester som ble kartlagt i disse undersøkelsene. I [54] er det vist hvorledes spørreskjema er utformet i de to undersøkelsene.

For hver tjeneste som er beskrevet, skal ekspertene svare på de angitte spørsmål. Dersom det er spesielle kommentarer eller forutsetninger som eksperten har i tilknytning til spørsmålene, skal vedkommende angi dette i kommentarfeltet.

Svarene fra undersøkelsen sendes så inn. Dette er da avslutningen på første runde i en Delphiundersøkelse, som vanligvis gjennomføres i to eller tre runder. Etter første runde behandles så de respektive spørreskjemaene. Det lages tabeller over svarene, og det beregnes medianer (50 %-fraktilen) og intervaller rundt denne for hvert enkelt spørsmål. Dette materialet sendes så tilbake til hver enkelt deltaker i Delphiundersøkelsen. I tillegg sendes det et utvalg med kommentarer. Dette er kommentarer fra ekspertene som de som arrangerer Delphiundersøkelsen mener er viktig informasjon og forutsetninger som kan ha betydning for synspunkter på tjenesten.

Runde nummer to i Delphiundersøkelsen går ut på å gjennomgå resultatene fra undersøkelsen så langt basert på det som sendes tilbake til hver enkelt ekspert. *Deretter skal ekspertene svare på de samme spørsmålene på ny.* Poenget med andre runde i Delphiundersøkelsen er å få ekspertene til å iterere seg fram til samme synspunkt basert på kommentarer og tidligere svar fra andre eksperter.

En Delphiundersøkelse gjennomføres vanligvis i to eller tre runder. Dette bestemmes av de som arrangerer undersøkelsen ut fra en vurdering om hvor store endringer som forventes fra en runde til neste.

Etter at siste runde i Delphiundersøkelsen er ferdig gjennomført, foretas nødvendige tabellutkjøringer og analyser som vil angi prognoser for de ulike tjenestene. Prognosene angis vanligvis i en Delphiundersøkelse ved medianen (50 % kvartilen) som er midtpunktet i observasjon-

ene eller i svarene og ved et usikkerhetsintervall som vanligvis utgjøres av 25 % persentilen og 75 % persentilen. Dette intervallet utgjør da 50 % av observasjonene eller svarene. I undersøkelsen fra 1976 var medianen for introduksjon av fjernarbeid 1990, mens usikkerhetsintervallet her var 1985–1995.

3.4 Scenariometoden

Det kan diskuteres om scenariometoden er en ren prognosemetode. Metoden går ikke ut på å lage enkeltprognoser, men det er en framstilling av flere variable som gjensidig varierer og som kan påvirke hverandre. Det er i og for seg en konstruktiv metode som kan benyttes for å se og vurdere nærmere årsakssammenhenger og gjensidig utvikling.

Metoden går ut på å ta utgangspunkt i et spesielt år, eksempelvis år 2005. Det skal så lages en forventet beskrivelse av situasjonen dette året. Hvis vi tar utgangspunkt i Televerket, vil det være naturlig å se på Televerket i en konkurransesituasjon dette året. Her åpnes det selvsagt for ulike situasjonsbilder, og dette er en av fordelene med scenariometoden. Det er mulig å foreta ulike beskrivelser avhengig av de forutsetninger som legges til grunn. Deretter går metoden ut på å gi en beskrivelse av utviklingen til de tilhørende variable i perioden fram til år 2005.

Altså gitt situasjonen i år 2005, hva slags utvikling har vi hatt fra nå og fram til år 2005 for å få den beskrevne utvikling i år 2005. Beskrivelse av denne utviklingen blir da prognoser for de angitte variable. Metoden kan i første omgang høres noe banal ut, men det er den ikke. Dette er en måte mange tenker på. Med andre ord, ut fra tanker og forutsetninger om framtiden tenker vi oss hvorledes vi skal komme til dette framtidspunktet. I scenariometoden er dette satt mer i system. Det er ikke snakk om tanker eller drømmebilder om en gitt framtid, men en reell beskrivelse av dem. Og deretter lages det også en realistisk beskrivelse av utviklingen mot dette framtidspunktet med de konsekvenser dette måtte ha.

3.5 Analogimetoden

Denne metoden kan også brukes som prognosemetode dersom en har manglende data. Det er imidlertid en vanskelig metode å bruke, og det bør nok vurderes meget nøye om det er noe å hente på å bruke metoden.

Analogimetoden går ut på å anta at utvikling for ulike produkter er sammenliknbar eller at utviklingen av like pro-

dukter i ulike områder, gjerne ulike land, kan være lik, men i ulike tidsintervaller.

Dersom det skal lages prognoser for utviklingen av en ny teletjeneste, kan det være naturlig å se på hvorledes utviklingen har vært i et land der denne tjenesten allerede har blitt introdusert. ISDN har eksempelvis nå blitt kommersielt tilbudt i Frankrike en periode, mens det i Norge ennå er en prøvetjeneste. Ved å se på utviklingen i Frankrike kan en så anslå en analog utvikling i Norge, alle forhold tatt i betraktning. En slik analogi vil i mange sammenhenger være svært farlig, fordi forholdene i de to land som sammenliknes kan være vidt forskjellige. Derfor er det ikke udelte lett å benytte en slik metode.

I utviklingsland har metoden blitt brukt noe ved å se på korrelasjon mellom ulike produkter. Eksempelvis kan det være aktuelt å se på etterspørselen etter kjøleskap for å vurdere etterspørselen etter telefonabonnement. Etterspørselen etter telefonabonnement vil være begrenset ut fra dekningsgrad i teleinfrastruktur, mens kjøleskap er avhengig av dekningsgrad i elverksinfrastruktur, som er lettere å få fram. Utviklingen i etterspørsel etter kjøleskap kan derfor fortelle noe om forventet potensial i etterspørsel etter telefonabonnement. I hvilken grad dette kan tilfredsstilles er selvsagt helt avhengig av pris på abonnement og av utbygningsmidler.

3.6 Markedsundersøkelser

Markedsundersøkelser kan også brukes som et underlag for prognoser. Markedsundersøkelser kan brukes både i tilknytning til kvantitative og kvalitative prognosemodeller. En måte som markedsundersøkelser kan benyttes på, er å kartlegge potensialet for bruken av en tjeneste.

Ved å foreta en representativ kartlegging av møtevirksomhet i norske bedrifter, kan det eksempelvis gjøres anslag for totalt antall møter i norske bedrifter. Deretter kan møtene sorteres ut fra antall personer på møtet, reisekostnader allokert til møtet og hva som foregikk på møtet, for å finne fram til et potensial for bildeoverføring – det vil si volum for bildetelefon og volum for konferansefjernsyn.

Ved å sammenlikne kostnadsutvikling for disse tjenestene basert på fallende pris på nettkomponenter og reisekostnadene som ble beregnet til de respektive møtene, er det mulig å lage prognoser for utvikling av tjenestene. Her må det

Tele-hjemmearbeid

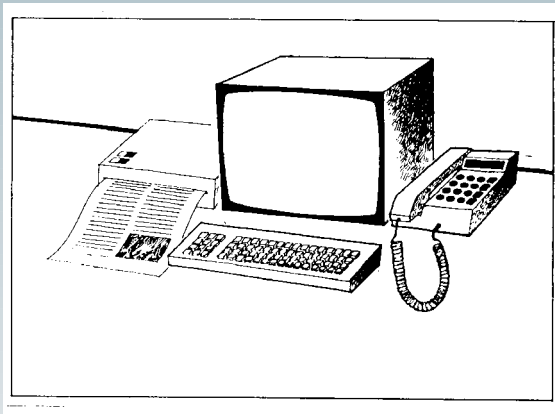
Kommunikasjon mellom arbeidsplass og hjemmet som hel eller delvis erstatning for fysisk oppmøte på arbeidsstedet. Kommunikasjonen kan foregå ved tekst, lyd og stillestående bilde.

Eksempler på anvendelser:

- Kan gjelde arbeidsaktiviteter for ulike personalgrupper
- lyd- (eventuelt bilde-) kommunikasjon som erstatning for vanlig møtevirksomhet
 - fjernskrivning etter diktat, puncharbeid, etc.
 - tilgang til bedriftens registre, datamaskin, etc.

Tekniske forutsetninger:

Det forutsettes at overføring av levende bilder holdes utenfor vurderingen, slik at kommunikasjonen kan foregå over telefonnettet.



Figur 3.1 Eksempel på beskrivelse av tjenesten fjernarbeid i en Delphiundersøkelse fra 1976

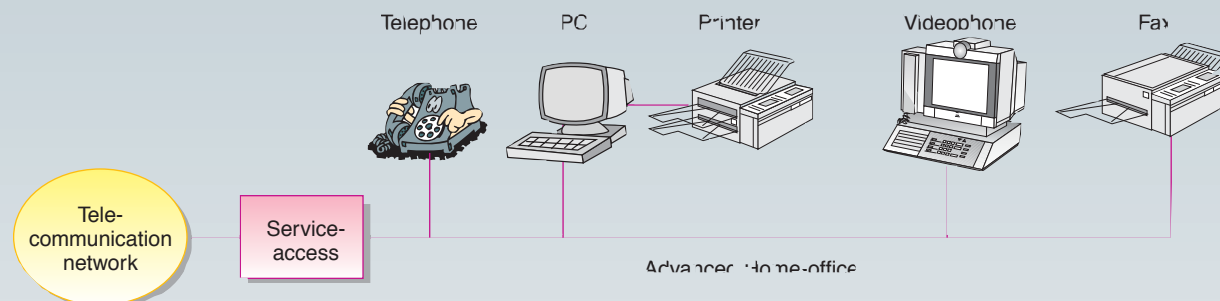
Application 4B: Home-office (advanced)

General description:

This is very similar to application 4A (Simple Home-office), but we have added some new communication features – extended multimedia possibilities on the PC, a videophone, and a fax. This will increase the possibilities for advanced and/or interactive communications. One of the reasons for going for such an advanced solution might be the type and characteristic of work to be done. The complexity of the work may be an example of such a characteristic.

Technical assumptions:

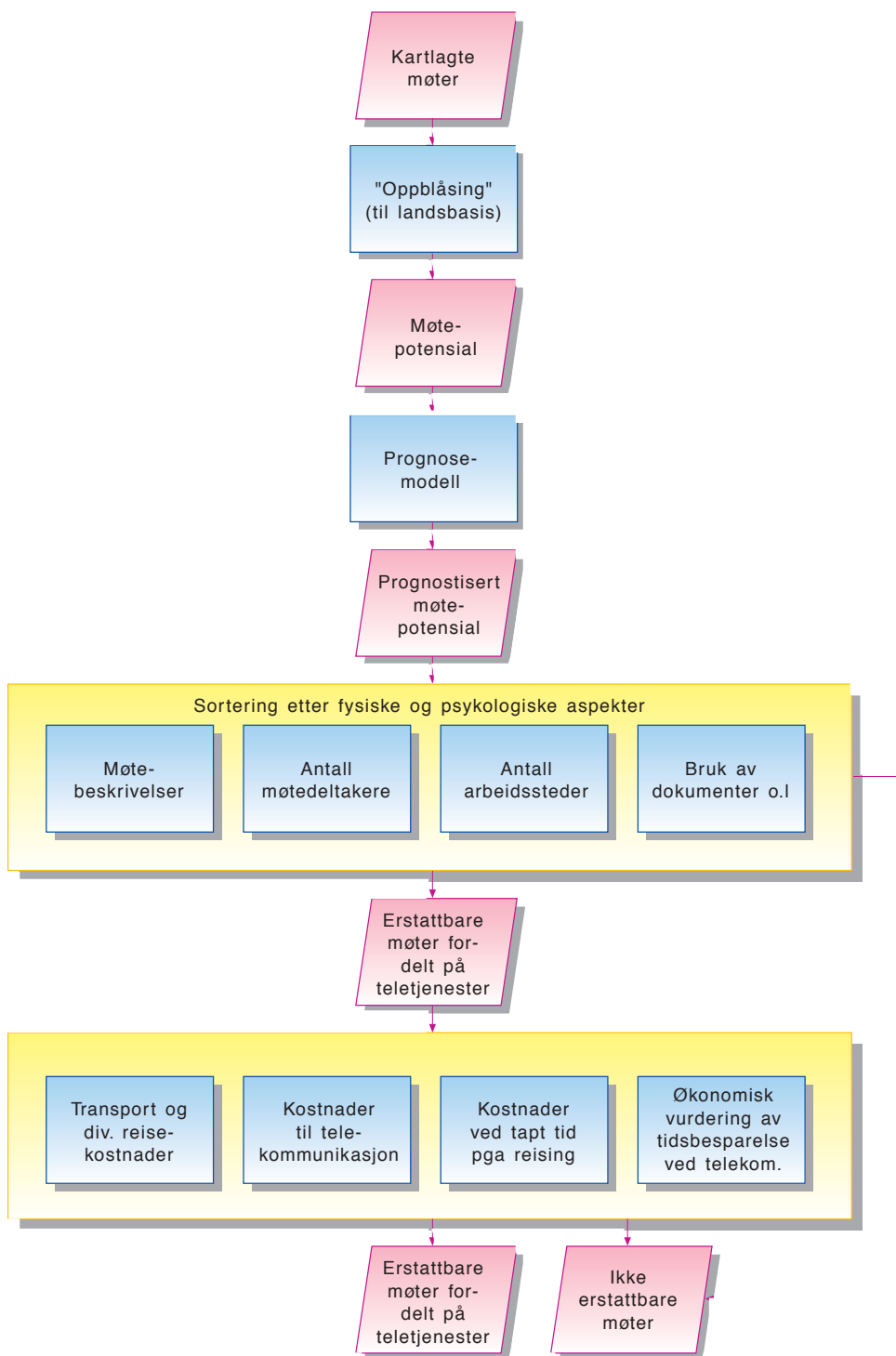
This office has a more advanced multimedia-PC with necessary hardware and software also to support video-presentations. It has in addition a videophone and a fax. The added features will probably require higher capacity from the teleservice. The household must subscribe to either a wideband or a broadband teleservice for support of these applications. The broadband-service will give faster transmission and better image qualities, but both the user-equipment and the teleservice will be more expensive. The wideband service is also of higher capacity compared to the one used in the simple home office.



Examples of use:

- Enhances possibilities for collaborative work and decision making because of the added multimedia features.
- Different kinds of consultancy work.

Figur 3.2 Eksempel på beskrivelse av tjenesten fjernarbeid i en Delphiundersøkelse fra 1993



Figur 3.3 Eksempel på hvorledes etterspørsel etter bildetelefon og konferansefjernsyn kan avledes ut fra møtevirkosomhet i norske bedrifter

imidlertid foretas justeringer, da ikke alle vil opptre rasjonelt og avholde et telemøte når dette lønner seg økonomisk og forutsatt at det som foregår på møtet er av en slik karakter at det ikke er nød-

vendig med personlig frammøte. Figur 3.3 viser prinsippet i en slik sortering av kartlagte møter [2].

I tillegg må en være klar over at nye tjenester også åpner mulighet for nye applikasjoner som vil komme.

Det har i lang tid vært store problemer med å kunne gi innvalgsserier til bedrifter på grunn av manglende nummer. Ved innføring av ny nummerplan har Televerket fått en større etterspørsel etter innvalgsnummer som nå kan tilfredsstilles. Etterspørsel etter innvalg er avhengig av husstratetype og eventuell utskifting, kostnader ved anskaffelse samt også besparelse av tid og kostnader for ekspedient. Markedsundersøkelser blant bedrifter vil kunne bidra til å kartlegge disse forholdene samt gi underlag til prognoser. I tillegg vil selvsagt de registrerte data i form av historiske data om etterspørsel etter innvalgsnummer være viktig underlagsmateriale for prognosene.

Generelt kan det sies at underlag fra markedsundersøkelser ofte er et supplement til prognoseunderlaget.

Et annet eksempel på bruk av markedsundersøkelse til underlag for prognoser, er å stille et sett av prisspørsmål for å få estimert en etterspørselskurve. Figur 3.4 viser et av resultatene fra en markedsundersøkelse som i 1976 ble gjort blant de første abonnentene i Norge som i et pilotprosjekt fikk tildelt knappsatsapparater og tilleggstjenester [3]. Etter at de hadde hatt tjenestene en periode ble de spurt om hva de var villige til å betale for å bruke tjenestene. Ut fra dette kunne det så settes opp en etterspørselskurve basert på svarene. I figur 3.4 er det vist hva abonnentene på den tiden og med den tids kroneverdi var villig til å betale for å ha et knappsatsapparat. Ut fra forutsetninger om prisendringer kunne det her blitt laget en prognose for anskaffelse av knappsatsapparat.

I den Delphiundersøkelsen som nå gjennomføres i RACE-prosjektet TITAN settes det også opp spørsmål for hver tjeneste som skal brukes til å anslå en etterspørselskurve. Dette ses i spørreskjemaet i figur 3.2. Her er det imidlertid ekspertene som vil angi forventet etterspørsel og ikke abonnenter som har brukt tjenesten.

Bruk av offentlige statistikker vil også sammen med markedsundersøkelser være til god hjelp i potensialberegninger. Dette vil bli nærmere omtalt i kapittel 8, der metningsmodeller introduseres som en av de kvantitative prognosemodellene. En viktig parameter som inngår i denne modellen er markedspotensialet for tjenesten. Dersom en har få eller ingen

observasjoner, fordi det er en ny tjeneste, må markedspotensialet anslås før det lages prognoser.

Når det gjelder bruk av markedsundersøkelser generelt skal en være klar over at dette er beheftet med usikkerhet. Resultatene her skal være representative for hele populasjonen (f.eks. norske husstander), men de som kartlegges er et utsnitt eller et utvalg av populasjonen. Usikkerheten i resultatene avtar med utvalgets størrelse og går mot 0 når størrelsen på utvalget blir lik hele populasjonen.

Anta at vi skal anslå usikkerheten i et svarresultat på et spørsmål der andelen som har valgt et gitt alternativ er p og der antallet i utvalget er n . Dersom populasjonen er stor i forhold til antallet i utvalget, vil en nedre grense for usikkerheten være gitt ved variansen som er:

$$p(1 - p)/n$$

Dette forutsetter at de som er i utvalget er trukket ut på en tilfeldig (representativ) måte. Kvadratrotten av det uttrykket som står over, er lik standardavviket. Vanligvis vil 2 ganger standardavviket til hver side for den estimerte p -verdien være et tilnærmet 95 % konfidensintervall. Dette intervallet sier noe om den statistiske sikkerheten i markedsundersøkelsen.

For å konkretisere dette ser vi på et eksempel. Anta at vi har foretatt en markedsundersøkelse og at utvalgets størrelse er $n = 900$. Vi har stilt et spørsmål der de spurte (respondentene) skulle angi hvilket alternativ de foretrakk. Det kan eksempelvis være et prisspørsmål for en tjeneste, der det er satt opp ulike prisalternativer eller det kunne være, slik vi kjenner igjen fra avisene, et spørsmål om å angi hvilket parti en vil stemme på. Anta at på ett av svaralternativene har vi estimert at $p = 0,10$. Det vil si at 10 % av de spurte har valgt dette alternativet. Hva er da den statistiske usikkerheten i dette resultatet? Standardavviket blir da 0,01; det vil si 1 %. Konfidensintervallet som er 2 ganger standardavviket til hver side for p blir da (8,0 % – 12,0 %). Dette viser at resultater selv fra representative markedsundersøkelser må behandles med varsomhet.

4 Dataunderlag og bearbeiding av data

4.1 Hvorfor er det viktig å bearbeide underlagsdata?

Den mest arbeidskrevende delen av prognosearbeidet er anskaffelse og bearbeiding av data. Anslagsvis 80 % av det totale prognosearbeidet går med til selve datahåndteringen. For å kunne lage gode prognoser er det av helt avgjørende betydning at datagrunnlaget er godt. Det betyr at dataene skal være representative, at de skal være konsistente, at de skal være riktige og at de skal være på en hensiktsmessig form.

Unormale data som er framkommet enten ut fra feil i dataene eller ut fra spesielle hendelser som ikke er representative for den generelle utviklingen, må det justeres for.

Det er også viktig å definere de ulike dataene på hensiktsmessig måte og deretter endre/vedlikeholde definisjonene i takt med utviklingen. Nye definisjoner kan eksempelvis være nødvendig i tilknytning til innføring av nye tjenester.

En modell er aldri bedre enn de data som benyttes. Er dataene dårlige, blir prognosen også dårlig.

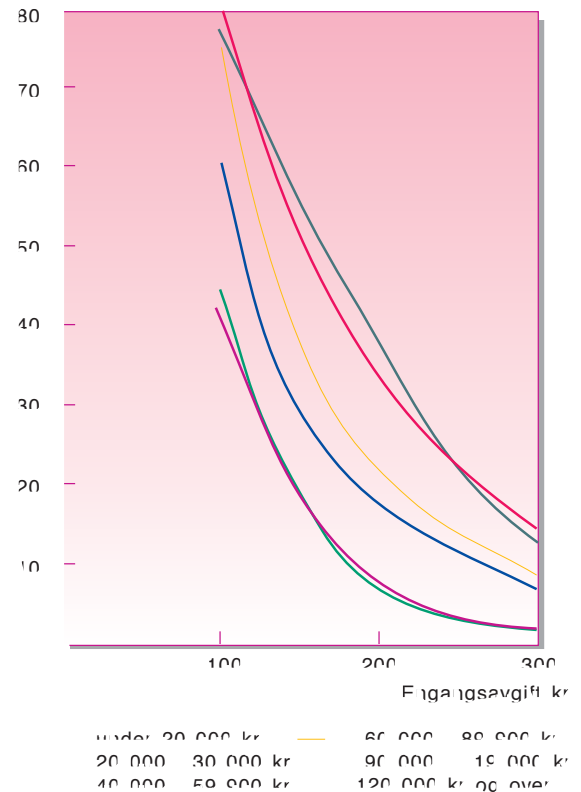
Få data gir også usikre prognoser. Dersom det er mulig, bør det i en slik situasjon arbeides med å framskaffe ytterligere data – enten ved observasjoner over et lengre tidsrom eller ved å se på en mer detaljert utvikling over samme tidsrom. Dette kan gjøres eksempelvis ved å gå over fra å lage prognoser for årlige data til å lage prognoser for kvartalsvise eller månedlige data. Det kan også være aktuelt å supplere de aktuelle data med data fra forklaringsvariable som vil bidra til sikrere prognoser.

I dataene som angir en gitt utvikling kan det være spesielle hopp (nivåskift), knekk (trendskift) eller spesielle unormale punkter (outliere) som det er viktig å justere for.

Figur 4.1 til 4.5 viser eksempel på en slik utvikling.

Det ses av figuren at antall tellerskritt på USA/Canada takst etter 1 januar 1992 gjør et markert hopp (nivåskift). Det er for øvrig her snakk om betydelige ekstrainntekter for Televerket som det var få som var klar over effekten av. Dette skyldtes at tellerskritt på innenlandsk opplysnings-tjeneste 1 januar 1992 ble lagt til USA/Canada takst – som skaffet Tele-

Prosentandel
av abonnentene



Figur 3.4 Estimert etterspørsel fra abonnenter i ulike inntektsgrupper i 1976 for anskaffelse av knappsettapparat

verket anslagsvis 300 mill kr i ekstrainntekter på årsbasis. Fra 1993 ble innenlandsk opplysnings-tjeneste lagt på en annen takstklasse.

For å kunne lage gode prognoser for tellerskritt, er det viktig å separere tellerskrittene på ulike takstklasser og lage prognoser på utviklingen innen hver enkelt takstklasse. For å kunne få en kontinuerlig utvikling relateres volumutviklingen til siste års takster, slik at en unngår hopp i tidsrekken hver gang lengden på tellerskrittet endres på de ulike takstklasser. I figuren vises det for øvrig at fra 1 januar 1991 opphørte rikstakst 1, slik at abonnentene etter det kun har tre ulike innenlandske takster. For å kunne ha data på sammenliknbar form før og etter endringen, er det nødvendig å justere dataene slik at det kan lages prognoser med samme sikkerhet som tidligere.

Det ses at utviklingen endrer seg markant. Dette skyldes at Televerket stoppet formidling av sex-telefoner over

sine Teletorgtjenester. Dette førte dermed til at denne type tjenester ble overført og formidlet fra andre steder som Hong Kong, de nederlandske Antiller og senere også Canada.

Figuren viser tellerskrittutviklingen på S12-sentralene i Region Oslo i januar 1991. Det er grunn til å tro at det var en liknende utvikling også i andre sentraler. Det ses at siste onsdag i januar (30 januar) var det kong Olavs begravelse, og at det var en markant nedgang i antall tellerskritt den dagen. Nedgangen var så stor at den også innvirket på antall tellerskritt på månedsbasis. Dette er et eksempel på en unormal begivenhet, og en slik observasjon kalles en outlier. I dataene bør det justeres for slike unormale hendelser for ikke å skape dårligere prognoser.

Også sportsbegivenheter påvirker abonnentenes bruk av teletjenester. Trafikkmålinger har vist nedgang i trafikk i takt med spennende sportsbegivenheter.

Det ses av figuren at netto-etterspørselen etter telefonabonnement i Region Oslo viste en markant økning de første månedene i 1993 og deretter en brå nedgang. Denne økningen, som var unormal, skyldtes imidlertid at inkassofunksjonen i Region Oslo ikke fungerte etter omorganiseringen. Dermed ble ikke en relativt stor del av abonnentene sagt opp på grunn av manglende betaling, slik som vanlig er hver måned. Dette førte da til at netto-etterspørselen økte markant selv om brutto-etterspørselen var som vanlig.

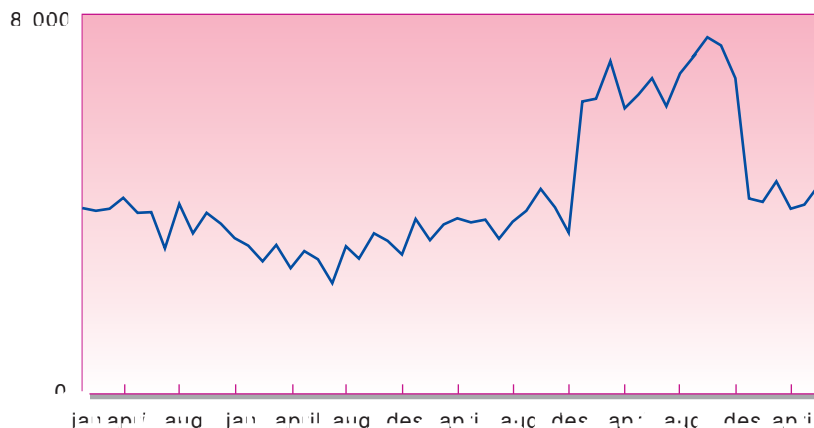
4.2 Vurdering av data

En forutsetning for å få gode underlagsdata er å forstå dataene som det skal lages prognoser for.

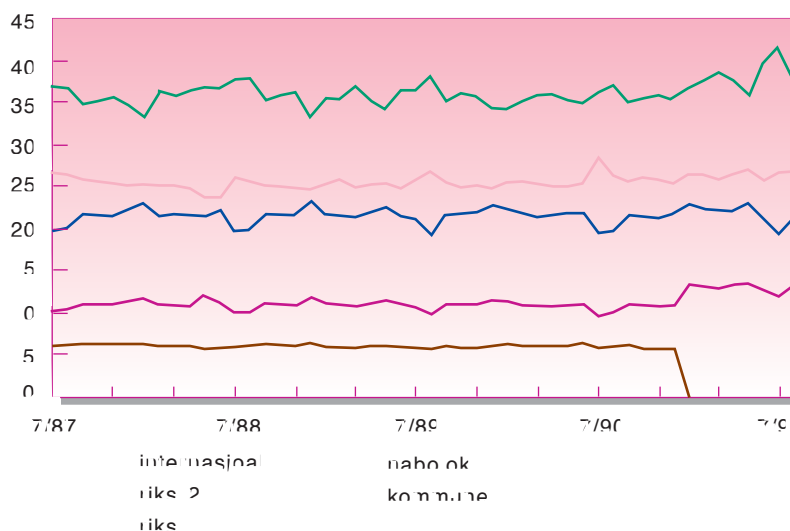
- Hvordan er dataene generert?
- Hva slags system/systemer er dataene hentet ut fra?
- Hvem har vært inne i bildet under håndtering av dataene?
- Hva beskriver dataene?
- Hvilken avhengighetsstruktur finnes mellom ulike variable?

Det er først når en arbeider med dataene at en får god forståelse for dataene – noe som er helt nødvendig for å kunne lage gode prognoser. Det er da en lærer om årsakssammenhenger som har betydning og om andre randbetingelser for variasjon i dataene.

Registrerte tellerskritt på USA/CND takst i Oslo.
Ikke justert for takstendringer

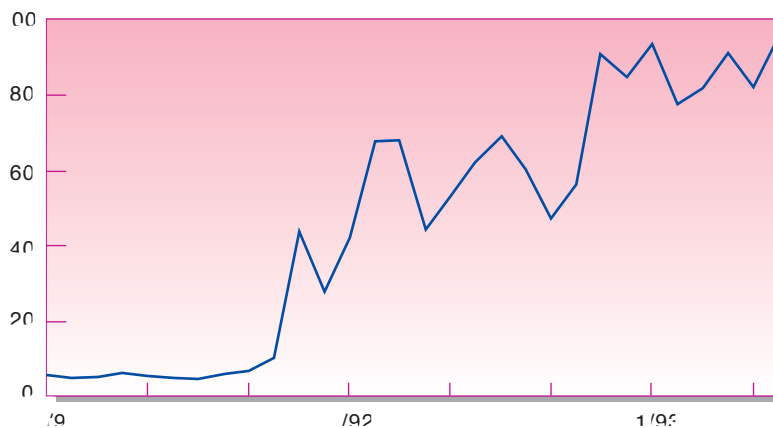


Figur 4.1 Antall tellerskritt på USA/Canada takst i Region Oslo

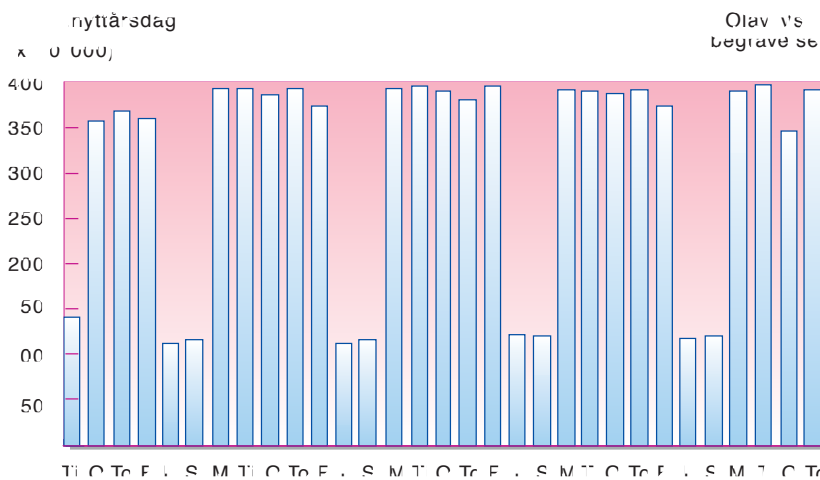


Figur 4.2 Prosentvis utvikling av telleskritt på ulike innenlandske takstklasser

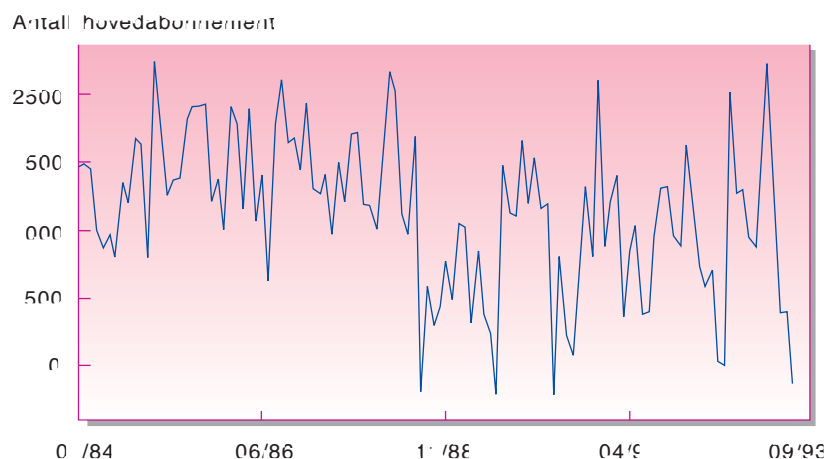
Samtallet inntetter (x 10 000)



Figur 4.3 Antall tellerskritt til Hong Kong



Figur 4.4 Utviklingen i tellerskritt januar 1991



Figur 4.5 Netto-etterspørsel etter telefonabonnement i Region Oslo fra januar 1984 til juni 1993

Det er alltid hensiktsmessig å benytte en grafisk presentasjon av dataene som et utgangspunkt for bearbeiding og analyse av dataene. Spørsmål som kan stilles, er:

- Hvor mye historiske data er tilgjengelig?
- Hvor gode/pålitelige er dataene?
- Er datene konsistente?
- Har det vært definisjonsendringer?
- Hva slags oppløsning er det mulig å få (årsdata, månedsdata, ukedata, etc)?
- Er det unormale variasjoner i dataene?
- Er det spesielle variable som påvirker utviklingen til dataene?

Ut fra spørsmålene må det i mange tilfeller gjøres supplerende arbeid for å få

underlagsdataene på en brukbar presentasjonsform.

4.3 Eksempler på bearbeiding av data

Hovedabonnement telefon

I [4] er det gitt en detaljert beskrivelse av hvorledes etterspørselsdata for hovedabonnement må oppdeles for å kunne få laget gode prognoser. Her foretas det kun en kort oppsummering.

For en del år siden var det kun nødvendig å se på brutto-etterspørsel, som da omtrent var lik netto-etterspørsel siden det var få som sa opp eller fikk sagt opp abonnementet. For de siste årene har denne utviklingen endret seg markant. Figur 4.6 viser utviklingen i antall opp-

sigelser. Det ses av figuren at antall oppsigelser nærmest har øket med 200 %. Denne økningen skjedde svært brått i 1988. Ser vi på de siste årene er antall oppsigelser omlag tre ganger så mange som netto-etterspørsel. Det betyr at brutto-etterspørsel, som er summen av netto-etterspørsel og antall oppsigelser, er om lag fire ganger så stor som netto-etterspørselen.

Det ses også at det i noen måneder er en negativ netto-etterspørsel.

Status for antall innkoplede abonnement fordelt på abonnemeststyper finnes i TELSIS/AB Rapport 35. Den totale etterspørsel er imidlertid gitt ved:

- antall installerte abonnement
- + antall ueffektuerte nybestillinger
- antall ueffektuerte oppsigelser
- + antall effektuerte tilflyttinger
- antall ueffektuerte fraflyttinger
- + antall effektuerte ikke oppdaterte nybestillinger
- antall effektuerte ikke oppdaterte oppsigelser
- + antall effektuerte ikke oppdaterte tilflyttinger
- antall effektuerte ikke oppdaterte fraflyttinger
- = Total etterspørsel etter abonnement

I dette systemet med effektueringer kan det skje opphopninger som da vil påvirke den totale etterspørselen. Det er derfor viktig å ha oversikt over hvordan etterspørselsdataene framkommer og likeledes benytte definisjoner som forhindrer at store bevegelser i dataene kan skjules.

I måned 7 1988 og i måned 3 1990 har det vært definisjonsendringer i TELSIS som har gjort det nødvendig å justere dataene.

Når det gjelder oppsigelser, så har det også vært nødvendig å skille antall oppsigelser i antall oppsigelser fra abonnement og antall oppsigelser fra Televerket grunnet manglende betaling (inkassofunksjon). Figur 4.7 viser utviklingen av disse kurvene. Det ses at antall oppsigelser grunnet manglende betaling har et meget stort omfang. Dersom Televerket over en periode ikke håndhever denne funksjonen, vil det på grunn av opphopning kunne føre til en negativ netto-etterspørsel. Når det skal lages prognoser, bør det, så fremt det er mulig, justeres for en slik utvikling.

Differansen mellom brutto-etterspørsel og antall oppsigelser er som nevnt netto-etterspørsel. Utviklingen i netto-etterspørsel er gitt i figur 4.5.

Til slutt nevnes at når det gjelder prognoser for abonnement for det enkelte sentralområde, er det viktig å gjennomgå underlagsdataene ytterligere. Her kan det være nødvendig å korrigere dataene for tidligere nummerlån. Det vil si nummer som er lånt eller lånt bort til andre sentralområder. Likeledes kan det være aktuelt å korrigere for eventuell endring på sentralområdegrensene som vi har hatt mye av.

I tillegg til dette er det også ønskelig å kunne segmentere disse dataene slik at vi får ut etterspørselsdata både for bedriftsmarkedet og for privatmarkedet. Her er det nødvendig å gjøre en god del anslag idet underlaget for disse dataene fra tidligere var svært mangelfullt.

Tellerskritt

Fremdeles kommer mer enn halvparten av Telegruppens inntekter fra tellerskritt. Dette markedssegmentet er derfor av usedvanlig stor betydning for Televerket. Helt til ganske nylig har det vært store problemer med å få gode tellerskrittdata fordelt på ulike takstklasser. Med innføring av digitale sentraler har det vært mulig å få fram denne type data.

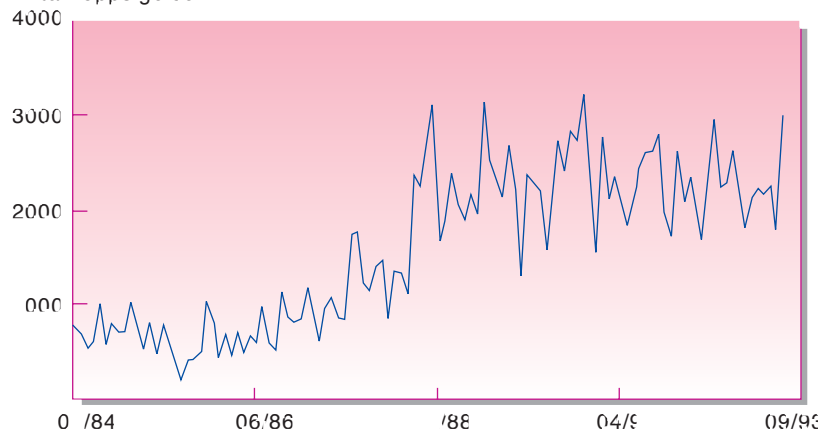
For i en overgangsfase å få en forholdsvis lang serie med observasjoner, noe som er nødvendig for å lage gode prognoser, har det vært nødvendig å hente data fra flere kilder for så å sette dette sammen på best mulig måte.

I Region Oslo ble prognosene i en periode basert på:

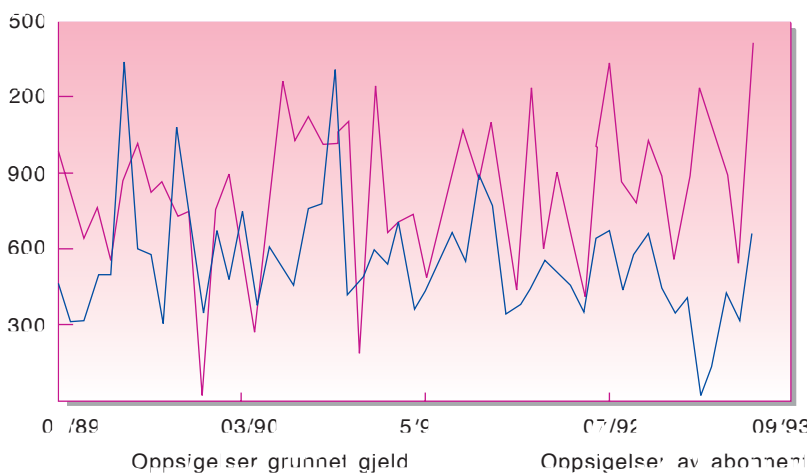
- Manuelle månedlige tellerskrittmålinger fra alle 10C sentraler
- STYRAN målinger fra alle S12 sentraler. Disse målingene ble senere foretatt i MEAS da dette målesystemet kom i funksjon
- RITEL målinger basert på fakturerte tellerskritt
- RUST målinger som er debiterte tellerskritt tatt ut av RITEL tellerarkiv.

De manuelle 10C målingene og STYRAN/MEAS målingene angir tellerskritt-etterspørsel totalt og fordelt på ulike takstklasser, mens RITEL og RUST målingene kun angir totalt antall tellerskritt. 10C målingene og STYRAN/MEAS målingene dekker imidlertid ikke alle

Antall oppsigelser



Figur 4.6 Utviklingen av antall oppsigelser fra januar 1984 til juni 1993 i Region Oslo



Figur 4.7 Utvikling i antall oppsigelser fordelt på oppsigelser av abonnent og oppsigelser på grunn av gjeld

abonnentene. Til gjengjeld har vi 10C målinger lengst tilbake i tid.

I figur 4.8 er det vist hvorledes det er mulig å integrere informasjon fra flere målinger for å få et underlag for å lage prognoser. Det henvises til artikkel [5] for nærmere beskrivelse.

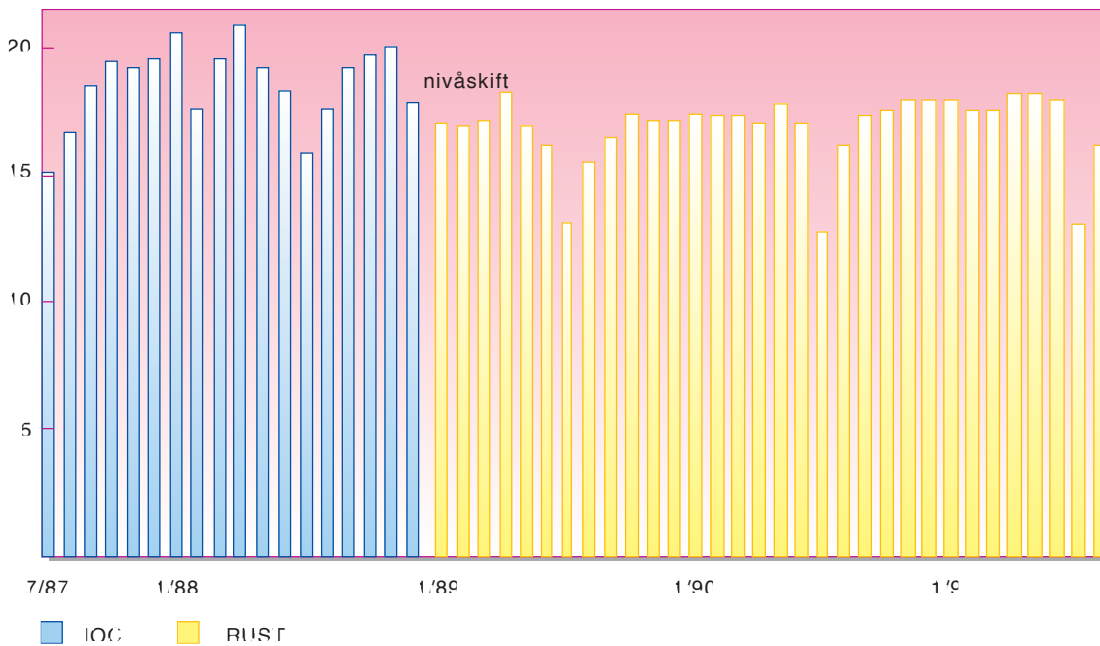
Etterspørselen etter tellerskrittene er avhengig av

- takstene, det vil si pris på tellerskritt og tellerskritt lengde
- samtalelengden
- forholdet mellom antall virkedager og antall helligdager
- antall abonnenter.

Det er derfor nødvendig å bygge inn disse variablene for å kunne lage prog-

noser på tellerskritt. Figur 4.9 viser hvorledes endring i takstene – det vil si tellerskritt lengden – påvirker volumet av tellerskritt. Siden abonnentene ikke reagerer momentant på takstendringer når tellerskritt lengden eksempelvis økes, ville det bli hopp i tellerskrittutviklingen dersom det ikke foretas justeringer basert på tellerskritt lengde.

Dersom tellerskritt lengden innen en takstklasse eksempelvis økte med 10 % fra en måned til neste, eksempelvis ved årsskiftet, skulle en tro at tellerskritt volumet ville avta med 10 %, dersom det ses bort fra månedlige variasjoner, generell vekst og eventuell rask respons på lavere pris. Dette er imidlertid ikke tilfelle fordi det ikke er noen uniform holdtidsfordeling. Figur 4.10 viser et eksempel på holdtidsfordeling på lokalsamtaler. Mellom



dere med antall abonnenter. Dermed framkommer en tidsrekke som angir tellerskritt volum pr abonnent. Når det skal lages tellerskrittprognoser, kan prognosen for tellerskritt volum pr abonnent så multipliseres med abonnementsprognosene.

Disse eksemplene fra hovedabonnement på telefon og tellerskritt viser

Figur 4.8 Integrasjon av tellerskrittdata fra ulike målesystemer

hvor viktig det er med håndtering av data før det kan lages prognoser.

kl 08 og kl 17 er tellerskritt lengden 3 minutter. Det er klart at selv om tellerskritt lengden ble økt med 10 %, så ville dette ikke ha betydning for de samtaler som er på under 3 minutter. Av figuren ser vi at mer enn 70 % av lokalsamtalene er på mindre enn 3 minutter. Mindre enn ett tellerskritt er det ikke mulig å betale. Figuren viser en tilnærmet eksponentiell fordeling. Det er derfor nødvendig å justere også for de samtaler som er på mer enn 3 minutter.

Trafikk- og tellerskritt målinger viser at det går atskillig mer trafikk i telenettet på virkedager enn på helligdager. Det er derfor nødvendig å justere trafikkvolumet både på måneds- og årsbasis ut fra forholdet mellom virkedager og helligdager.

Endelig vil tellerskritt volumet være avhengig av antall abonnenter. En framgangsmåte for å få en utvikling som er uavhengig av abonnentvekst, er å divi-

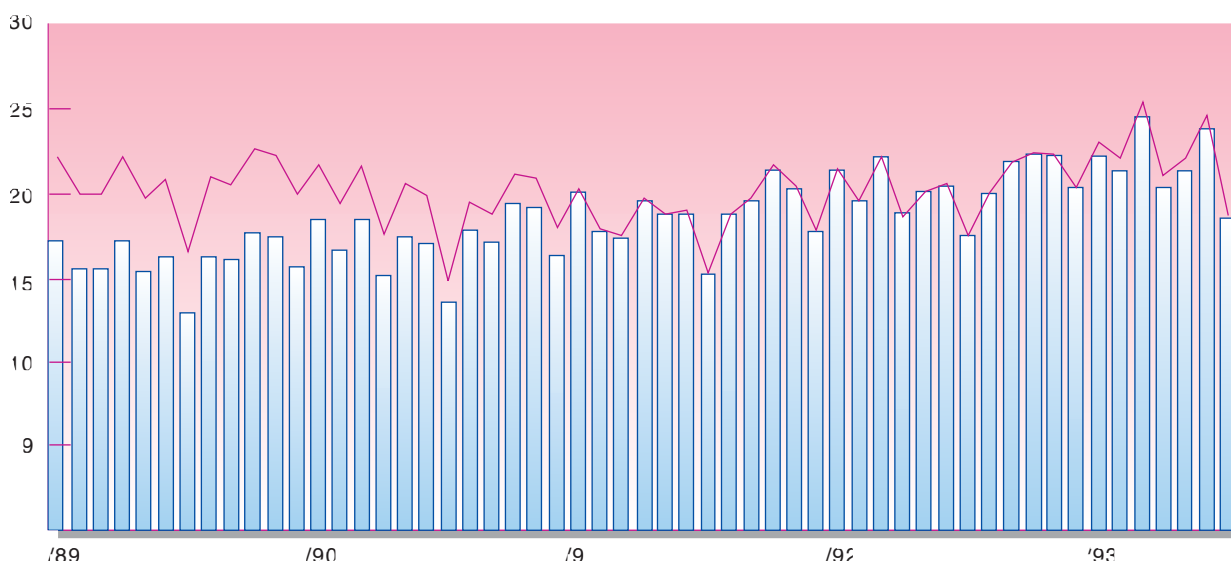
5 Tidsrekker og komponenter

5.1 Innledning

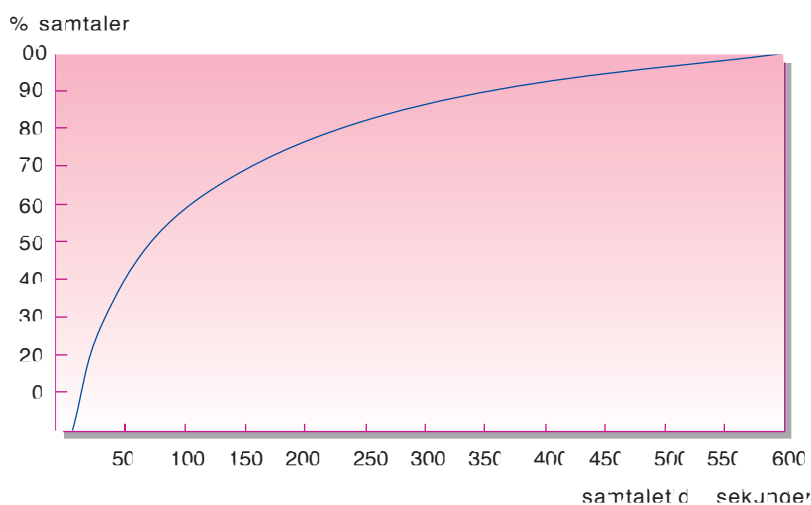
I dette kapittelet vil det bli vist en del ulike tidsrekker, slik at vi skal få et forhold til tidsrekkene. Før analysen starter er det mye en tidsrekke kan indikere.

Derfor vil vi i første del av kapittelet vise ulike tidsrekker som hver for seg har ulike egenskaper.

(x 100)



Figur 4.9 Tellerskritt volum både justert og ikke-justert for takstendringer



Figur 4.10 Estimert akkumulert holdetidsfordeling for kommunetakst

I tilknytning til veksten i en tidsrekke er det en rekke faktorer som spiller inn. Faktorer som har betydning, er:

- Sesongkomponenter
- Ulike former for trender
- Tilfeldige variasjoner
- Outliere (unormale observasjoner)
- Nivåskift (hopp i tidsrekken)
- Andre spesielle endringer.

Det er viktig å kjenne til de ulike faktorene og kunne identifisere hvilke faktorer som er til stede i ulike tidsrekker. Først da har vi mulighet til å modellere tidsrekken på en profesjonell måte.

Vi ser ofte utviklingen beskrevet ved ulike tidsrekker og da særlig som tilveksten ved hver observasjon eller som det aggregerte tallet ved hver observasjon. Et eksempel her er årlig etterspørsel etter en tjeneste. Her kan tidsrekken enten framstilles som den årlige tilveksten eller som det akkumulerte volumet.

Når det gjelder å beskrive et problem, kan det også være flere angrepsmåter der ulike tidsrekker benyttes. Dette kommer vi tilbake til både under eksemplene i dette kapitlet og senere.

5.2 Eksempler på tidsrekker

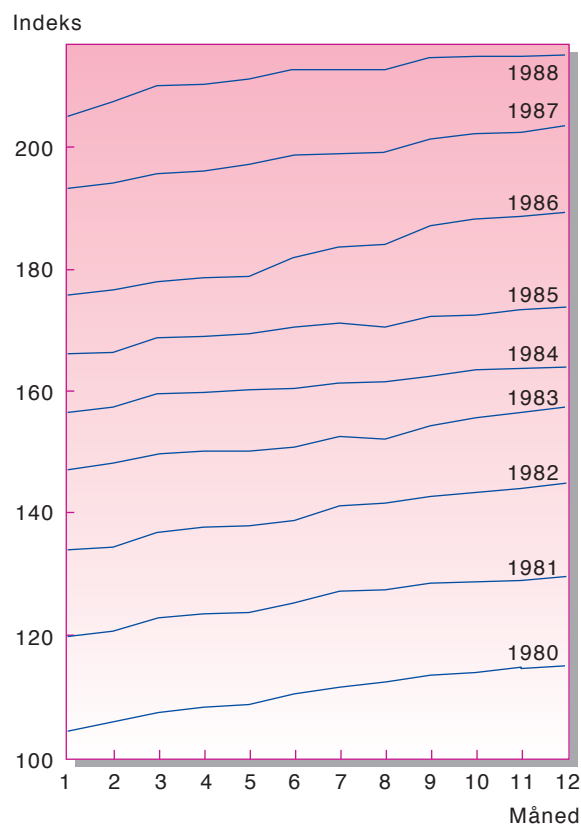
Det vil nå bli gitt eksempler på ulike typer tidsrekker for å gi en bedre forståelse og gjerne grunnlag for en bedre strukturert tankegang i tilknytning til tidsrekker og prognostisering.

En tidsrekke som forteller om kostnadsutviklingen i Norge er konsumprisindeksen. Indeksen er basert på at beregningen av kostnader for kjøp av et representativt utvalg av "produkter", det vil si mat, klær, bensin, etc, hver måned. Denne angir da hvor mye dyrere det er én måned i forhold til en annen.

I mange prognosemodeller brukes konsumprisindeksen som en forklaringsvariabel. Det betyr at denne indeksten er med på å forklare utviklingen i det vi skal prognostisere. Et eksempel kan være etterspørselen etter en tjeneste eller et produkt. Vi kjenner de framtidige takstene eller prisene på dette produktet. Dermed kan disse takstene eller prisene brukes som forklaringsvariabel. For å få satt takstene/prisene inn i en mer reell sammenheng kan det imidlertid være lurt å dividere med konsumprisindeksen. Den faktoren som da kommer fram forteller reelt sett om takstene er dyrere de kommende måneder i forhold til nå, og dette vil åpenbart ha betydning for den framtidige etterspørsel etter produktet.

Derfor er konsumprisindeksen en meget viktig størrelse.

La oss nå se på hvorledes konsumprisindeksen har utviklet seg i løpet av 80-årene. I figur 5.1 er konsumprisindeksen tegnet inn fra januar 1980 til desember 1988. I 1979 ble indeksten normert slik at



Figur 5.1 Konsumprisindeks 1980-1988. Månedlig utvikling

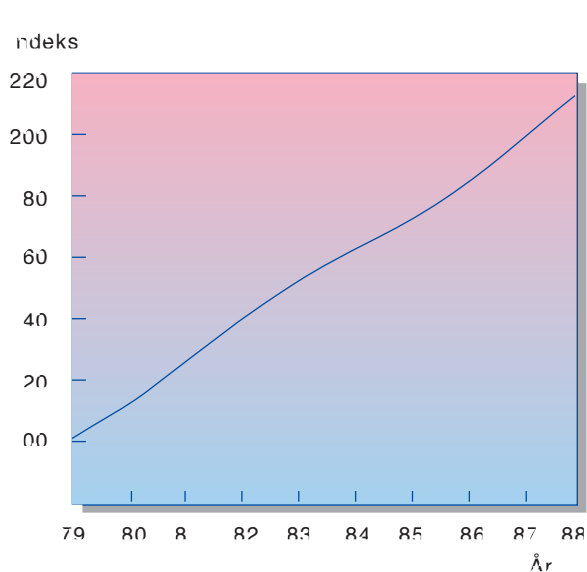
gjennomsnittlig indeks dette året var 100. I januar 1980 var indeksten 104.0.

Av figuren ser vi at konsumprisindeksen er jevnt stigende. Det er noe sesongbetoning i utviklingen. Med det menes at det er samme mønster i utviklingen samme måned i ulike år. Dette kommer nok best fram i det hoppet som kan registreres fra desember til januar hvert år. For øvrig kan det legges merke til at det er stagnasjon i veksten, stort sett, fra juli til august hvert år.

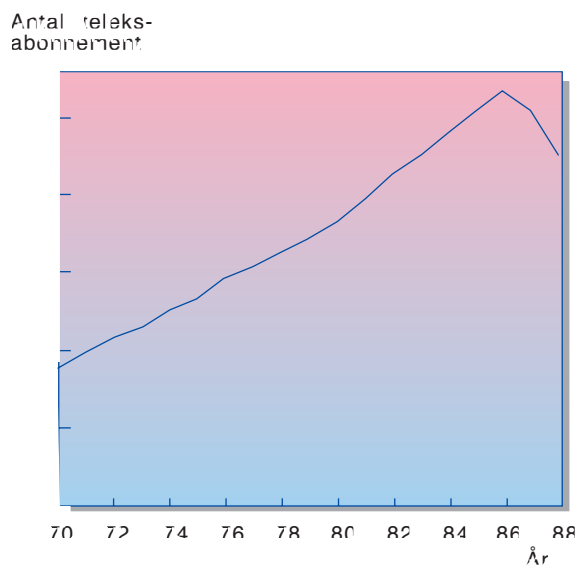
Vi ser også at i 1986 øker konsumprisindeksen kraftig og da særlig fra mai måned. Generelt er det en meget jevn utvikling i tidsrekken.

Når det gjelder framstilling av en tidsrekke med sesongvariasjoner, er det to muligheter: enten å framstille hele tidsrekken i ett, eller slik som vist på figur 5.1. I mange tilfeller vil en framstillingsform som i figur 5.1 lettere vise sesongvariasjonene.

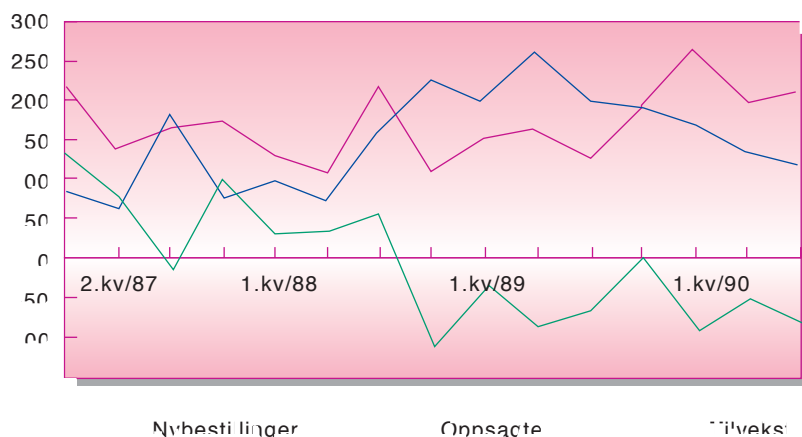
Figur 5.2 viser den årlige utviklingen av konsumprisindeksen. Vi ser at den stiger jevnt. Det er en tendens til at den flater



Figur 5.2 Konsumprisindeks 1980–1988. Årlig utvikling



Figur 5.3 Etterspørsel etter teleksabbonnement. Akkumulert antall abonnemnt, hele Norge



Figur 5.4 Etterspørsel etter datexabbonnement – tilvekst, nybestillinger og antall oppsagte

noe ut i 1982–1984, mens den i 1987 og 1988 har steget kraftig.

Dersom vi skulle lage prognoser og hadde både månedlige og årlige data, ville det være gunstig å utnytte den informasjon som ligger i de månedlige dataene i prognosene.

I figur 5.3 er den årlige utviklingen av etterspørsel etter teleksabbonnement vist. Vi ser av figuren at det er en jevn stigning fram til 1986, deretter avtar tidsrekken.

Dette er en meget interessant tidsrekke. Tidsrekken angir et “turning point”. Dette vil alltid være en utfordring for de som lager prognoser. I de fleste tilfeller vil det ikke være nok å se på utviklingen av selve tidsrekken for å kunne prog-

nostisere slike markante endringer. Dette kan ligge i årsakssammenhenger som må modelleres inn i selve prognosemodellen.

I dette tilfellet vet vi hva som er årsaken til denne utviklingen. Det skyldes en konkurrerende tjeneste: telefaks. Denne tjenesten har tatt en del av markedet til teleks, og denne utviklingen har av de som lager prognoser vært forutsett i lang tid.

Vi ser av figuren at det ikke bare er veksten i antall abonnemnt som avtar. Siden figuren angir totalt antall abonnemnt, ser vi at fra 1986 har vi nedgang i det totale antall abonnemnt. Det betyr at vi har en negativ etterspørsel, hvilket vil si at vi har avgang av teleks-abbonnenter.

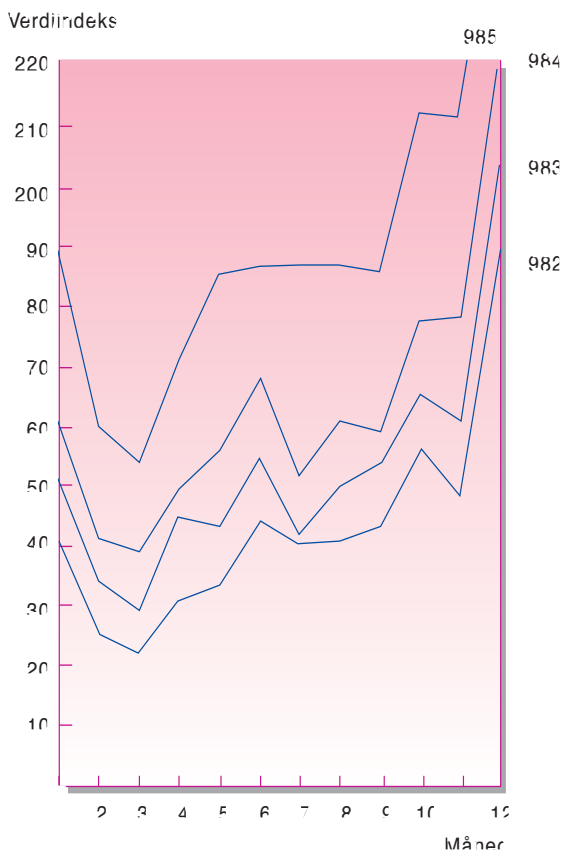
Når det skal lages prognoser videre her, er vi i en vanskelig situasjon. Det må da legges inn modellforutsetninger om hvorledes telefaks, og på sikt andre tjenester som EDI, vil ta over markedet og eventuelt hvilke markedsandeler som ikke vil gå over til de substituerende tjenestene.

Figur 5.4 viser etterspørselen etter datexabbonnement for Region Oslo fra 1987 til 3. kvartal 1990. Dette er et eksempel på en tidsrekke som er sesongbetont med kvartalsvise observasjoner.

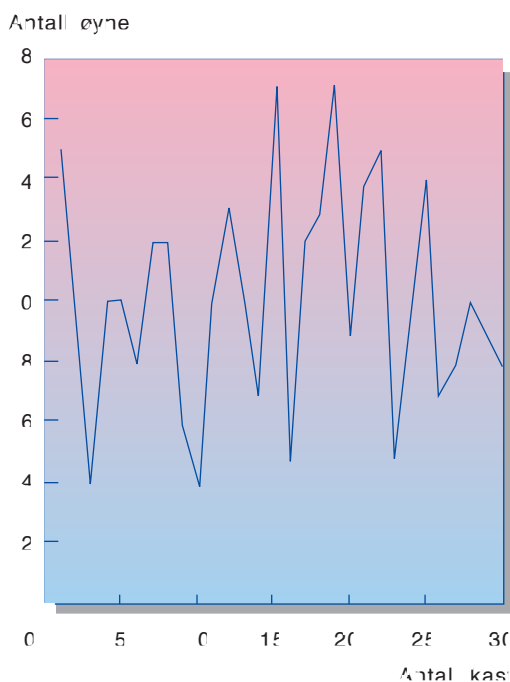
Det ses at etterspørselen er avtagende – og ikke nok med det – den er også negativ. Det betyr, som vi ser, at Region Oslo etter 1988 stort sett har hatt en avgang i totalt antall abonnemnt.

I figuren er det også tegnet inn antall nybestillinger hvert kvartal og antall oppsagte abonnemnt hvert kvartal. Det er interessant å legge merke til at antallet nybestillinger holder seg på samme nivå over hele perioden. Dette er ikke usedvanlig. Det som imidlertid har skjedd med oppsagte abonnemnt er mer uvanlig. Fra å ligge på et forholdsvis lavt nivå i 1987 og 1988, øker så antall oppsagte abonnemnt kraftig i 1989 og 1990.

Det er dette som er årsaken til den markante nedgangen i tilveksten som i denne perioden faktisk er negativ. Dette skyldes en rekke faktorer. Når det gjelder datexabbonnement, er det mange kunder som bestiller høyere hastighet. I tillegg har det vært økonomiske nedgangstider,



Figur 5.5 Verdiindeks for detaljomsetning



Figur 5.6 Antall øyne ved kast av tre terninger. Eksempel på en tilfeldig tidsrekke

som også avspeiler seg på etterspørselen. I den anledning har nok mange bedrifter ryddet opp i sin datakommunikasjon for å få den mer effektiv. Dette har også ført til oppsigelser av abonnement.

Det er også viktig å være klar over at de ulike datatjenestene konkurrerer med hverandre. En effektivisering fra bedriftens side vil da gå ut over etterspørselen.

For øvrig er det viktig når det skal lages prognoser, å se tjenester i sammenheng der disse fungerer som substitutter for hverandre. Dette gjelder i høy grad dataoverføring ved tjenestene datex, datapak, digital og utleide samband.

I figur 5.5 er verdiindeks for detaljomsetningen tegnet inn. Også dette er en tidsrekke med sesongvariasjoner. Vi ser at det er en trend i tidsrekken og klare sesongvariasjoner. For 1985 er det en større vekst enn de øvrige år. Derfor ser vi at den karakteristiske knekken ned i august og september er borte. Dette skyldes ikke manglende sesongvariasjon disse månedene, men en vekst som skjuler sesongvariasjonen.

Også i denne figuren ser vi at sesongvariasjonene kommer godt fram når de framstilles oppdelt på år.

I figur 5.6 har vi angitt resultatene av kast med tre terninger over en periode på 30 kast. For hvert kast er summen av antall øyne på terningen beregnet.

Vi ser at det minste som er oppnådd i perioden, er 4 og det meste 17. Det er tydelig at antall øyne i hvert kast er tilfeldig. Det betyr at det foregående eller de foregående kast ikke influerer på neste kast. Dermed har vi en tilfeldig tidsrekke som varierer mellom 3 og 18.

Begrepet tilfeldig er av stor betydning i statistisk modellbygging. Vi vil behandle dette begrepet nærmere senere.

En annen måte å generere en tilfeldig tidsrekke på er å trekke tall fra en slumptalltabell. Dette er en tabell som inneholder tilfeldige tall. Det kan også gjøres ved å bruke en slumptallgenerator i en datamaskin som så genererer de nødvendige tall.

I figur 5.7 er det generert 30 tilfeldige tall. De er generert fra en normalfordeling med forventning 0 og standardavvik lik 1. Det karakteristiske ved fordelingen er den symmetriske "klokkeformen" rundt midtpunktet som i vårt tilfelle er 0. Dette betyr at tyngdepunktet ligger rundt 0 og at sannsynligheten for store avvik fra tyngdepunktet er liten.

Av figur 5.7 ser vi at kun én av de 30 observasjonene ligger mer enn to ganger standardavviket (dvs 2) utenfor forventningen, som er 0. Likeledes ser vi her som i figur 5.6 at observasjonene er tilfeldige. Det vil igjen si at verdien av tidligere observasjoner ikke påvirker senere observasjoner.

5.3 Egenskaper til tidsrekkene

Vi har nå sett på en rekke forskjellige tidsrekker for å få en bedre forståelse av hvorledes de utvikler seg og hvorledes de kan illustreres. Det er klart at det er en rekke ulike årsaker til utviklingen av de ulike tidsrekkene. Dette er årsaken til at det kreves et grundig arbeid for å kunne utvikle gode prognosemodeller for tidsrekkene.

Underlaget for en god prognosemodell vil alltid være de systematiske forklaringer om utviklingen av tidsrekken kombinert med framtidige vurderinger. Det viktigste vil derfor være å lage en modell som på analytisk måte beskriver utviklingen av tidsrekken fram til nå.

Denne modellen vil ha som egenskap at den ikke bare beskriver utviklingen fram til nå; den vil også beskrive den framtidige utviklingen, hvilket er ekvivalent med prognoser.

For å kunne lage en god modell og dermed gode prognoser for en tidsrekke, er det nødvendig å kjenne til egenskapene til tidsrekken. La oss på grunnlag av blant annet eksemplene foreta en oppsummering så langt.

De grunnleggende egenskapene eller komponentene til tidsrekkene er:

- Trend
- Sesong
- Tilfeldighet (støy).

I tillegg kan det forkomme spesielle karakteristikkene i tidsrekkene. Her nevnes:

- outlier
- nivåskift
- trendskift
- turning point.

Disse karakteristikkene er illustrert i figur 5.8 til 5.11.

Outlier

En outlier (kalles også uteligger) er en observasjon som ligger langt unna det som er forventet verdi. Med langt unna

menes at det statistisk ikke er trolig at observasjonen er riktig. Dersom vi ikke har så mange observasjoner, vil en observasjon som ligger 3 – 4 ganger unna forventet verdi være en slik observasjon.

Et eksempel på en slik verdi er gitt i figur 5.8 Vi ser av figuren at en av observasjonene ligger langt under de øvrige observasjonene.

Spørsmålet er så hva som er årsaken til at denne observasjonen er så forskjellig i forhold til de øvrige observasjonene. Det kan være flere årsaker til dette, som feil avlesning, feil koding av dataene, feil innlegging eller punching, etc.

En annen forklaring kan være at selve observasjonen er helt riktig, men den har oppstått på grunn av unormale hendelser. Det kan eksempelvis ha vært en streik, eller helt spesielle værforhold, eller andre ekstreme forhold som har vært skyld i observasjonen. Da det i den videre utvikling ikke kan forventes liknende observasjoner, er det ikke naturlig å ta hensyn til slike observasjoner under modelleringen. Det som imidlertid er viktig, er å identifisere slike observasjoner slik at de ikke påvirker modelleringen.

Nivåskift

I figur 5.9 er det vist et sprang i tidsrekken. Det kan skje i løpet av en eller flere på hverandre følgende observasjoner. Vi har også sett tendenser til det i noen tidligere figurer. Dersom spranget eller nivåskiftet i tidsrekken skjer ved at veksten i tidsrekken før og etter spranget er noenlunde det samme og at nivåskiftet har en spesiell forklaring, eksempelvis en takstnedsettelse, er det viktig å være klar over dette under modelleringen.

I noen tilfeller bør tidsrekken justeres for nivåskiftet under modelleringen slik at prognosene ikke påvirkes for mye av spranget. I andre tilfeller ligger forklaringen på spranget i modellen gjennom forklaringsvariable.

Trendskift

I figur 5.10 er det vist et trendskift i tidsrekken. Etter hvert vil dette identifiseres og automatisk ligge inne i modellen til tidsrekken. Dersom trendskiftet er meget markant og skjer over kort tid og gjerne at det er naturlige forklaringer til det, bør det identifiseres og legges inn i modellen.

Turning point

I figur 5.11 er turning point til en tidsrekke vist. Dette er også vist i figur 5.3 og kommentert der. Dette er ofte en vanskelig sak. Det er meget viktig å prognostisere turning point for en tidsrekke. Dersom dette skjer ut fra en kontinuerlig avtagende vekst eller ut fra markante endringer i forklaringsvariable, er det greit. Verre er det hvis turning pointet kommer ut fra akutte uforutsigbare hendelser som eksempelvis den første oljekrisen. Men uansett, det er viktig å følge opp modelleringsprosessen for å kunne fange dette opp på best mulig måte.

6 Oppbygging og glatting av tidsrekker

6.1 Oppbygging av en tidsrekke

Når vi lager modeller for en tidsrekke og deretter prognoser, er utgangspunktet å “dissekere” tidsrekken for å finne fram til hvilke komponenter den består av. Etter en inngående analyse for å finne fram til de ulike komponenter samt beregning av de aktuelle parametrene, lages det så en modell for tidsrekken. Det er denne modellen som brukes til å lage prognoser.

Som en innledning skal vi nå prøve å illustrere hvorledes en tidsrekke kan være sammensatt av ulike komponenter. I dette tilfellet nøyter vi oss med å anta at tidsrekken er forholdsvis stabil og består av

- trend
- sesong
- tilfeldige variasjoner.

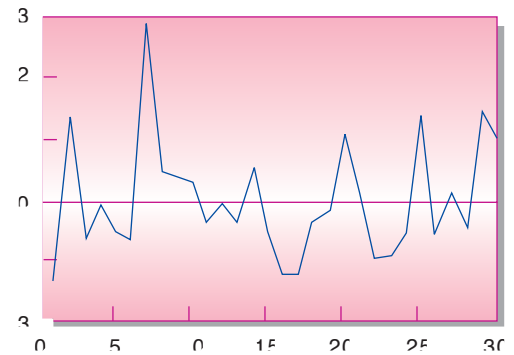
Det antas at det ikke forekommer outlier, nivåskift eller trendskift i tidsrekken.

I figur 6.1, 6.2 og 6.4 er henholdsvis trenden, sesongvariasjonene og de tilfeldige variasjonene vist.

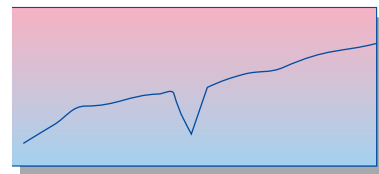
Trenden

Vi kunne her vist en rekke forskjellige trender. Vi har valgt å vise en lineær trend. Det kunne vært en avtagende trend i forskjellige former eller det kunne vært en mer voksende trend i forskjellige former. Siden det er prinsippene som det er viktig å formidle, har vi her holdt oss til en konstant voksende trend som er lineær og som har formen:

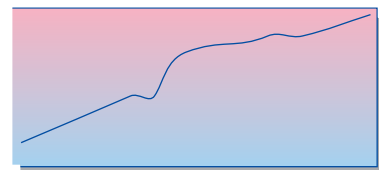
$$y = a + bt$$



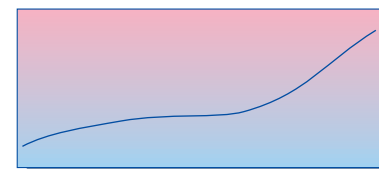
Figur 5.7 Tilfeldig tidsrekke generert av normalfordeling med forventning 0 og standardavvik 1



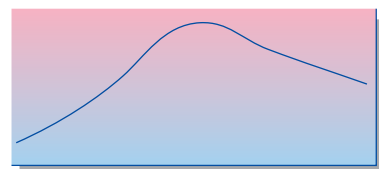
Figur 5.8 Outlier



Figur 5.9 Nivåskift



Figur 5.10 Trendskift



Figur 5.11 Turning point

Vi har brukt verdiene:

$$a = 10,0$$

$$b = 0,5$$

Det vil si at formen på trenden er:

$$y = 10,0 + 0,5t$$

Dette ses av figur 6.1. Det kan også legges merke til at vi har 16 observasjoner som er nummerert ved tidspunktene $t = 1, 2, 3, \dots, 16$.

Sesongvariasjoner

Vi ser så på sesongvariasjonene. De er vist i figur 6.2. Dette er de rene sesong-

variasjonene der alle andre variasjoner er fjernet. Som vi ser er det meget regelmessige variasjoner. Det finnes to ulike mønstre. Enten øker sesongvariasjonene med økende nivå på trenden eller så er sesongvariasjonene uavhengig av nivået på trenden. Det er sistnevnte tilfelle vi har i figur 6.2.

Som det framgår av figuren er det kvartalsvise sesongvariasjoner. Det betyr at det er 4 observasjoner i løpet av året – en observasjon i hvert kvartal. Verdien på de enkelte sesongutslagene er:

- Kvartal nr 1: 2,0
- Kvartal nr 2: -0,5
- Kvartal nr 3: -1,0
- Kvartal nr 4: -0,5

Dette er faste utslag i den tid det er foretatt observasjoner. For å se på den samlede effekten av trenden og sesongutslagene, så summeres disse. Dette er gjort i figur 6.3.

Tilfeldige variasjoner

Det vil alltid være tilfeldige variasjoner knyttet til en tidsrekke. Ellers ville jo tidsrekken være deterministisk – uten usikkerhet, og slike situasjoner har vi jo ikke.

I modellbyggingen er det et mål å få en så lav usikkerhet som mulig. Usikkerheten, det vil si de tilfeldige variasjonene, representerer det vi ikke greier å forklare. Vi må derfor søke å redusere de tilfeldige variasjonene til et minimum når vi lager en modell for tidsrekken.

Dersom vi lager en dårlig modell, vil de tilfeldige variasjonene være store. Dette vil da også føre til usikre prognoser.

I figur 6.4 er det vist et sett med tilfeldige variasjoner. Disse er generert som observasjonene i figur 5.7, fra en normalfordeling med forventning lik 0 og standardavvik 1. Vi ser av figuren at ingen av de 16 observasjonene er større enn to ganger standardavviket som er lik 2. Normalt vil 5 % av observasjonene (det vil si én av 20) ligge utenom to ganger standardavviket.

For øvrig ser observasjonene ut som tilfeldige variasjoner uten noe mønster, med hopp opp og ned.

Ved å summere trenden, sesongvariasjonene og de tilfeldige variasjonene som er vist i figurene 6.1, 6.2 og 6.4, får vi så den aktuelle tidsrekken. Det er også mulig å summere dataene fra figur 6.3 og 6.4. Den aktuelle tidsrekken er vist i figur 6.5.

Vi ser at formen på tidsrekken nå er mer ugjenkjenkelig enn i figur 6.3 der vi ikke hadde med komponenten med tilfeldige variasjoner. Vi har ikke lenger den regelmessigheten som vi hadde tidligere. Imidlertid synes sesongvariasjonene fremdeles. Samtidig er det også klart at vi har en stigende trend. Dette greier de tilfeldige variasjonene ikke å skjule.

Dersom de tilfeldige variasjonene var en god del større, kunne det blitt problemer med å skille ut både trend og sesong. Det som er poenget, er imidlertid ikke bare å kunne identifisere om det er en trend og om det er sesongvariasjoner. Det er også av stor betydning å kunne kvantifisere størrelsene på trenden og på sesongvariasjonene på en brukbar måte.

Selve identifiseringen av trenden og sesongvariasjonene vil vi se nærmere på i den påfølgende del av kapittelet, mens selve prognosemodellene er presentert senere.

Tabell 6.1 gir en oppsummering av dataene fra figurene.

Som vist i tabellen har vi følgende størrelser:

t :	tiden
y :	tidsrekken
$a + bt$:	trenden der $a = 10,0$ og $b = 0,5$
S :	sesongkomponenten der $S = 2,0, -0,5, -1,0$ eller $-0,5$
e :	tilfeldige variasjoner

Vi får da følgende likning:

$$y = a + bt + S + e \quad (6.1)$$

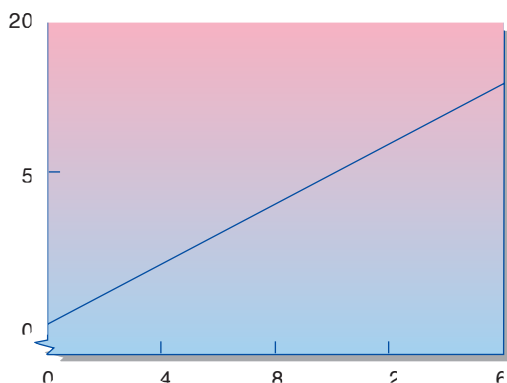
I dette tilfellet har parametrene a og b samt sesongutslagene S vært kjent på forhånd. Slik er det ikke i en vanlig situasjon. Da må vi beregne disse parameterverdiene.

Vi introduserer nå metoder for glatting av tidsrekker som vil hjelpe oss i arbeidet med å identifisere de ulike komponentene til tidsrekker.

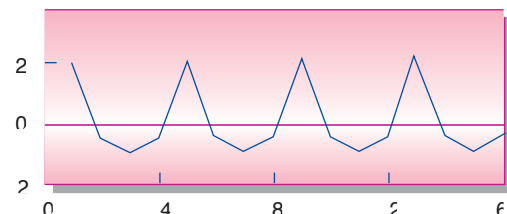
6.2 Glatting av tidsrekker

Det finnes en rekke ulike metoder til å glatte tidsrekker. Metodene varierer sterkt i kompleksitet. I dette kapittelet vil vi kun se på noen enkle metoder.

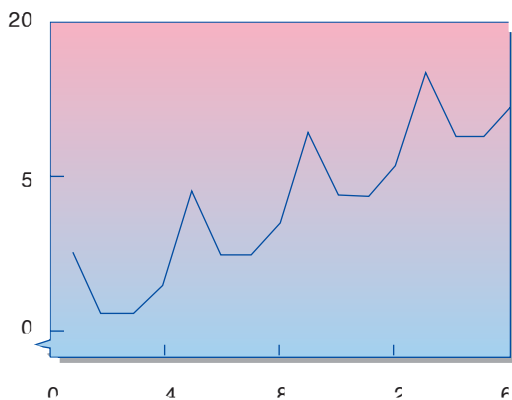
Glattingmetoder har blitt brukt i lang tid. De ble blant annet brukt i tilknytning til dekomponeringsmodeller som var den



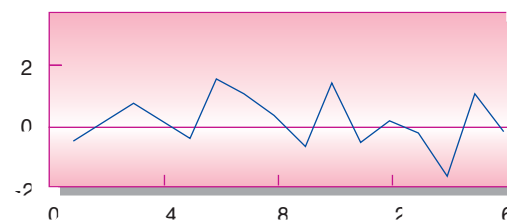
Figur 6.1 Trend



Figur 6.2 Sesongvariasjoner



Figur 6.3 Trend og sesongvariasjoner



Figur 6.4 Tilfeldige variasjoner

klassiske prognosemetoden i den første del av det 20. århundre (1900–). Dekomponeringsmodellene baserer seg, slik navnet sier, på å separere tidsrekken i ulike komponenter. Dette vil vi vise i kapittel 6.3. For å kunne gjøre dette, er det nødvendig å ha metoder for glatting av tidsrekker.

Før vi går videre med metoder for glatting nevnes det at det også i dag finnes anerkjente dekomponeringsmodeller. Den mest kjente er X11. I dette kompendiet vil vi imidlertid ikke konsentrere oss om denne type modeller. Det som kort kan sies, er at det i de bedre dekomponeringsmodellene manipuleres mye med dataene ved å bruke glattingsmodeller en rekke ganger for å få skilt ut komponentene. Deretter settes så komponentene sammen til en prognosemodell.

La oss nå illustrere poenget med glatting av en tidsrekke.

Vi ser at denne tidsrekken får et glattere forløp etter at det er foretatt en glatting av observasjonene.

Tidsrekken i figur 6.6 bestod av tilfeldige variasjoner på begge sider av 0. Disse tilfeldige variasjonene var trukket ut av en normalfordeling med forventning 0 og standardavvik 1. Vi vet fra før at disse observasjonene vil gruppere seg på begge sider av 0.

I eksempelet med rekruttene fra kapittel 5 grupperte observasjonene seg på begge sider av 180, som var forventningsverdien. Det vi imidlertid la merke til, var at da vi summerte et sett med observasjoner og tok gjennomsnittsverdien fikk vi en verdi som lå nærmere 180 enn om vi så på en enkelt observasjon.

Dette er også prinsippet med glatting av observasjoner. Vi summerer et sett med observasjoner og finner gjennomsnittet. Vi vet da at en del av effekten av de tilfeldige variasjonene elimineres. Dermed vil det være lettere å identifisere trend og sesong.

I figuren har vi tatt gjennomsnittet av fem og fem observasjoner. Vi ser tydelig at de glattede observasjonene ligger nærmere forventningsverdien 0 enn det selve observasjonene gjorde.

Som tidligere nevnt har vi ulike glattingsmetoder. Følgende nevnes her:

- Glidende gjennomsnitt
- Veiet glidende gjennomsnitt
- Eksponentiell glatting.

Den siste metoden kan også brukes direkte som en prognosemetode. Forutsetningen er imidlertid at tidsrekken ikke har sesongvariasjoner, og at den i rimelig grad ligger på samme nivå.

Glidende gjennomsnitt

Glidende gjennomsnitt går som nevnt ut på å summere et sett med på hverandre følgende observasjoner, det kan være to, tre, fire ..., og så beregne gjennomsnittet av disse.

Anta at vi bestemmer oss for å ta gjennomsnittet av tre observasjoner. Dette betyr at vi starter med å ta gjennomsnittet av observasjon nr 1, 2 og 3. Deretter tar vi gjennomsnittet av observasjon nr 2, 3 og 4. Deretter tar vi gjennomsnittet av observasjon nr 3, 4 og 5, osv helt til vi har tatt gjennomsnittet av de tre siste observasjonene.

Den første gjennomsnittsverdien eller glattingen plasseres i punkt nr 2, som er senterpunktet for de tre første observasjonene. Den neste i punkt nr 3, som er senterpunktet for de neste observasjonene, osv.

Dette er også illustrert i tabell 6.2.

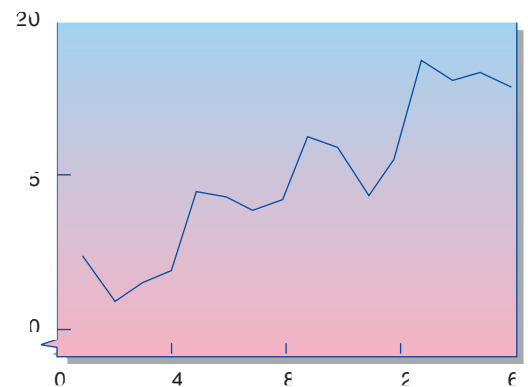
Vi ser av tabellen at tidsrekken som er angitt til venstre er voksende, men den hopper mye opp og ned. Ved å glatte observasjonene får vi et jevnere forløp. Både glattingen ved bruk av tre og av fem observasjoner gir en jevnere stigning. Trenden i tidsrekken framkommer nå ut fra de glattede observasjonene.

Fra tabellen kan det også registreres at vi mister observasjoner når vi glatter dem. Når vi glatter med tre observasjoner, mister vi en glattet på hver side. Når vi glatter med fem observasjoner, mister vi to glattede på hver side.

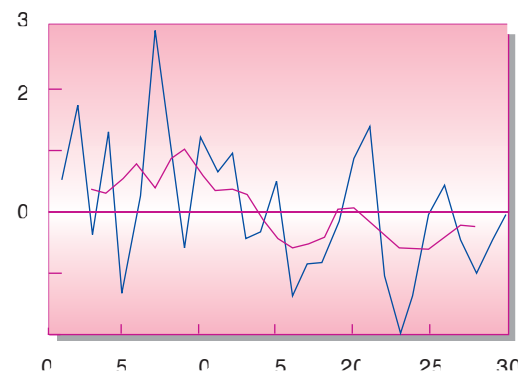
Dette er en grunn til at det er vanskelig å lage prognoser med glattings- og dekomponeringsmodeller. Det er lite ønskelig å miste de siste observasjonene som jo er de viktigste, da de forteller om den siste utviklingen. I de dekomponeringsmodellene som brukes i dag, tas det inn andre klassiske prognosemetoder i kombinasjon for å dekke opp de siste observasjonene.

I dette kompendiet vil vi imidlertid kun bruke glattings-

modeller til identifikasjon av de ulike komponentene.



Figur 6.5 Tidsrekke med trend, sesong og tilfeldige variasjoner



Figur 6.6 Glatting av en tidsrekke

Tabell 6.1 Tidsrekke med trend, sesong og tilfeldige variasjoner

Tid t	Trend a + bt	Sesong S	Tilfeldige var. e	Tidsrekke a + bt + S + e
1	10,5	2,0	-0,3	12,2
2	11,0	-0,5	0,2	10,7
3	11,5	-1,0	0,8	11,3
4	12,0	-0,5	0,2	11,7
5	12,5	2,0	-0,3	14,2
6	13,0	-0,5	1,6	14,1
7	13,5	-1,0	1,1	13,6
8	14,0	-0,5	0,4	13,9
9	14,5	2,0	-0,5	16,0
10	15,0	-0,5	1,0	15,5
11	15,5	-1,0	-0,5	14,0
12	16,0	-0,5	0,2	15,7
13	16,5	2,0	-0,2	18,3
14	17,0	-0,5	1,1	17,6
15	17,5	-1,0	1,2	17,7
16	18,0	-0,5	-0,1	17,4

Tabell 6.2 Eksempel på glatting av observasjoner

Tid	Tidsrekke	Glatting med 3 obs	Glatting med 5 obs
1	5		
2	4	5,3	
3	8	6,3	6,5
4	7	8,0	7,6
5	9	8,7	8,4
6	10	9,0	9,6
7	8	10,7	10,4
8	14	11,0	11,6
9	11	13,3	
10	15		

Veiet glidende gjennomsnitt

Veiet glidende gjennomsnitt er basert på den samme tankegangen som glidende gjennomsnitt. Den eneste forskjellen er at hver enkelt observasjon gis en vekt. Den vekten som ble gitt de enkelte observasjoner i glidende gjennomsnitt var den inverse av antall observasjoner. Med andre ord, når vi benyttet tre observasjoner fikk hver vekt $1/3$, og når vi hadde fem observasjoner fikk hver vekt $1/5$.

Vi legger merke til at summen av alle vektene er lik 1. Dette er også tilfelle i veiet glidende gjennomsnitt, men her er det ikke noe krav om at alle observasjonene har samme vekt. I prinsippet kan alle vektene være helt ulike bare summen av vektene er lik 1.

Det som i praksis har vist seg lønnsomt, er å ha vekter som er symmetriske rundt midtpunktet. Et eksempel på slike vekter er:

Observasjon	Vekt
observasjon nr 1:	0.10
observasjon nr 2:	0.20
observasjon nr 3:	0.40
observasjon nr 4:	0.20
observasjon nr 5:	0.10
Sum vekter	1.00

Vi ser at midtobservasjonen har den høyeste vekten og at vektene på begge sider er symmetriske om denne observasjonen. Summen av vektene er lik 1.

Det som må bestemmes når observasjoner skal glattes, er hvor mange observasjoner som skal være med i glattingen og hvorledes de skal vektlegges. Dersom denne glattingen kun skal brukes for identifisering av ulike komponenter,

er det ingen grunn til å være raffinert med hensyn til valg. Dersom glattingen skal brukes i en dekomponeringsmodell, må det imidlertid foretas et sett med ulike glattinger. Dette er imidlertid godt beskrevet i litteratur om dekomponeringsmodeller.

Det som vi imidlertid bør merke oss når vi glatter tidsrekker med sesongvariasjoner, er at antall

observasjoner i glattingen da er bestemt. I disse tilfellene skal antall observasjoner i glattingen være lik antall observasjoner i sesongen. Dersom vi eksempelvis har månedsobservasjoner, har vi 12 observasjoner i løpet av året. Da skal antall observasjoner i glattingen være 12. Årsaken til dette er at glattingen av tidsrekken skal gi oss trenden. Når vi så summerer over observasjoner i en hel sesong vil de positive og negative sesongutslagene oppheve hverandre. Dermed vil trenden komme godt til syne. Dette ses godt på figur 6.2 som bare representerer sesongvariasjoner. Når vi glatter denne, vil samtlige av de glattede verdier bli lik 0.

Ekspontiell glatting

I motsetning til i vanlig glatting brukes det en parameter som vi betegner med a i eksponentiell glatting.

Ekspontiell glatting er også en prognosemodell. Dersom denne modellen skal kunne brukes, må det ikke være sesongvariasjoner i observasjonene og samtidig bør tidsrekken være tilnærmet stasjonær.

Modellen er basert på følgende:

n er nummeret på siste observasjon

y_n er siste observasjon

$\hat{y}_{n-1}(1)$ er prognosen et skritt fram fra tidspunkt $n-1$

$\hat{y}_n(1)$ er prognosen et skritt fram fra tidspunkt n

Vi har da følgende prognosemodell:

$$\hat{y}_n(1) = (1 - a)y_n + a\hat{y}_{n-1}(1) \quad (6.2)$$

Dette er det samme som å si:

Ny prognose = $(1 - a) * (\text{ny observasjon}) + a * (\text{gammel prognose})$.

Algoritmen for å lage prognosene blir da:

- Først lages en prognose
- Deretter kommer det inn en ny observasjon
- Deretter lages det en ny prognose, etc.

Det gjelder å finne startverdier for å starte denne algoritmen og så bestemme best mulig verdi på glattingsparameteren a . Dette kommer vi nærmere tilbake til i kapittel 8.

Det er viktig å være klar over at denne glattingsmodellen ikke må brukes når vi har en sterkt stigende trend. Vi får da en kraftig estimering av tidsrekkens utvikling.

Glattingparameteren a (a mindre enn 1) kan interpreteres ved at siste observasjon gis vekten a , nest siste observasjon gis vekten a opphøyd i andre, etc, slik at vektene på de følgende observasjonene avtar eksponentielt. De avtagende vektene er da:

$$a^2 \ a^3 \ a^4 \ \dots$$

Glatting ved forskjellig usikkerhet

Til slutt skal vi se hvorledes glattingen kan brukes til å få innsikt i strukturen på en tidsrekke under ulik grad av usikkerhet.

I tabell 6.3 har vi data for en tidsrekke som består av komponentene

- trend
- sesong
- støy.

Det er forutsatt en lineær trend som stiger i samme takt som tiden t . I tillegg har vi en kvartalsvis sesongkomponent som er lik

- første kvartal: 8
- andre kvartal: -1
- tredje kvartal: -5
- fjerde kvartal: -2

Tidsrekken basert på lineær trend og sesongkomponenter, men uten støy, er vist i figur 6.7.

Tabell 6.3 gir en oversikt over dataene for tre ulike tidsrekker med samme trend og samme sesongkomponenter, men med ulike støykomponenter. Tabellen viser følgende:

Tabell 6.3 Tidsrekker med forskjellige støykomponenter

Tid	Trend	Sesong	Trend + Sesong	Støy	T.rekke1	T.rekke2	T.rekke3
1	1	8	9	0,6	9,6	10,2	10,8
2	2	-1	1	1,7	2,7	4,4	6,1
3	3	-5	-2	-0,3	-2,3	-2,6	-2,9
4	4	-2	2	1,3	3,3	4,6	5,9
5	5	8	13	-1,3	11,7	10,4	9,1
6	6	-1	5	0,2	5,2	5,4	5,6
7	7	-5	2	2,9	4,9	7,8	10,7
8	8	-2	6	1,0	7,0	8,0	9,0
9	9	8	17	-0,5	16,5	16,0	15,5
10	10	-1	9	1,2	10,2	11,4	12,6
11	11	-5	6	0,7	6,7	7,4	8,1
12	12	-2	10	1,0	11,0	12,0	13,0
13	13	8	21	-0,4	20,6	20,2	19,8
14	14	-1	13	-0,3	12,7	12,4	12,1
15	15	-5	10	0,5	10,5	11,0	11,5
16	16	-2	14	-1,3	12,7	11,4	10,1
17	17	8	25	-0,8	24,2	23,4	22,6
18	18	-1	17	-0,8	16,2	15,4	14,6
19	19	-5	14	-0,2	13,8	13,6	13,4
20	20	-2	18	-0,9	17,1	16,2	15,3
21	21	8	29	1,4	30,4	31,8	33,2
22	22	-1	21	-1,0	20,0	19,0	18,0
23	23	-5	18	-1,8	16,2	14,4	12,6
24	24	-2	22	-1,3	20,7	19,4	18,1

- Kolonne 1: tiden
- Kolonne 2: lineær trend
- Kolonne 3: sesongkomponentene
- Kolonne 4: lineær trend + sesongkomponent
- Kolonne 5: støyen
- Kolonne 6: lineær trend + sesongkomponent + støy * 1
- Kolonne 7: lineær trend + sesongkomponent + støy * 2
- Kolonne 8: lineær trend + sesongkomponent + støy * 3

De tidsrekkeene som er dannet i kolonne 6, 7 og 8 består av samme trend og samme sesongkomponent. Når det gjelder støyen, så er den forsterket for tidsrekkeene i kolonne 7 og 8 ved at støykomponenten multipliseres med henholdsvis 2 og 3.

Støyen er trukket ut fra en tabell med normalfordelte verdier. Den er vist i figur 6.8.

Figur 6.9, 6.10 og 6.11 viser tidsrekkeene T.rekke1, T.rekke2 og T.rekke3 fra tabell 6.3. Det ses at det etter hvert blir vanskeligere å anslå sesongkomponentene og

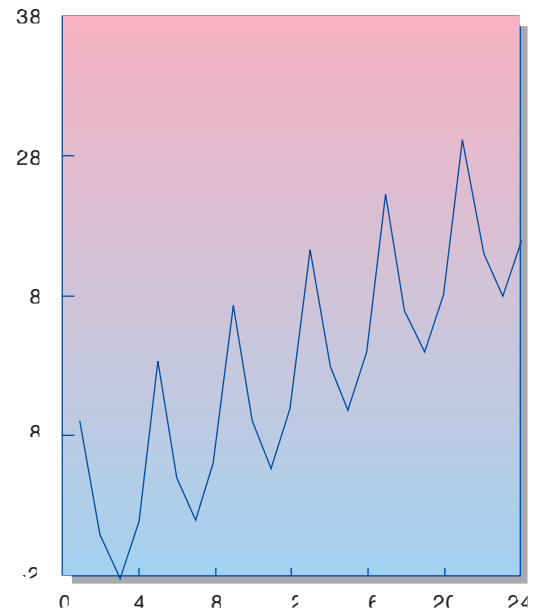
trenden etter hvert som støyen får en økende effekt.

I figur 6.12, 6.13, 6.14 og 6.15 er tidsrekkeene fra figur 6.7, 6.9, 6.10 og 6.11 glattet.

I figur 6.12 har vi glattet tidsrekken fra figur 6.7 som kun bestod av en lineær trend og sesongkomponenter, men uten støykomponenter. Vi ser at glattingen av tidsrekken gir eksakt den underliggende trenden tilbake. Dette skyldes at det ikke er støy i tidsrekken. I figur 6.13 ser vi at glattingen fjerner seg noe fra den lineære trenden, men at den fremdeles ligger nær opp til en lineær trend. I figur 6.14 forsterkes denne effekten, og det samme ser vi av figur 6.15. Her har åpenbart støyen hatt stor innvirkning.

Vi registrerer imidlertid at glattingen er et hensiktsmessig og slagferdig virkemiddel for bruk til tentativ analyse av tidsrekker.

Glattingen viser hva slags trend det er i tidsrekken. I tillegg kan den gi foreløpige anslag for sesongkomponentene, og når det er gjort kan det gis anslag for støyen. Dette er imidlertid en foreløpig analyse som kun utføres for at vi skal få kjennskap til tidsrekke før vi starter selve modellbyggingen.



Figur 6.7 Tidsrekke med lineær trend og kvartalsvise sesongkomponenter uten støy

Alle utkjøringer som er vist i figurene 6.7–6.15, er gjort ved bruk av programpakken Statgraphics.

7 Prinsipper for modellering

7.1 Innledning

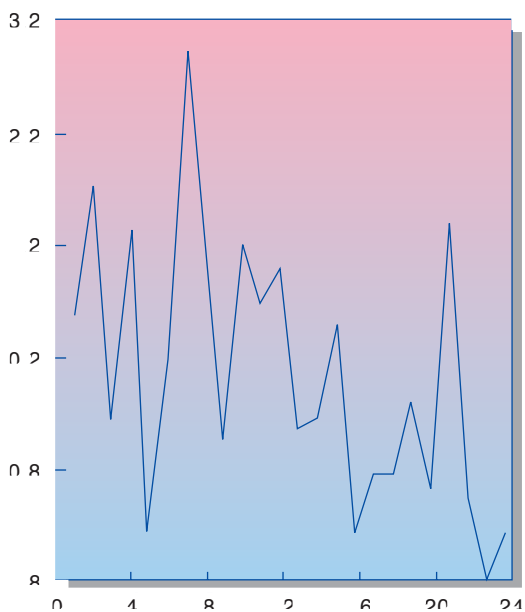
Det er laget flere lærebøker i prognoser som på sine felt utdyper det stoffet som presenteres her. De viktigste bidragene finnes i [6], [7], [8], [9], [10], [11], [12], [13].

I dette kapitlet vil vi se på de grunnleggende forutsetningene under utvikling av prognosemodeller. Alle profesjonelle prognostisører bør kjenne til og benytte de prinsipper som presenteres her.

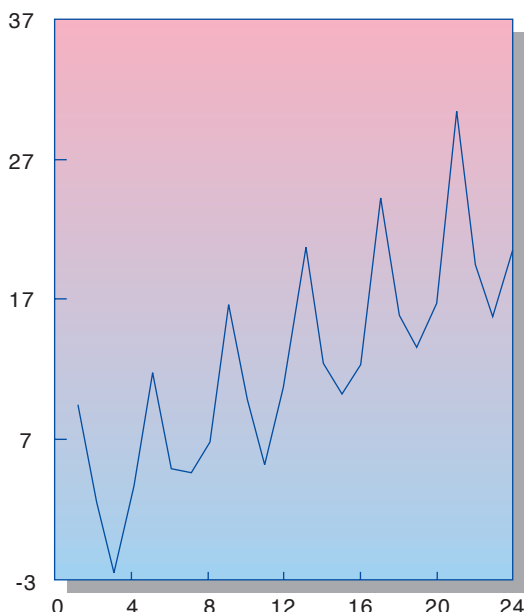
Det er lagt vekt på å gjennomgå hvorledes arbeidet med utvikling av prognosemodellene gjøres og hvilke prinsipper som benyttes før de ulike typer prognosemetoder presenteres i de neste kapitler.

7.2 Støyen som beskriver de tilfeldige variasjonene

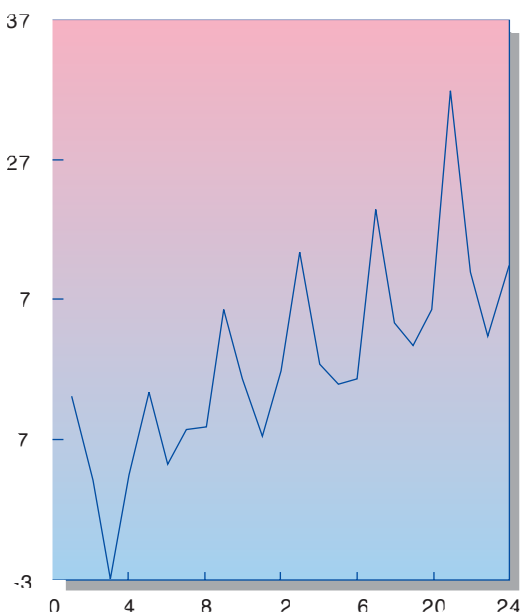
Et nøkkelbegrep i modellbyggingen er de tilfeldige variasjonene som vi heretter kaller støyen. Det er den delen av tidsrekkenes variasjon som vi ikke greier å forklare.



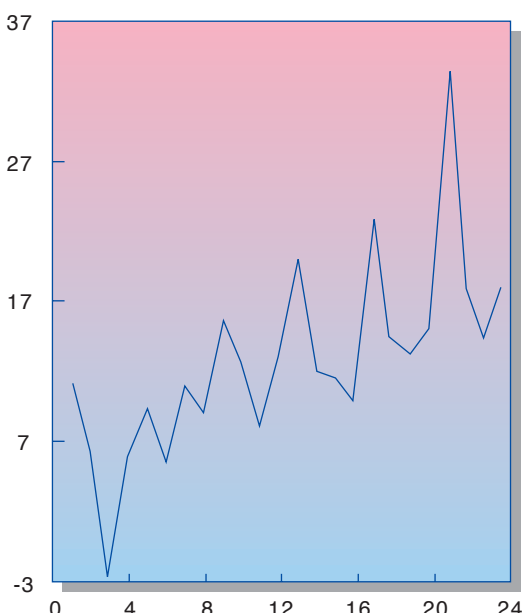
Figur 6.8 Støyleddene til tidsrekken



Figur 6.9 Tidsrekke 1 (T.rekke1)



Figur 6.10 Tidsrekke 2 (T.rekke2)



Figur 6.11 Tidsrekke 3 (T.rekke3)

Anta eksempelvis at vi har en tidsrekke som består av både trend og sesongkomponent. Når vi har beregnet eller laget modell for trenden og for sesongen, vil vi stå tilbake med støyen. Det som er hovedmålet under modelleringen, er å gjøre støykomponenten så liten som mulig samtidig som den skal tilfredsstillе bestemte krav.

Vi er nå inne på det mest sentrale punkt under modellbyggingen. Dersom vi lykkes med å tilfredsstillе kravene til

støyen, er vi sikret en god modell og gode prognoser.

De kravene som stilles til støyen, er:

- 1) Forventningen til støyen (de tilfeldige variasjonene) skal være lik 0. Det vil si:

$$E e_t = 0 \quad (7.1)$$

- 2) Variansen til støyen skal være konstant og ikke variere med tiden. Det vil si:

$$Var e_t = \sigma_t^2 \quad (7.2)$$

- 3) Det skal ikke være noen avhengighet (*Cov*) mellom de ulike støyleddene. Dette skrives på matematisk form ved:

$$Cov(e_t, e_{t+k}) = 0 \quad (7.3)$$

- 4) Støyleddene skal være normalfordelt.

Det første kravet om at forventningsverdien til støyen skal være lik 0, er innlysende. Vi må være klar over at støyen skal representere tilfeldige variasjoner. Disse skal ikke ha noe mønster og selvfølgelig ha en forventning Δ som er lik null. Anta imidlertid at $\Delta = 10$. Det betyr i så fall at støyleddene fluktuierer rundt 10. Vi har da laget en modell som systematisk underestimerer alle verdiene i tidsrekken med 10. Det samme vil være tilfelle med prognosene. I et slikt tilfelle burde $\Delta = 10$ adderes inn i modellen.

A priori bør prognosemodellen være slik at det ikke er noen systematisk skjevhet; det vil si at forventningen til støyen må være lik 0.

Dersom forutsetningen om at variansen til støyen ikke er konstant, vil vi få usikre anslag på de parametrene som skal estimeres. Dermed vil vi også få introdusert større usikkerhet i prognosene enn nødvendig.

På samme måte vil det være dersom det er avhengighet mellom de ulike støyleddene. Dette vil også føre til større usikkerhet i de parametrene som skal estimeres og dermed også større usikkerhet i prognosene.

Forutsetningen om at støyleddene skal være normalfordelte er ikke så kritisk. Det er mulig å foreta det meste av modellbyggingen uten at denne forutsetningen er til stede. En del av testene for signifikans i parametre (det vil si at parametrene er forskjellig fra 0) kan imidlertid ikke brukes dersom støyen ikke er tilnærmet normalfordelt.

En illustrasjon av støyledd som er tilfeldige er gitt i figur 5.7 og 6.4

7.3 Enkel modellbygging

Utgangspunktet for modellbyggingen er at vi har et sett med data. Det vil være selve tidsrekken og eventuelt andre forklaringsvariable. Ut fra disse dataene lages det så en modell som så godt som mulig beskriver den tidligere tidsrekken. Differansen mellom tidsrekken og modellverdiene vil da være det vi ikke greier å forklare – nemlig støyen.

Dette er illustrert i figur 7.1

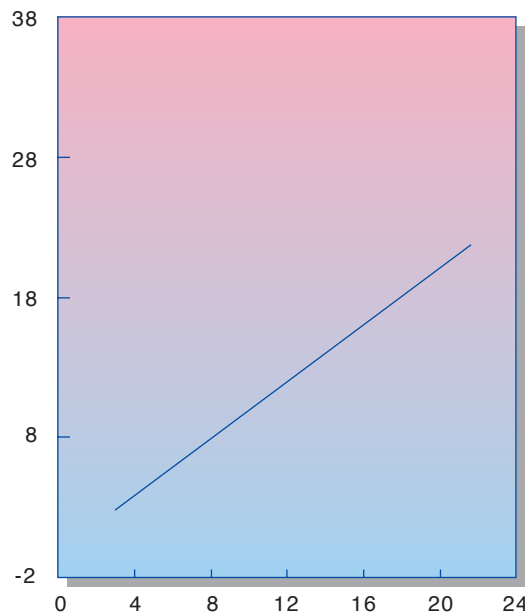
Figuren viser grovt hvorledes det arbeides med modelleringen. På venstre side har vi input-dataene. Ut fra disse lages det en modell som på best mulig måte beskriver den aktuelle tidsrekken. Output-dataene på høyre side er støyleddene.

Dersom vi har laget en god modell, skal støyleddene tilfredsstillende krav som vi har angitt i kapittel 7.2. Det finnes en rekke teknikker som kan brukes slik at støyen skal få de angitte egenskaper. Likeledes finnes det flere tester som verifiserer om støyleddene har de gitte egenskaper. Dette vil bli omtalt i mer detalj senere i kapitlet.

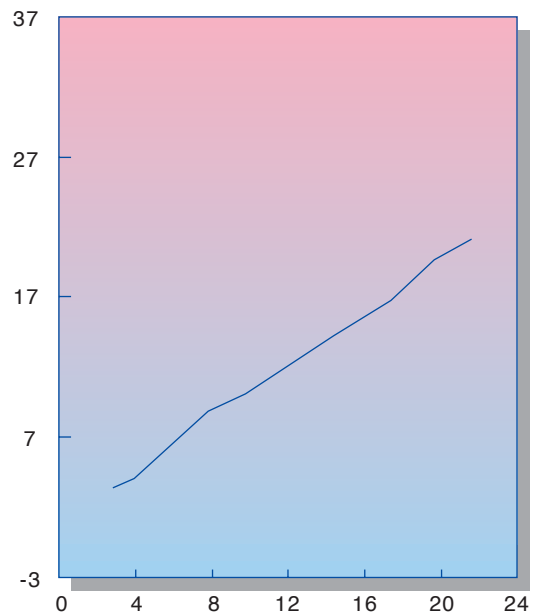
Vi introduserer nå et nytt begrep som brukes mye i tilknytning til modellbygging. Det er residualer (avvik mellom estimert verdi og observert verdi) som er realiseringen av støyen. Dette betyr at vi skiller på begrepet støy og på residualen. Vi snakker med andre ord om støyen helt til vi kjenner verdien på støyen, og det er denne verdien som kalles residualen.

Anta eksempelvis at vi betrakter tidspunktet januar 1989. Den observerte verdien for tidsrekken i dette tidspunktet er 1179. Vi lager så en modell for tidsrekken slik som blant annet vist i figur 7.4 og 7.5. Når vi lager disse modellene, ønsker vi blant annet å gjøre støyen så liten som mulig samtidig som vi ønsker at kravene i kapittel 7.2 skal være tilfredsstillende. Anta nå at modellen vår for januar 1989 gir verdien 1148. Da har vi funnet en verdi på støyen ved dette tidspunktet. Vi sier da at residualen for januar 1989 er $1179 - 1148$ som er lik 31.

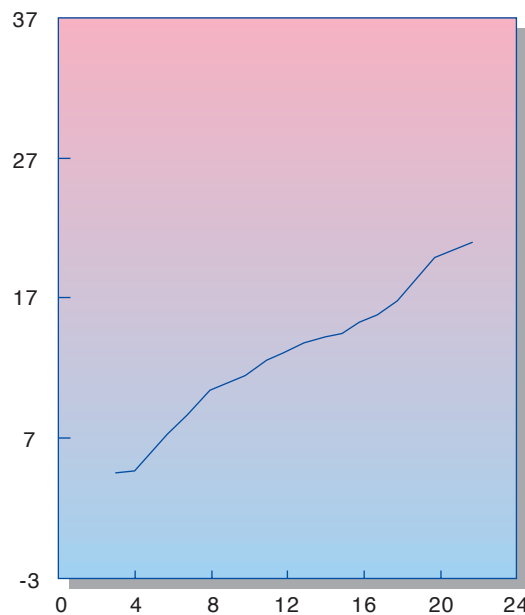
Vi vil naturlig nok få forskjellig verdi på residualene avhengig av hvilken modell



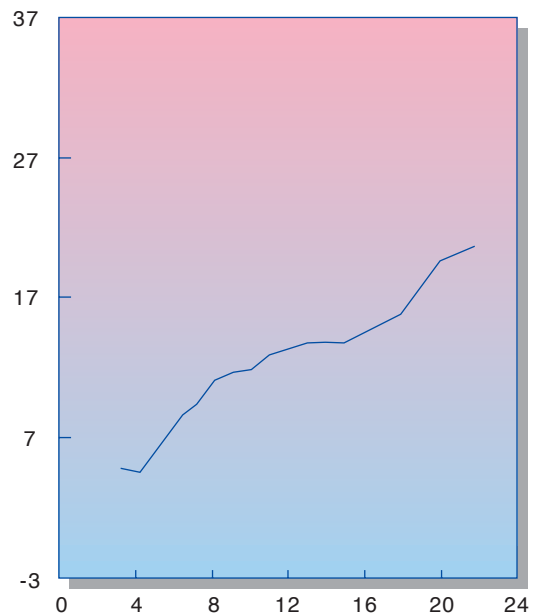
Figur 6.12 Glatting av tidsrekken med lineær trend og sesongkomponenter, uten støy fra figur 6.7



Figur 6.13 Glatting av tidsrekke 1 fra figur 6.9



Figur 6.14 Glatting av tidsrekke 2 fra figur 6.10



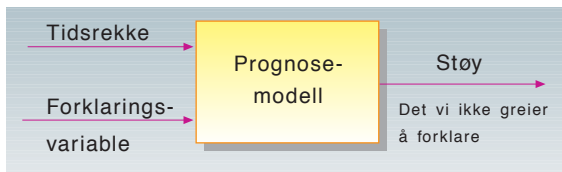
Figur 6.15 Glatting av tidsrekke 3 fra figur 6.11

som brukes. Dette er også vist i figur 7.4 og 7.5.

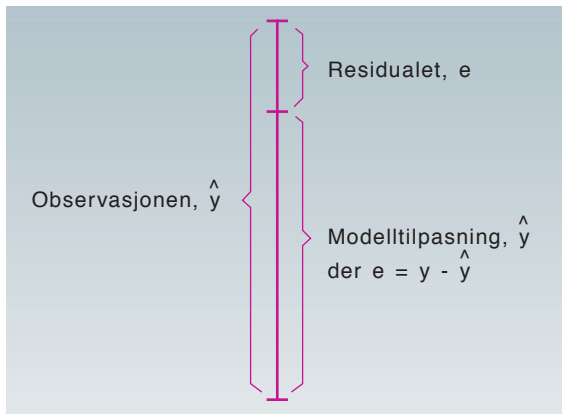
Residualen er definert som differansen mellom den observerte verdi og den verdi som modellen gir (modelltilpassningen). Residualen blir dermed realiseringen av støyen ut fra den modellen som brukes. Det er selvsagt viktig å velge eller arbeide seg fram til modeller som

gjør at støyen får de egenskapene som vi har beskrevet tidligere, og det er ved analyse av residualene – den realiserede støyen – vi kan undersøke om egenskapene tilfredsstilles.

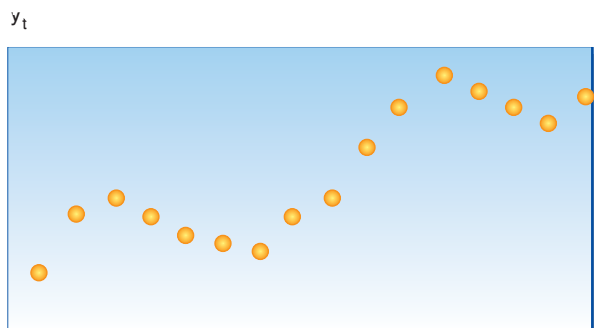
I figur 7.2 er det gitt en illustrasjon av residualen. Vi legger merke til at støyen betegnes med e , mens residualen betegnes med ϵ .



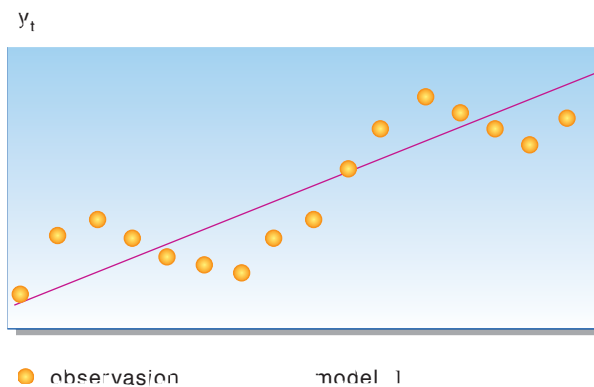
Figur 7.1 Modellering



Figur 7.2 Residualet, observasjonen og modelltilpasningen



Figur 7.3 En tidsrekke



Figur 7.4 Tidsrekken, modell 1 av tidsrekken og residualene (støyen)

Før vi går nærmere inn på de ulike teknikkene og testene, skal vi se nærmere på hvorledes selve modellbyggingen foretas ved introduksjon av et sett med figurer.

Figur 7.3 viser en tidsrekke som vi skal lage prognoser for. Første skritt blir da å lage en modell for tidsrekken. Dette er illustrert i figur 7.4. Vi har i første omgang laget en enkel modell av tidsrekken. Modellen er en rett linje – for eksempel en regresjonsmodell med tiden som forklaringsvariabel som også var omtalt i kapittel 5.8.

Vi registrerer at den rette linjen ikke greier å tilpasse tidsrekken vår på en fullgod måte. I figur 7.4 er også realisasjonen av støyen tegnet inn. Vi pleier å kalle disse for residualene som da er konkrete tall på de observerte støyleddene. Vi ser at disse er forholdsvis store. Det burde være mulig å lage en bedre modell.

En bedre modell er vist i figur 7.5.

Vi ser av figurene 7.4 og 7.5 at modell II gir en bedre tilpasning enn modell I på den aktuelle tidsrekken. Dermed blir støyleddene mindre. Spørsmålet er imidlertid om de krav vi har satt opp til støyleddet nå er tilfredsstillt?

Vi legger merke til at det fremdeles er slik at er det først et positivt residual, så er det stor sannsynlighet for at også neste residual er positivt; og det samme når det er negative residualer. Dette betyr at kravet om at støyleddene skal være uavhengige ikke er tilfredsstillt. Dette kan vi også påvise ved å analysere autokorrelasjonen. Dette er vist senere i kapitlet.

7.4 Plotting av residualene

Det finnes mange måter å plote residualene på. Vi skal ta for oss et par av mulighetene her. Vi ser på plott av residualene som funksjon av tiden (tidssekvensplott) og plott av residualene som fordeling.

Med plott av residualene som funksjon av tiden menes at ett og ett av residualene plottes fra tidspunkt 1, 2, ... etc. Residualene plottes altså i samme rekkefølge som observasjonene.

Med plott av residualene som fordeling, menes at residualene kan plottes i histogram slik som vist for rekrutter i kapittel 5 for om mulig å påvise at residualene er normalfordelt.

I figur 7.6 – 7.11 vises det en rekke ulike tidssekvensplott av residualene. Meningen med disse plottene er å illustrere hvorledes plott av residualene kan hjelpe oss i modellfasen til stadig å lage bedre modeller.

Av residualene i figur 7.6 ser vi at det er en systematisk feil i den modellen som er valgt. De første residualene er for små, de siste er for store. Feilene kan uttrykkes som en lineær trend. Det betyr at disse feilene delvis kan elimineres ved introduksjon av en lineær trend med tiden som forklaringsvariabel.

Det er klart at de modellene som er brukt i tilknytning til både figur 7.6 og 7.7 ikke er gode nok. Det er grunnlag for forbedringer av modellene, og tidssekvensplottet av residualene gir her verdifull informasjon.

Figur 7.8 viser et plott av residualer der det er en annen systematisk variasjon. Det viser at samme type avvik gjentar seg i faste perioder. Dette er sesongperioder. Det betyr at det i modellen ikke er tatt hensyn til sesongvariasjonene i modellen.

Figur 7.9 viser at variansen til støyen øker med tiden. Dette strider mot en av de forutsetninger som ble angitt i kapittel 7.2. Dette plottet forteller at det må foretas noe i modelleringsfasen for å endre dette forholdet. Vi vil senere i kapitlet vise hvorledes dette kan forbedres blant annet ved bruk av transformasjoner.

I figur 7.10 ser vi at mønsteret på residualene stort sett er tilfredsstillende, men at det er et stort avvik. Denne type avvik kalles outlier fordi det ligger unormalt langt unna et sannsynlig avvik. Når vi lager en modell uten å korrigere for slike unormale avvik, kan vi komme svært uheldig ut i modelleringsfasen fordi store outlierer kan influere meget sterkt på modellen, slik at vi kan få en modell som på en dårlig måte beskriver den faktiske tidsrekken.

En del prognosemodeller og statistiske modeller generelt baserer seg på et prinsipp som heter minste kvadraters metode. Når denne metoden benyttes, blir modellen bestemt ved at summen av kvadratet av residualene er minst mulig. Sagt på en annen måte: Vi tar det første residualet og kvadrerer det, deretter kvadreres det neste osv, til alle residualene er kvadrert. Deretter summeres dette.

Når vi bruker minste kvadraters metode bestemmes modellen ut fra at summen av kvadratet av residualene er minst mulig.

Når vi bruker denne type metoder er det derfor viktig å gå nøye igjennom residualene for å undersøke om det forekommer outliere fordi disse vil påvirke modellen i sterk grad. Dette skyldes at når selve residualen er stort, vil kvadratet av residualen bli dominerende og dermed tvinge fram en modell som gir en dårlig beskrivelse av tidsrekken. Dette er illustrert i figur 7.11 og 7.12.

Vi må derfor være varsomme når vi oppdager spesielt store avvik i residualene. I første omgang bør det stilles spørsmål om hvorfor de er så store. Det kan skyldes feilmålinger, annen type registreringsfeil, eventuelle misforståelser/kommunikasjonssvikt, innleggingsfeil av data, etc. I slike tilfeller må dataene rettes slik at de blir korrekte.

En outlier kan også skyldes spesielt oppståtte, men unormale situasjoner. Det kan eksempelvis bli registrert lav etterspørsel på grunn av en streik e.l. I slike tilfeller er dataene korrekte, men de gir ikke noe representativt bilde av situasjonen. Det vil ikke være riktig å bruke denne observasjonen som grunnlag for videre prognostisering når det ikke vil være noen streik de kommende perioder. Dersom vi imidlertid vet at det kommer en streik, må vi selvsagt ta hensyn til det i prognosene.

I figur 7.11 og 7.12 er det vist hvorledes en stor outlier påvirker modellen og dermed også prognosene. Vi har i figur 7.11 en enkel regresjonsmodell med tiden som forklaringsvariabel og hvor det er en spesielt stor outlier i begynnelsen av observasjonene. Vi har i figuren foretatt en modelltilpasning uten at det er korrigert for outlieren. I figur 7.12 har vi de samme observasjonene, men her er det korrigert for outlieren, noe som må gjøres under modellbyggingen. Vi ser nå at vi får en helt annen modell og dermed også helt andre prognoser.

I noen tilfeller kan vi klart se at vi har en outlier fordi den ligger svært langt borte fra det som er normalt. I andre tilfeller kan det være vanskeligere å avgjøre om vi har en outlier eller ikke. Noen programpakker framstiller tidssekvensplottet av residualene med et usikkerhetsintervall som er lik to ganger standardavviket til residualene. Dette er vist i figur 7.13.

I kapittel 5.6.2 ble egenskaper til normalfordelingen gjennomgått. Vanligvis vil støyen (residualene) tilnærmet tilfreds-

stille en normalfordeling. Da vil omlag 5 % av residualene ligge utenfor to ganger standardavviket, mens kun 0,3 % av residualene vil ligge utenfor tre ganger standardavviket.

Dette betyr at dersom ett av residualene avviker i størrelsesorden tre ganger standardavviket eller mer, er det grunn til å stille spørsmålsteget ved den observasjonen som har generert residualen. Denne må karakteriseres som en outlier.

Det siste som bør ses på, er i hvilken grad residualene tilfredsstillende normalfordelingen. Dette kan gjøres på flere måter. I noen programpakker er det lagt inn et spesielt normalfordelingsplott. Dette kan da benyttes. Plottet gjøres på en spesiell form slik at når residualene ordnes i stigende rekkefølge skal de tilnærmet ligge på en rett linje. Skalaen i plottet er transformert slik som vist i figur 7.14.

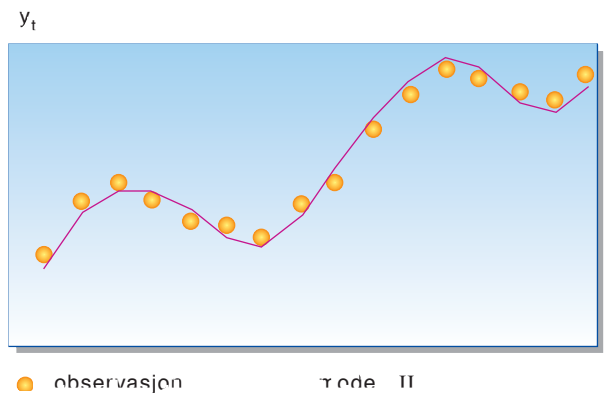
Det er også mulig å plote residualene slik som vist i figur 5.2 for å få et inntrykk av fordelingen til residualene. Det anbefales imidlertid å benytte et normalfordelingsplott dersom dette er tilgjengelig i programpakken.

Det bør også sies at forutsetningen om normalfordelte støyledd ikke er det mest kritiske. Det er fullt mulig å foreta mesteparten av modellbyggingen uten forutsetning om at støyleddene er normalfordelte.

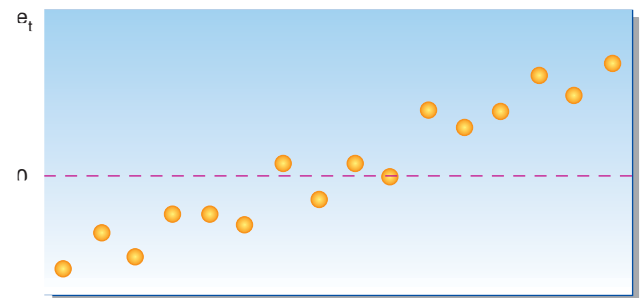
Vi har nå vært igjennom en del ulike mønstre som vi kan ha på plottet av residualene. Vi forstår nå hvor viktig det er å foreta slike plott etter at vi har laget en modell. Figur 7.1 beskriver nettopp dette. På grunnlag av dataene lager vi en modell, deretter sammenlikner vi observasjonene i tidsrekken med de tilpassede dataene fra modelleringen. Differansen mellom disse er residualene som så plottes ut. Dersom disse ikke tilfredsstillende krav som er gitt i kapittel 7.2 for støyen, må vi prøve å finne fram til en bedre modell.

Selv om vi ennå ikke har beskrevet de aktuelle prognosemetodene, har vi så langt fått et godt grunnlag for hvorledes det må arbeides med modellbyggingen for å få fram gode prognosemodeller.

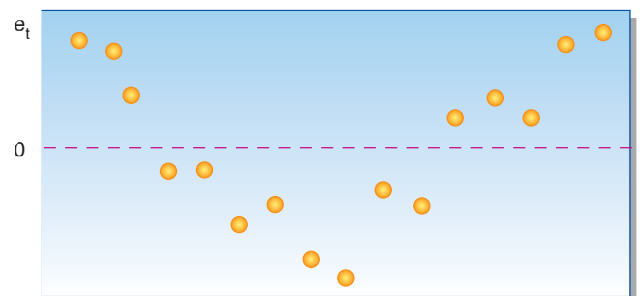
Det som vi skal stå tilbake med til slutt er et residualplott slik som vist i figur 7.15, der residualene er tilfeldig



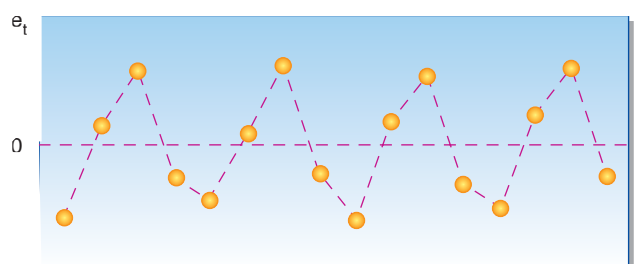
Figur 7.5 Tidsrekke, modell II av tidsrekken og residualene (støyen)



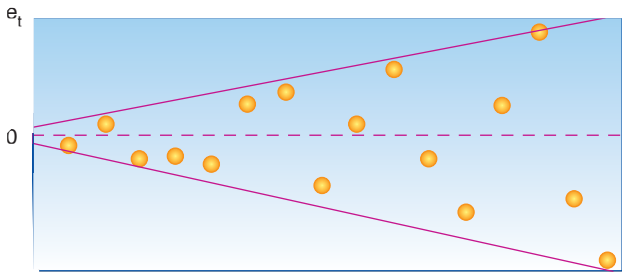
Figur 7.6 Plott av residualer fra en modell som bør ha en trend



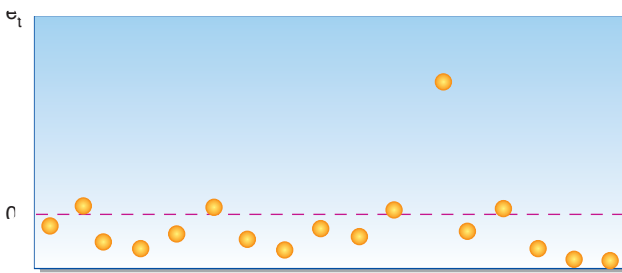
Figur 7.7 Plott av residualer fra en modell som gir dårlig tilpasning



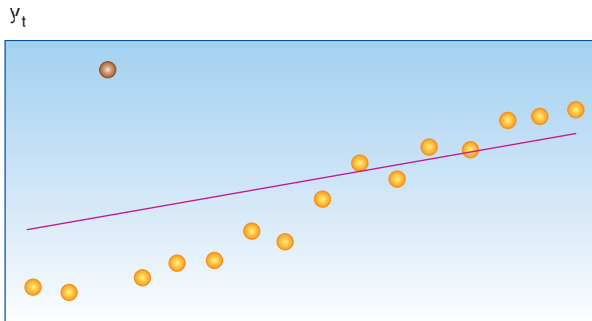
Figur 7.8 Plott av residualer fra en modell som ikke har bygget inn sesongvariasjoner



Figur 7.9 Plott av residualer med økende varians

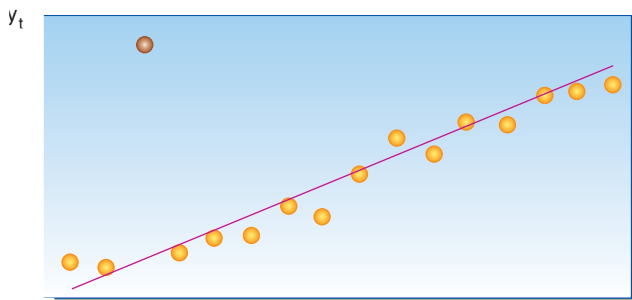


Figur 7.10 Plott av residualer med en outlier



● observasjon ● outlier — kode

Figur 7.11 Modell der det ikke er tatt hensyn til outlier



● observasjon ● outlier — modell

Figur 7.12 Modell der det er tatt hensyn til outlier

fordelt slik at de tilfredsstillende krav.

7.5 Transformasjoner

Det første som bør gjøres før vi starter med å analysere og modellere en tidsrekke, er å studere forløpet til tidsrekken. Vi får da et visuelt bilde av tidsrekken som kan fortelle mye.

Spesielt må vi være oppmerksom på om tidsrekken varierer med tiden. Dersom det er slik at selve tidsrekken øker med tiden samtidig som variasjonen øker med tiden, er det grunnlag til å anta at forutsetningen om at variansen til tidsrekken støy er konstant, ikke er tilfredsstillende.

Dette undersøkes ved et variasjons-gjennomsnittsplott.

Observasjonene deles da opp i grupper. I hver gruppe foretas det en beregning av gjennomsnittet og variasjonen på observasjonene. Variasjonen defineres her som differansen mellom laveste og høyeste observasjon i gruppen. Disse verdiene plottes så i en todimensjonal figur der gjennomsnittet er angitt på den ene aksene og variasjonen på den andre aksene.

Dette illustreres nå på et sett med kvartalsvise observasjoner.

Totalt har vi 40 observasjoner over en 10-årsperiode. Gjennomsnittet for hvert år og variasjonen (differansen mellom høyeste og laveste verdi) er angitt til høyre i tabellen.

Det er viktig at de gruppene som det regnes ut gjennomsnitt og variasjon for, er observasjoner over en hel sesong, slik at vi får med alle variasjonene over året. Dersom det er kvartalsvise observasjoner blir det 4 observasjoner i hver gruppe. Dersom det er månedlige observasjoner blir det 12 observasjoner i hver gruppe. Observasjonene må grupperes på denne måten, ellers kan vi få uheldige grupperinger. Anta eksempelvis at vi i tabell 7.1 grupperte observasjonene i grupper på 3. Vi ser av tabellen at observasjonene i kvartal 2 er høyere enn for de øvrige månedene. Dersom vi grupperer observasjonene 3 og 3, vil vi i noen av gruppene

ikke ha observasjoner fra kvartal 2. Dette forårsaker at variasjonen i gruppen blir betraktelig lavere – noe den ikke skulle vært. Dette skyldes imidlertid en uheldig gruppering av observasjonene.

Vi kan nå se nærmere på variasjons-gjennomsnittsplottet for observasjonene i tabell 7.1.

Vi ser av figuren at når gjennomsnittet øker, så øker også variasjonen i observasjonene. Dette er også illustrert ved den linjen som er trukket opp i figuren.

Dette betyr at forutsetningen om konstant varians i tidsrekken ikke er tilfredsstillende. Det er derfor grunnlag for gjøre noe med tidsrekken for bedre å kunne imøtekomme kravet.

Det som vanligvis gjøres, er å transformere tidsrekken for å dempe variasjonen. Det finnes en rekke transformasjoner som kan brukes. De viktigste transformasjonene er den logaritmiske transformasjonen og kvadratrots-transformasjonen. Det finnes for øvrig en klasse transformasjoner kalt Box-Cox transformasjoner, men i dette kompendiet vil vi begrense oss til de mest vanlige transformasjonene.

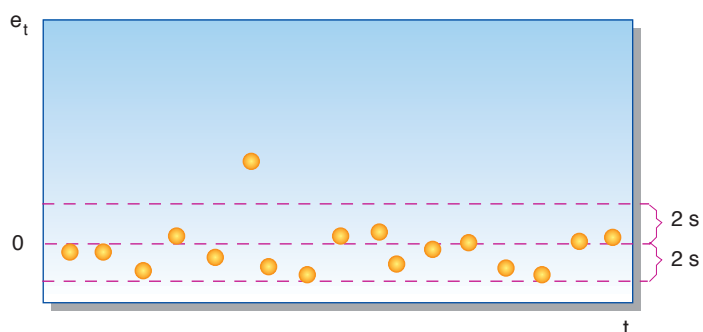
La oss nå foreta en logaritmisk transformasjon av de kvartalsvise observasjonene i tabell 7.1. Vi får da:

$\ln 35 = 3,56$
$\ln 52 = 3,95$
$\ln 28 = 3,33$
$\ln 33 = 3,50$
$\ln 38 = 3,64$
.
.
$\ln 42 = 3,74$

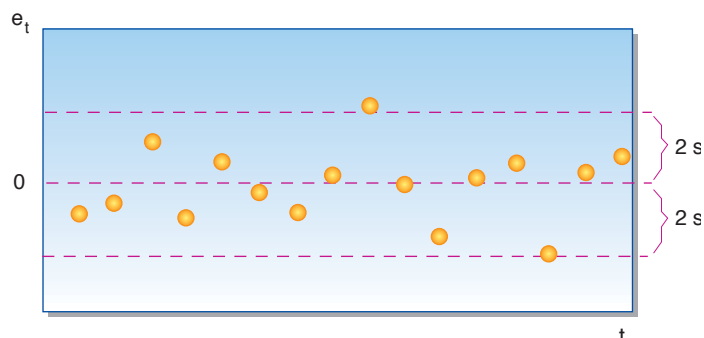
Den originale tidsrekken og den transformerte tidsrekken er vist i figur 7.16 og 7.17.

Det ses av figur 7.17 at variasjonen i tidsrekken øker med økende nivå i tidsrekken. I figur 7.18 er tidsrekken transformert ved bruk av logaritmisk transformasjon, og vi ser at variasjonen i tidsrekken er dempet. I dette tilfellet har den logaritmiske transformasjonen hatt en god effekt.

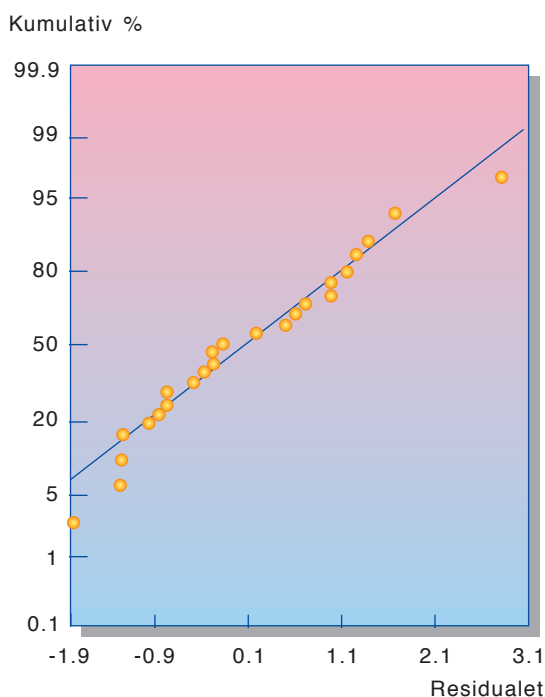
Til slutt vises det eksempel på et variasjons-gjennomsnittsplott i figur 7.19, der variasjonen ikke øker med gjennomsnittet. Dersom punktene som er avmerket i figuren ikke varierer etter et mønster, men er mer eller mindre tilfeldig fordelt slik som vist på figuren, er



Figur 7.13 Plott av residualer med outlier og to ganger standardavvik som usikkerhetsintervall



Figur 7.15 Residualplott der støyen er tilfeldig fordelt



Figur 7.14 Normalfordelingsplott

det ikke grunnlag for å foreta transformasjon av tidsrekken.

Det bør imidlertid nevnes at ikke alle prognosespesialister er for å transformere tidsrekken under modellbyggingen. Dersom dette gjøres må all modellbygging foregå med de transformerte observasjonene. Når så analysene, modellen og prognosene er laget, transformeres det tilbake, og vi får de originale prognosene. Under tilbake-transformasjon mistes noen statistiske egenskaper. Det er derfor en avveining om en skal transformere tidsrekken eller ikke. I dette kompendiet anbefales det transformasjoner av tidsrekken dersom variasjonen i tidsrekken øker markant med økende nivå på tidsrekken.

7.6 Autokorrelasjon

Ett av de viktige kriterier for støyen er at støyleddene skal være ukorrelerte – det vil si at det ikke er autokorrelasjon i støyleddene. Dette betyr at det ikke skal være avhengighet mellom støyleddene.

Hvordan opptrer avhengighet i en serie med observasjoner? Dette kan beskrives ved at størrelsen av en observasjon påvirker størrelsen på neste observasjon.

Hvis vi har en positiv avhengighet eller positiv korrelasjon, vil en vekst i en observasjon påvirke til vekst i den andre observasjonen. Analogt vil en avtagning av en observasjon påvirke til avtagning i den andre observasjonen.

Dersom vi har negativ avhengighet eller negativ korrelasjon, vil en vekst i en observasjon påvirke til avtagning i den andre observasjonen og vice versa.

Vanligvis brukes begrepet korrelasjon, men hvis de observasjonene som vi analyserer ligger på tidsaksen – det vil si at de er observert etter hverandre i tid, benyttes isteden begrepet autokorrelasjon.

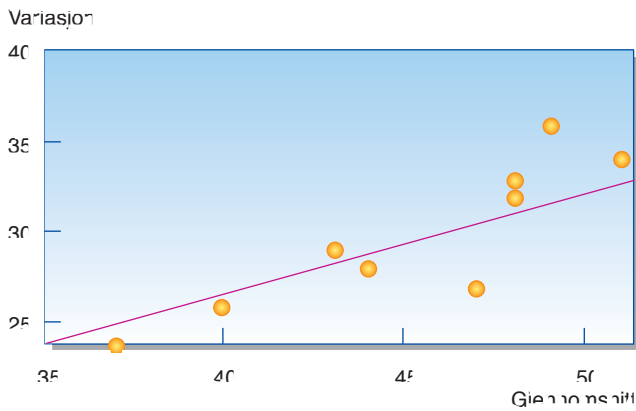
Hvis vi ønsker å analysere avhengigheten mellom to på hverandre følgende observasjoner, så uttrykker vi det som autokorrelasjonen i avstand 1. Hvis vi ønsker å analysere avhengigheten mellom to observasjoner som ligger i avstand to fra hverandre, så uttrykker vi det som autokorrelasjon i avstand 2, etc. Den rekken med autokorrelasjoner som vi får i avstand 1, avstand 2, avstand 3, ..., kalles autokorrelasjonsfunksjonen.

Vi vil i det følgende forutsette at tidsrekken fluktuierer rundt en gjennomsnittsverdi.

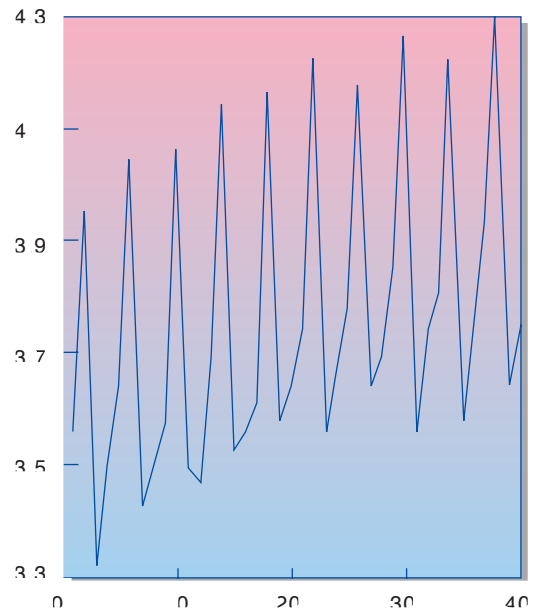
Den beregnede autokorrelasjonsfunksjonen er gitt ved:

Tabell 7.1 Gjennomsnitt og variasjon på et sett med kvartalsvise observasjoner som er delt opp i grupper på 4

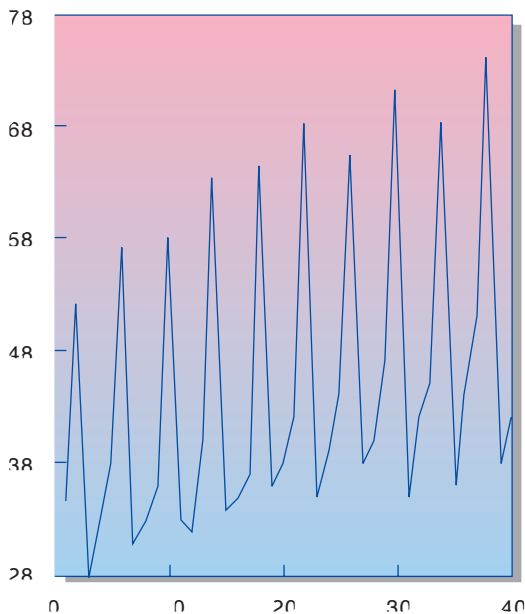
År	Kvartal1	Kvartal2	Kvartal3	Kvartal4	Gj.snitt	Variasjon
1981	35	52	28	33	37	24
1982	38	57	31	33	40	26
1983	36	58	33	32	40	26
1984	40	63	34	35	43	29
1985	37	64	36	38	44	28
1986	42	68	35	39	48	33
1987	44	65	38	40	47	27
1988	47	71	35	42	49	36
1989	45	68	36	44	48	32
1990	51	74	38	42	51	34



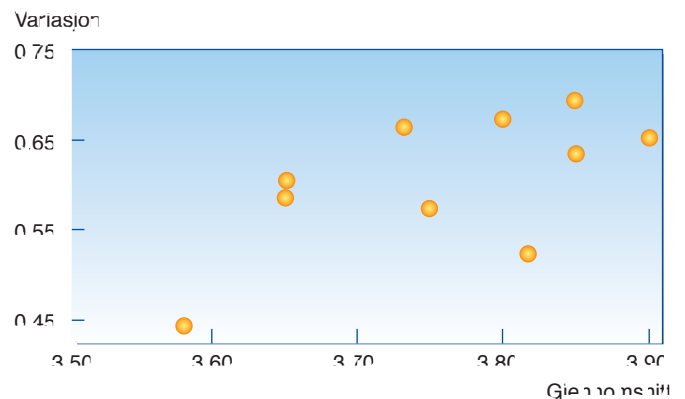
Figur 7.16 Variasjons-gjennomsnittsplott for tidsrekken i tabell 7.1



Figur 7.17 Den originale tidsrekken



Figur 7.18 Den transformerte tidsrekken



Figur 7.19 Variasjons-gjennomsnittsplott for den logaritmisk transformerte tidsrekken

$$r_k = \frac{c_k}{c_0} \quad k = 0, 1, 2, 3, \dots \quad (7.4)$$

der

$$c_k = \frac{1}{n} \sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y}) \quad (7.5)$$

Vi beregner så autokorrelasjonsfunksjonen – det vil si autokorrelasjonene – til en tidsrekke med flere observasjoner. Vi tar utgangspunkt i tabell 7.1 og tar med de 12 første observasjonene. Disse er vist grafisk i figur 7.20.

Vi beregner så autokorrelasjonene til denne tidsrekken. Denne gangen gjør vi det ikke for hånd, men vi lar en program-pakke beregne autokorrelasjonene.

I figur 7.21 finner vi de beregnede autokorrelasjonene.

Vi gjentar denne prosessen og beregner autokorrelasjonsfunksjonen til tidsrekken som består av samtlige 40 observasjoner fra tabell 7.1. Tidsrekken er framstilt i figur 7.17. Den beregnede autokorrelasjonsfunksjonen er gitt i figur 7.22.

Hvis vi ser på de 4 første autokorrelasjonene, så ser vi klare likhetspunkter i figur 7.21 og figur 7.22. Dette betyr at vi har større usikkerhet i beregningene når vi har få observasjoner i forhold til mange observasjoner.

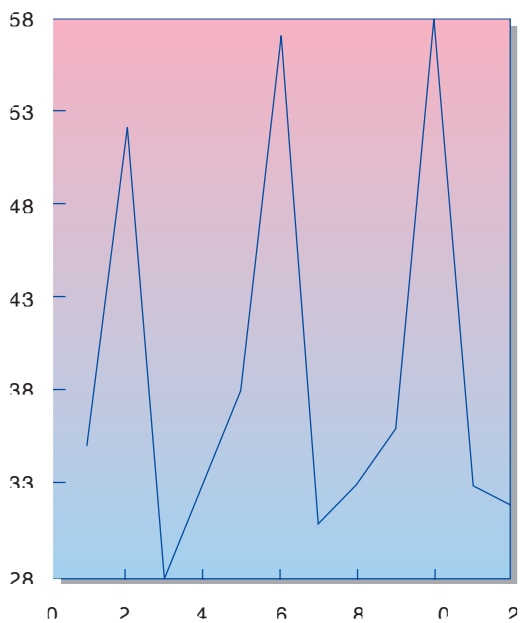
I figurene 7.21 og 7.22 er det tegnet opp usikkerhetsgrenser. Vi ser at disse avtar med økende antall observasjoner. Usikk-

erhetsgrensene er basert på følgende tilnærmede formel:

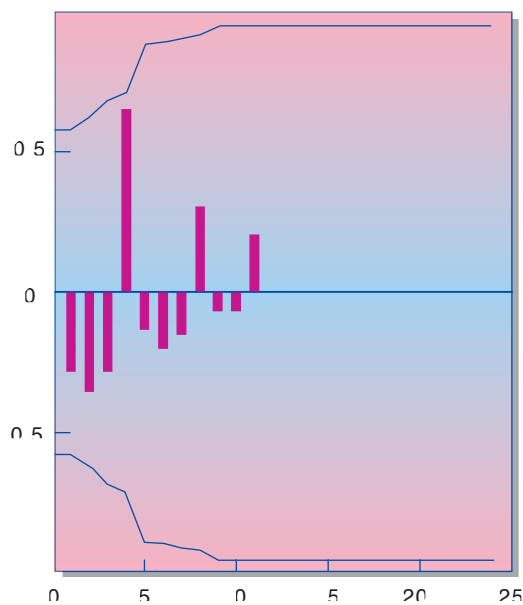
$$\text{var}(r_k) = \frac{1}{n} \left(1 + 2 \sum_{i=1}^q r_i^2 \right) \quad (7.6)$$

Her er q avstanden der det er autokorrelasjon mellom observasjonene. Når vi skal analysere støyleddene, er målsettingen at det ikke skal være autokorrelasjon mellom dem. I så fall blir $q = 0$, og vi får følgende tilnærmede formel for variansen til autokorrelasjonene:

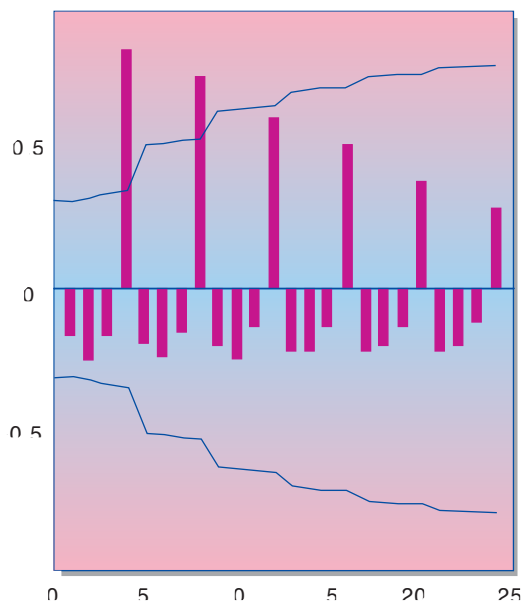
$$\text{var}(r_k) = \frac{1}{n} \quad (7.7)$$



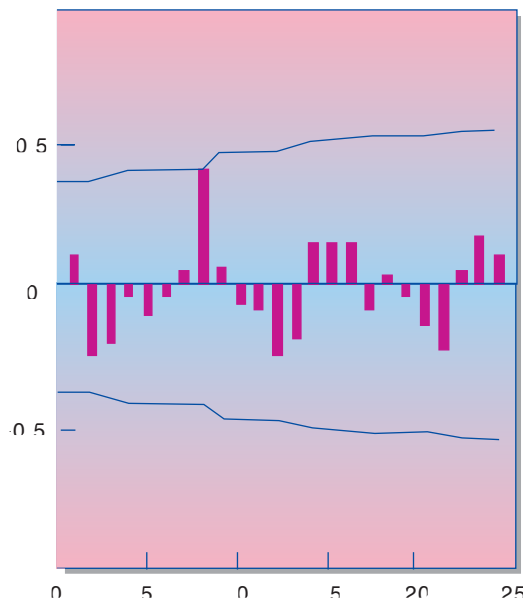
Figur 7.20 En tidsrekke med 12 observasjoner



Figur 7.21 Beregnet autokorrelasjonsfunksjon for tidsrekken med 12 observasjoner



Figur 7.22 Beregnet autokorrelasjonsfunksjon for tidsrekken med 40 observasjoner



Figur 7.23 Beregnet autokorrelasjonsfunksjon for støyleddene i en enkel sesongmodell

Som vi husker fra kapittel 5 er standardavviket lik kvadratroten av variansen. En tilnærmet 95 % usikkerhetsgrense eller konfidensintervall er da gitt som to ganger standardavviket som i dette tilfellet er 1 dividert med kvadratroten av det antall observasjoner vi har.

Dersom vi eksempelvis har 36 støyledd ($n = 36$) og antar at det ikke er autokorrelasjon i støyen, så vil et tilnærmet 95 %

konfidensintervall være lik pluss minus $2 * 1/6 = 0,33$. Dersom det imidlertid er autokorrelasjon i støyleddene, vil usikkerhetsgrensen bli enda høyere.

La oss igjen se på figur 7.23 som angir den beregnede autokorrelasjonsfunksjonen for tidsrekken med 40 observasjoner. Vi ser at det er meget høye utslag på autokorrelasjonen i avstand 4, 8, 12, etc. Dette betyr at vi har en tidsrekke som

er sesongbetont med en sesongperiode på 4. Så langt har vi påstått at vi har følgende modell:

$$y_t = e_t$$

Dette er helt klart en dårlig modell fordi vi får store avhengigheter i støyleddene som nå er identisk med observasjonene.

Vi foreslår nå en modell som sier at observasjonen i dette kvartalet er den

samme som observasjonen i forrige kvartal pluss støyleddet. Vi får da:

$$y_t = y_{t-4} + e_t$$

Figur 7.23 viser autokorrelasjonsfunksjonen til støyleddene i denne enkle sesongmodellen.

Vi ser nå at de sterke sesongmessige autokorrelasjonene er fjernet. I tillegg ser vi at usikkerhetsgrensene har avtatt fordi vi har fjernet en del autokorrelasjon i støyleddene.

Det er fullt mulig å arbeide videre for å finne en enda bedre modell. Dermed vil kravene til støyleddene tilfredstilles ytterligere.

Under modellbyggingen kan det være hensiktsmessig å differensiere tidsrekken på sesong slik som vist her. Det kan også være hensiktsmessig å differensiere i avstand 1. Det vil si:

$$y_t = y_{t-1} + e_t$$

For øvrig bør det tas med parametere i modellen som gjør tilpasningen til de aktuelle observasjonene enda bedre. Autokorrelasjonsfunksjonen til støyleddene – den variasjonen som er tilbake – skal da i utgangspunktet illustrere tilfeldig støy.

8 Regresjonsmodeller

I de kommende kapitlene vil følgende prognosemodeller bli omtalt:

- Regresjonsmodeller
- Metningsmodeller
- Glattingmodeller
- Tidsrekkemodeller
- Kalmanfiltermodeller.

Det vil bli pekt på spesielle egenskaper til de ulike modellene, uten at det går ned på spesielt detaljert nivå.

Regresjonsmodeller er en meget benyttet type modeller som brukes til prognostisering. I mange tilfeller inkluderes det økonomiske variable i disse modellene for å kunne modellere etterspørselen. For nærmere studier i regresjon henvises det til [9], [14], [15], [16], [17] og [53].

8.1 Enkel regresjon

Det er vanlig å skille mellom enkel regresjonsmodell og multiplere regresjonsmodell. I den enkle regresjonsmodellen angis det en relasjon mellom to variable, mens

i den multiple regresjonsmodellen angis det relasjoner mellom flere variable.

La y være den variable som det skal lages prognoser for. Den andre variable kan eksempelvis ha betegnelsen t . Vi kaller gjerne variabelen t for forklaringsvariabelen fordi det er denne variable som forklarer utviklingen av den andre variabelen y .

For å konkretisere ser vi på et eksempel. I tabell 8.1 har vi angitt abonnementsstettheten for telefon i Norge fra 1975 til 1989. Abonnementsstettheten er definert som antall abonnement pr. 100 innbyggere. Variabelen y står i tabellen for abonnementsstettheten, mens variabelen t , som er tiden i tabellen, er forklaringsvariabelen.

I dette tilfellet vil regresjonsmodellen, som angir sammenhengen mellom utvikling i abonnementsstettheten og tiden, være gitt ved:

$$y = a + bt + \varepsilon \quad (8.1)$$

der a og b er parametere i modellen, mens e er støyleddet. Vanligvis brukes det greske bokstaver for parametrene.

I figur 8.1 er regresjonslinjen tegnet opp og de ulike verdiene på abonnementsstettheten som funksjon av tiden. Vinkelen på regresjonslinjen er den estimerte verdien på b , mens avstanden fra origo og opp til regresjonslinjen når tiden $t = 0$, er den estimerte verdien på a . Hvordan blir så verdiene på a og b beregnet eller estimert?

Det gjøres ved en bergingsmetode som heter *minste kvadraters metode*. Denne statistiske metoden bestemmer verdien på a og b slik at summen av kvadratavstanden fra punktene y til regresjonslinjen blir minimert. Populært kan vi si at verdien på a og b bestemmes slik at regresjonslinjen går på best mulig måte mellom punktene.

Denne regresjonslinjen blir da vår *regresjonsmodell*. Ved å forlenge regresjonslinjen lineært fra $t = 15$ og framover, får vi prognoser for abonnementsstettheten.

I figur 8.1 ser vi også noen stiplede linjer. Det er konfidensintervall eller usikkerhetsgrenser. Disse intervallene sier hvor sikker beregningen av parametrene a og b er. I dette tilfellet ser vi at de estimerte verdiene av parametrene er svært gode fordi de angitte intervallene gir liten frihetsgrad til å endre på vinkelen på kurven og høyden på kurven innen de intervaller som er satt opp. Forlengelse av det ytterste intervallet som vi

ser krummer noe, angir konfidensintervallet for prognosen. Krummingen betyr at usikkerheten i prognosene øker som funksjon av tiden.

I den angitte regresjonsmodell er det en lineær sammenheng mellom y og forklaringsvariabelen t . Det kan også tenkes helt andre sammenhenger. Et eksempel er:

$$y = a + b^t + \varepsilon \quad (8.2)$$

Her vil vi ha en eksponentiell økning som funksjon av tiden. En slik utvikling kan være riktig over en periode, men så vil stigningsgraden etter hvert avta (da intet vokser inn i uendelig).

8.2 Multiplere regresjon

I eksemplene så langt har tiden t vært forklaringsvariabelen. I de fleste regresjonsmodeller vil vi ha supplerende eller andre forklaringsvariable. Det som ubetinget er en fordel med å ha med tiden som forklaringsvariabel, er at utviklingen til tiden er helt deterministisk og uten usikkerhet. Benytter vi andre forklaringsvariable, er det viktig å være klar over at når det skal lages prognoser, må det også lages prognoser for utviklingen av disse forklaringsvariable som da i seg selv vil være beheftet med usikkerhet. Dette induserer en ekstra usikkerhet i prognosene. Eksempler på slike forklaringsvariable er pris for tjenesten og konsumprisindeksen.

Som nevnt kalles regresjonsmodeller med flere forklaringsvariable for multiple regresjonsmodeller. Et eksempel på slike modeller er:

$$y = a + bu + cv + dw + \varepsilon \quad (8.3)$$

Her er u , v og w tre forklaringsvariable og a , b , c og d parametrene i den multiple regresjonsmodellen. På samme måte som i den enkle regresjonsmodellen skal nå parametrene bestemmes på best mulig måte. Dermed brukes minste kvadraters metode som gir optimale verdier på parametrene.

I selve utviklingen av prognosemodellen går mye av arbeidet med til å bestemme hvilke variable som bør tas med og hvilke variable som bør utelates fra modellen. I utviklingen benyttes de samme prinsippene som i forrige kapittel. I tillegg benyttes spesielle teknikker for regresjonsanalyse. *Den multiple korrelasjonskoeffisient* er viktig med hensyn til å sammenlikne de ulike modellene. Denne størrelsen beregner i hvor stor grad regresjonsmodellene greier å forklare varia-

sjonene eller utviklingen til y . Valget mellom de ulike regresjonsmodeller med ulike forklaringsvariable baseres blant annet på hvor god den multiple korrelasjonskoeffisient er.

Det finnes også metoder som *trinnvis regresjon* som suksessivt beregner seg fram til den beste regresjonsmodellen ut fra et sett med forklaringsvariable. Den velger da ut hvilke forklaringsvariable som skal tas med i modellen.

Det ses nå på en problemstilling der vi har flere forklaringsvariable. Det skal lages en prognosemodell for årlig etterspørsel etter hovedabonnement for telefon. Tabell 8.2 viser etterspørselen.

Det ses at den årlige etterspørselen har økt fram til 1981, deretter har den avtatt. Spørsmålet er så hva slags forklaringsvariable som bør inkluderes i regresjonsmodellen. Etterspørselen vil være avhengig av prisen på tjenesten. Aktuelle forklaringsvariable er:

- Kvartalsavgift
- Innmeldingsavgift
- Teletakstindeks
- Tiden.

Teletakstindeks var tidligere en sentral indeks for telefontjenesten som ble beregnet av Økonomiavdelingen. Tabell 8.3 angir verdien på disse forklaringsvariable.

Den taksten som er angitt her, er ikke justert for den generelle prisutvikling. Dette kunne vært gjort ved å dividere de aktuelle forklaringsvariable med konsumprisindeksen. Dette ville også være relevante regresjonsmodeller som må prøves ut i en evalueringsfase.

Ved å benytte en trinnvis regresjon, ble resultatet at: Tiden, Kvartalsavgiften og Innmeldingsavgiften gav den beste regresjonsmodellen. Dette gav følgende regresjonsmodell:

$$\begin{aligned} & \text{Abonnementsetterspørsel} \\ & = 175565 \\ & + 10340 * \text{tid} \\ & + 17,95 * \text{innmeldingsavgift} \\ & + 127,25 * \text{kvartalsavgift} \quad (8.4) \end{aligned}$$

De tallene som vi har i likning (8.4) er nå verdiene på de estimerte parametrene a , b , c og d . Tilpasningen her må sies å være brukbar med en multiplert korrelasjonskoeffisient på 0,94 og Durbin Watson observator på 1,79 som indikerer liten autokorrelasjon i residualene.

Så langt har dette eksemplet vært en god reklame for regresjonsmodeller. Siden vi

har brukt statistikk fram til 1984, kjenner vi resultatene videre. Utviklingen fra 1985 til 1988 var som vist i tabell 8.4 for forklaringsvariablene.

Vi legger merke til at tiden $t = 1$ i 1976 og tiden følgelig har verdi 10 i 1985, etc. Vi kan nå sette verdien til forklaringsvariablene fra 1985 til 1988 inn i likning (8.4). Vi får da prognoser for 1985–1988. Dette er imidlertid en tildragelse, fordi vi egentlig ikke kjenner verdien på de aktuelle forklaringsvariable – unntatt tiden. Det betyr at den prognosen som nå lages har mindre usikkerhet enn det som er realistisk. Ser vi på de aktuelle forklaringsvariable, så er innmeldingsavgiften uforandret unntatt i 1988, da den ble redusert med 25 %. Kvartalsavgiften har hatt en noe rar stigning: Uforandret fra 1985 til 1986, for deretter å stige. Noe av denne policyen kunne nok vært avdekket før en lagde prognosene.

Resultatet av prognosene er nå gitt sammen med de faktiske verdier i tabell 8.5.

Tabellen viser at den faktiske etterspørselen når en topp i 1985 for så å falle dramatisk etterpå. Årsaken til dette fallet er at markedet på hovedabonnement nærmer seg metning. Dermed avtar også den årlige etterspørselen. Dette ble også forsterket av en depresjon i næringslivet fra 1988.

Konklusjonen fra dette eksempelet er *ikke at regresjonsmodeller er ubrukbare prognosemodeller*, men at det er viktig å sette seg nøye inn i et produkts og en tjenestes utvikling når det skal lages prognoser. De prognosene som er laget her, har ikke noe med Televerkets offisielle prognoser å gjøre. De er vist som en argumentasjon for neste kapittel der vi skal se på metningsmodeller, som er en meget aktuell modelltype for prognostisering.

9 Metningsmodeller

9.1 Generelt

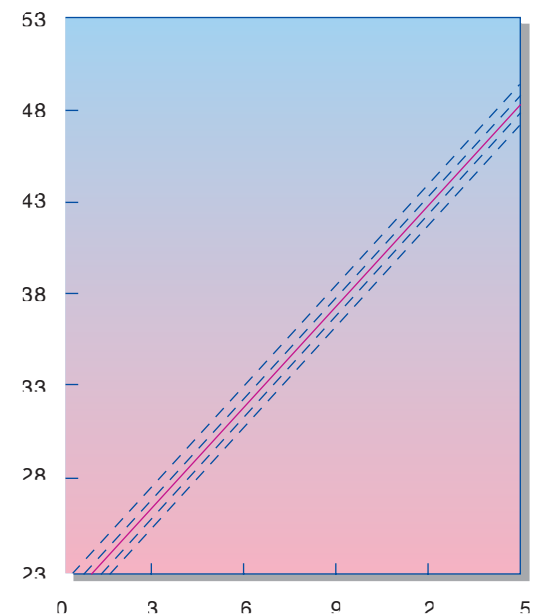
Det finnes mange navn på disse modellene. I litteraturen kalles de både metningsmodeller, vekstmodeller, diffusjonsmodeller og S-kurver. Modellene kjennetegnes ved at de har en forklaringsvariabel som er tiden t . Modellene inneholder fra 2 til 4 parametere. I en mer komplisert form kan modellene også inneholde flere forklaringsvariable.

Modellene beskriver et forløp der etterspørselen til å begynne med øker til-

nærmet eksponentielt for så i avslutningsfasen, når det nærmer seg metning, å avta tilnærmet eksponentielt. Modellen baserer seg på akkumulerte data. Det betyr at inngangsdataene eksempelvis er totaletterspørsel etter abonnement og

Tabell 8.1 Abonnementstetthet for telefon 1975–1989

Tid	År	Abonnementstetthet
1	1975	23,60
2	1976	24,87
3	1977	26,38
4	1978	27,95
5	1979	29,75
6	1980	31,48
7	1981	33,64
8	1982	35,55
9	1983	37,43
10	1984	39,27
11	1985	41,58
12	1986	43,83
13	1987	45,67
14	1988	47,14
15	1989	48,23



Figur 8.1 Regresjonsmodell for abonnementsstetthet for telefon med tiden som forklaringsvariabel

Tabell 8.2 Årlig etterspørsel etter hovedabonnement for hele landet

Tid	År	Årlig etterspørsel
1	1976	55644
2	1977	65028
3	1978	67739
4	1979	77145
5	1980	74999
6	1981	93170
7	1982	84116
8	1983	81932
9	1984	80407

Tabell 8.3 Forklaringsvariable for etterspørselen etter hovedabonnement

År	Teletakstindeks	Kvartalsavgift	Innmeldingsavgift
1976	59,1	574	3000
1977	62,9	622	3000
1978	66,7	682	3000
1979	66,7	682	3000
1980	71,3	765	3000
1981	77,5	849	3000
1982	91,5	1061	2000
1983	97,0	1123	2000
1984	100,0	1171	2000

Tabell 8.4 Forklaringsvariable for abonnementssetterspørsel 1985–1988

År	Kvartalsavgift	Innmeldingsavgift	Tid
1985	1171	2000	10
1986	1171	2000	11
1987	1254	2000	12
1988	1372	1500	13

Tabell 8.5 Prognostisert og realisert etterspørsel etter hovedabonnement 1985–1988

År	Faktisk etterspørsel	Prognostisert etterspørsel
1985	101149	94946
1986	100078	105286
1987	83629	104964
1988	68505	108364

ikke årlig etterspørsel (tilvekst) av abonnement.

Det finnes ulike typer vekstmodeller. Den mest kjente er den Logistiske modellen. Det er også den vi vil gjennomgå i størst detalj.

Metningsmodellene benytter tiden som forklaringsvariabel på en slik måte at modellen over en lengre periode beskriver et S-forløp.

Spørsmålet er så når metningsmodellene bør brukes?

Metningsmodellene er laget slik at de kan brukes til å prognostisere utviklingen til en tjeneste eller et produkt i en tidlig utviklingsfase. Da finnes det lite data-

grunnlag og det er stor usikkerhet ved å bruke andre typer prognosemodeller. Disse modellene har også en form og de er basert på forutsetninger som er svært realistiske for en tjenestes tidlige utvikling.

Et eksempel på bruk er prognostisering av ISDN-eterspørsel. Dette er en tjeneste som er i en startfase og hvor bruk av en metningsmodell er et godt utgangspunkt.

Metningsmodellene kan også brukes til å lage prognoser for tjenester og produkter som nærmer seg en metning. Vi har tidligere vist at det er aktuelt å beskrive etterspørselen etter hovedabonnement for telefon ved en metningsmodell ganske enkelt fordi etterspørselen etter denne tjenesten nærmer seg metning. Generelt kan det sies at når en tjeneste nærmer seg denne metningsgrensen er det aktuelt å bruke denne type modeller.

Metningsmodeller er særlig benyttet i en tidlig fase og i en sen fase av en tjeneste eller et produkts utvikling. Vanligvis brukes andre prognosemetoder når et produkt er i mellomfasen. Imidlertid, hvis en tjeneste er i mellomfasen og det skal lages en langsiktig prognose, kan det også være fornuftig å benytte en metningsmodell.

Det må også skilles på ulike prognosesituasjoner. Det er eksempelvis stor forskjell på å prognostisere abonnementssetterspørselen etter en tjeneste og trafikketterspørselen. For trafikkvolumet vil det for de fleste tjenestene ikke finnes noen naturlig metningsgrense. Dette vil imidlertid være tilfelle for abonnementsvolumet.

Metningsgrensen eller markedspotensialet defineres som den markedsdekning en tjeneste eller et produkt på lang sikt

forventes å få. Spørsmålet er imidlertid hvor konstant dette potensialet er over tid. Dersom modellene skal brukes til langsiktige prognoser, må markeds- og potensialpotensialet på mange års sikt estimeres. Et markeds- og potensialpotensial vil blant annet være avhengig av befolkningsutviklingen, som på lang sikt vil endre seg. Likeledes kan en gradvis forbedring av tjenesten ha betydning for markeds- og potensialpotensialet. Markante prisendringer som følge av ny teknologi kan gjøre det samme. I tillegg kan nye anvendelsesområder gjøre

tjenesten mer attraktiv og generere nye markedssegmenter.

9.2 Ulike metningsmodeller

For supplerende informasjon henvises det til [12], [18], [19], [20] og [21]. Et eksempel på bruk av denne modelltypen til prognostisering av ISDN-eterspørsel er gitt i [22].

Figur 9.1 viser en eksponentiell vekst slik vi har beskrevet den. I figur 9.2 har vi det motsatte – avtagende eksponentiell vekst. Metningsmodellene er sammensatt av disse to elementene: først en eksponentiell vekst og deretter en avtagende eksponentiell vekst.

I figur 9.3 er forløpet til den Logistiske modellen vist. Vi ser hvorledes den øker eksponentielt til å begynne med for senere å avta eksponentielt mot et metningsnivå.

En annen metningsmodell er vist i figur 9.4. Den er kalt Gompertz modell og har en liknende form som den Logistiske modellen, men med en annen matematisk beskrivelse.

Når en tjeneste kommer på markedet, vil den bli tatt i bruk av stadig flere inntil en får en form for metning eller full markedsdekning. Dette er tjenestens markeds- og potensialpotensial. Metningsmodellenes form er basert på å benytte informasjon om

- andelen av potensialet som har tatt tjenesten i bruk
- andelen av potensialet som ikke har tatt tjenesten i bruk.

Det er naturlig at etterspørselen er avhengig av hvor mange som har tjenesten. Jo flere som har tjenesten, desto flere er

det kommunisere med. Dermed blir tjenesten også mer attraktiv.

Etterspørselen vil også være avhengig av de som ikke har tjenesten. Det er disse som er det gjenstående markedspotensialet.

Vekstraten er den årlige tilveksten eller etterspørselen. Selve kurven er den akkumulerte etterspørselen fram til det gitte tidspunktet. Eksempelvis angir kurven ved tidspunkt t totalt antall solgte abonnement fra tidspunkt 0 til og med tidspunkt t for tjenesten. Mange empiriske undersøkelser har vist at den akkumulerte etterspørselen har et forløp som en S, og den kalles da også for S-kurve.

Selve utledningen (den matematiske beskrivelse) av den Logistiske modell kan finnes ved å sette opp at endringen i etterspørselen er proporsjonal med de som har tjenesten og de som ikke har tjenesten, og så løse den differensiallikningen som da framkommer.

9.3 Eksponentiell modell

I en tidlig fase vil etterspørselen etter en tjeneste kunne være eksponentiell. Det betyr at veksten kun er proporsjonal med de som har fått tjenesten.

Vi har da følgende modell:

$$y_t = a e^{-bt} \quad (9.1)$$

Formen på denne modellen er vist i figur 9.1. Kurven har intet vendepunkt.

9.4 Modell med avtagende eksponentiell vekst

Denne modellen har en analog form med den eksponentielle – bare at den avtar istedenfor å vokse. En slik beskrivelse vil være riktig når etterspørselen etter et produkt eller en tjeneste nærmer seg metning. La M være metningsnivået eller markedspotensialet. Vi antar så at etterspørselen ved det tidspunktet da vi starter å bruke modellen ($t = 0$) er m .

Modellen er da gitt ved:

$$y_t = M - (M - m) e^{-at} \quad (9.2)$$

9.5 Logistisk modell

Den Logistiske modellen er beskrevet ved:

$$y_t = \frac{M}{(1 + e^{(a-bt)})^g} \quad (9.3)$$

Her er M markedspotensialet og a , b og g parametere i modellen.

Det som er et problem ved den Logistiske modell, er å estimere parametrene. Det finnes ulike teknikker som kan brukes. En måte er å anta faste verdier på to av parametrene og så estimere de to øvrige. Deretter kan det settes nye verdier på de to første parametrene og prosessen kan gjentas. Det er også mulig å benytte ikke-lineære estimeringsprosedyrer. Vi går imidlertid ikke her nærmere inn på hva dette er for noe.

9.6 Andre metningsmodeller

I tillegg til den Logistiske modellen, er Gompertz modell og Weibull-modellen kjente metningsmodeller.

Gompertz modellen er gitt ved:

$$y_t = a(e^b)^{e^{gt}} \quad (9.4)$$

Her er a større enn 0, b mindre enn 0 og g mindre enn 0. Når g er mindre enn 0, vil kurven flate ut mot et metningsnivå når tiden t blir stor.

Weibull-modellen er gitt ved:

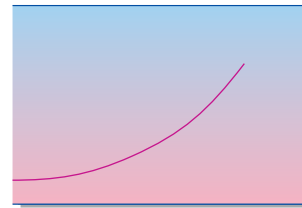
$$y_t = M - (M - m)e^{at^b} \quad (9.5)$$

Her er a mindre enn 0 og b større enn 0. Siden a er mindre enn 0, flater også denne kurven ut når tiden t blir stor.

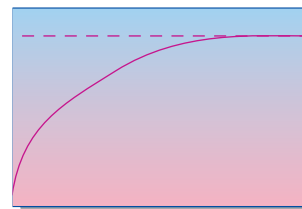
10 Glattingsmodeller

Det karakteristiske ved regresjonsmodeller er at det er mulig å bygge modeller med flere forklaringsvariable – eksempelvis økonomiske og demografiske forklaringsvariable. Metningsmodeller har den egenskapen at de gir god tilpasning til etterspørsel i en tidlig fase og når etterspørselen nærmer seg metning.

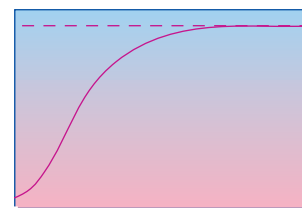
De modellene som presenteres videre – Glattingsmodeller, Tidsrekkemodeller og Kalmanfiltermodeller – er modeller som er basert på tidsrekkeanalyse. Her forklares utviklingen av etterspørsel primært ut fra tidligere verdier på tidsrekken. (Det er imidlertid mulig å utvide Tidsrekkemodellene og Kalmanfiltermodellene med ekstra forklaringsvariable.) Men som sagt, det vesentlige med disse modellene er at de direkte bygger på de tidligere observasjoner i tidsrekken. Modellene forutsetter også at avstanden mellom hver observasjon er like lang (ekvidistante observasjoner). Det kan



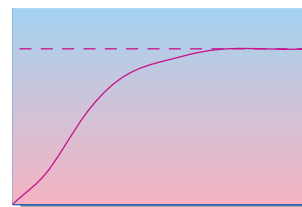
Figur 9.1 Eksponentiell vekst



Figur 9.2 Avtagende eksponentiell vekst



Figur 9.3 Logistisk modell



Figur 9.4 Gompertz modell

eksempelvis være ukentlige, månedlige eller årlige observasjoner.

Denne type modeller egner seg spesielt til å lage prognoser på kort og middels sikt. Metodene er svært gode til å beskrive sesongbetonte variasjoner og utnytter derfor eksisterende observasjoner på en effektiv måte.

I dette avsnittet vil vi se på følgende glattingsmodeller:

- Eksponentiell glatting
- Holts metode
- Holt-Winters metode.

For mer detaljerte studier henvises det til [7], [10], [23], [24] og [42].

10.1 Eksponentiell glatting

Denne metoden skal bare brukes når tidsrekken er stasjonær – det vil si at det ikke eksisterer noen trend i tidsrekken. Det kan være hopp i tidsrekken, men det antas at tidsrekken etter hvert hopp ligger i et stabilt leie. Metoden kan ikke brukes til å modellere sesongsvingninger. Metoden er også kalt Browns metode [25].

I eksponentiell glatting veies de tidligere observasjonene eksponentielt med en glattingsparameter a som har verdi mellom 0 og 1. De relative vektene blir da, sett fra siste observasjon og bakover i tid:

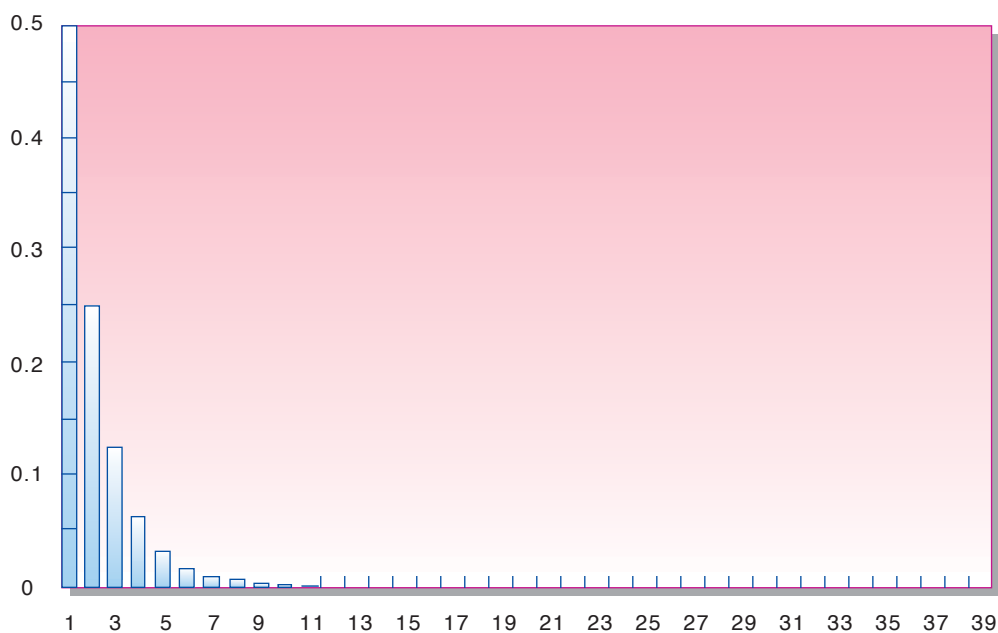
$$1, a, a^2, a^3, \dots \quad (10.1)$$

Det stilles imidlertid et krav om at summen av alle vektene skal være lik 1. Summen av en eksponentiell rekke med a som kvotient er lik $1/(1-a)$. Følgelig må hver av vektene deles med denne summen. Vi får da følgende vektorer som allokeres til observasjonene:

$$(1-a), (1-a)a, (1-a)a^2, (1-a)a^3, \dots \quad (10.2)$$

Figur 10.1 viser vektene når a er lik 0,5.

Det ses at når $a = 0,5$ blir vektene raskt små. Det betyr at det spesielt er de siste observasjonene som har betydning for den videre utviklingen av tidsrekken. Hadde a fått verdien 0,9, er det et tegn på at tidsrekken er mer stabil og atskillig flere observasjoner i større grad påvirker



Figur 10.1 Eksempel på vektorer i en eksponentiell glattingsmodell $a = 0,5$

videre utvikling. Dette ses også på vektningen som er gitt i figur 10.2.

Den eksponentielle glattingsmodellen er uttrykt ved:

$$\hat{y}_{t+1} = (1-a)y_t + (1-a)ay_{t-1} + (1-a)a^2y_{t-2} + \dots \quad (10.3)$$

Her er y den glattede verdien ved tidspunkt $t+1$.

En alternativ modellform som det er lettere å foreta beregningen med, er:

$$\hat{y}_{t+1} = (1-a)y + a\hat{y}_t \quad (10.4)$$

Dette kan uttrykkes verbalt på følgende måte: Glattet verdi ved tidspunkt $t+1$ er lik $(1-a)$ ganger siste observasjon + a ganger glattet verdi ved tidspunkt t .

Ved rekursiv utvikling av (10.4) kan det vises at (10.3) og (10.4) er tilnærmet like (og identiske hvis vi har med uendelig antall ledd).

Den glattede verdien fra likning (10.4) blir da også prognosen som vil være en konstant størrelse uavhengig av tiden. Dette fordi denne modellen skal brukes på stasjonære tidsrekker.

Glattingsparameteren a bestemmes enten ved å forsøke seg fram med ulike verdier eller ved bruk av ikke-lineær estimeringsmetode.

10.2 Holts metode

Holts metode er en utvidelse av eksponentiell glatting der det også inkluderes en trend. Det betyr at denne metoden kan brukes til å lage modeller av tidsrekker som er ikke-stasjonære. Metoden kan imidlertid ikke brukes på tidsrekker med sesongsvingninger.

Modellene er bygget på et nivå som endrer seg med tiden og på en trend eller stigningskoeffisient som også endrer seg med tiden. Disse størrelsene er betegnet med:

- nivået, μ
- stigningen, β .

Vi forutsetter her at det er samme tidslengde mellom hver observasjon. Stigningskoeffisienten i denne sammenheng er definert som hvor mye tidsrekken øker (eller avtar) i løpet av en tidslengde. Der- som vi eksempelvis har årlige observasjoner, vil stigningskoeffisienten reflektere årlig endring. Når nivået og stigningen ved tidspunkt t er kjent, vil et anslag for nivået ved tidspunkt $t+1$ være gitt som summen av nivået ved tidspunkt t og stigningen ved tidspunkt t . Rent matematisk kan dette skrives som:

$$\mu_{t+1} = \mu_t + \beta_t \quad (10.5)$$

På samme måte som for eksponentiell glatting oppdateres både nivået og stigningen rekursivt. I metoden dobbel eksponentiell glatting, som imidlertid ikke omtales her, brukes en glattingsparameter. I Holts metode, som er en mer fleksibel metode, brukes to glattingsparametere. Parameteren a benyttes til glatting av nivået, og parameteren b brukes til glatting av stigningen.

Glatting av nivået gjøres etter følgende resonnering: Nivået ved tidspunkt $t+1$ kan uttrykkes ved likning (10.5), som da er basert på tidligere observasjoner. Imidlertid har vi også fått inn en observasjon av tidsrekken ved tidspunkt $t+1$. Begge disse størrelser sier noe om nivået ved tidspunkt $t+1$. Analogt med hva som ble gjort i eksponentiell glatting multipliseres siste observasjon med $(1-a)$ og det tidligere glattede nivå med a . Vi får da følgende modell for oppdatering av nivået på tidsrekken:

$$\hat{\mu}_{t+1} = (1-a)y_{t+1} + a(\hat{\mu}_t + \hat{\beta}_t) \quad (10.6)$$

På samme måte oppdateres stigningen. Vi multipliserer anslaget for stigningen som er differansen mellom nivået ved $t+$

I og ved t med $(1 - b)$ og summerer dette med den tidligere glattede verdi som multipliseres med b . Dette gir da følgende uttrykk:

$$\hat{\beta}_{t+1} = (1 - b)(\hat{\mu}_{t+1} - \hat{\mu}_t) + b\hat{\beta}_t \quad (10.7)$$

Figur 10.3 illustrerer prinsippene i Holts modeller. Figuren viser hva som skjer med oppdateringen av parametrene når vi er i tidspunkt t i figuren og mottar en ny observasjon y ved tidspunkt $t + 1$. Vi ser at den observasjonen ligger høyere enn prediksjonen av nivået og stigningen. Det nye estimerte nivået ved tidspunkt $t + 1$ blir da liggende mellom den nye observasjonen og det predikerte nivået. I eksemplet i figuren har vi valgt $a = 0,5$, slik at det estimerte nye nivået blir liggende midt mellom de to nevnte punkter.

Likningene (10.6) og (10.7) er basert på rekursive beregninger. Problemet er at for å starte opp prosessen og samtidig utnytte alle observasjonene, er det nødvendig med noen startverdier. Måter å gjøre dette på er enten å benytte noen godt valgte subjektive verdier, eller å bruke regresjonsanalyse for å finne verdiene eller å foreta backcasting som betyr at det lages prognoser omvendt vei på tidsrekken. Ved å benytte en av disse metodene finnes så startverdiene for de rekursive beregningene.

Anta at vi kjenner observasjoner fram til tidspunkt n . En prognose k tidsenheter framover vil da være gitt som summen av det nivået som vi har funnet ved tidspunkt n og k multiplisert med den siste stigningskoeffisienten (ved tidspunkt n) som vi har beregnet. Det betyr at prognosen blir gitt i form av en lineær framskrivning.

10.3 Holt-Winters metode

Holt-Winters metode er som navnet indikerer en utvidelse av Holts metode basert på de samme prinsippene [26]. Forskjellen er at denne metoden også modellerer sesongmessige svingninger. Det er her to hovedmodelltyper.

Additiv modell

Den additive modellen forutsetter at sesongutslagene ikke påvirkes av nivået på tidsrekken. Med andre ord: hvis etterspørselen som reflekteres av tidsrekken over en periode har øket eksempelvis til det dobbelte, forutsettes det at de sesongmessige utslagene ikke påvirkes av denne økningen, men at virkningen av sesongutslagene kan adderes inn i modellen. Den additive modellen uttrykkes ved:

$$\hat{y}_{t+1} = \hat{\mu}_t + \hat{\beta}_t + \hat{S}_t \quad (10.8)$$

Her har vi følgende størrelser:

- Nivået, μ
- Stigningen, β
- Sesongutslaget, S .

Multiplikativ modell

Dersom det er grunn til å tro at sesongutslagene vil øke etter hvert som observasjonene øker i verdi, skal den multiplikative modellen brukes. I svært mange tilfeller er det grunn til å tro at dette er riktig. Den multiplikative modellen er gitt ved:

$$y_{t+1} = (\hat{\mu}_t + \hat{\beta}_t)\hat{S}_t \quad (10.9)$$

Vi ser av likningen at vi får en multiplikativ effekt av sesongutslagene.

Med hensyn til valg mellom den additive og den multiplikative modell, vil det også være mulig å benytte den additive modell dersom observasjonene transformeres, for eksempel ved bruk av en logaritmisk transformasjon, slik som vist i kapittel 7.

I modellene vil det være et gitt sesongutslag for hvert tidspunkt i sesongperioden. Dersom vi har månedlige observasjoner, vil vi ha 12 distinkte sesongkomponenter som rekursivt oppdateres etter hvert som vi får inn månedlige observasjoner. Det introduseres imidlertid kun en ny glattingsparameter c som vil gjelde for samt-

lige sesongutslag. Når vi får inn eksempelvis en ny mars-observasjon, vil vi veie dette utslaget og samtidig det tidligere glattede mars-utslaget. Dermed får vi et nytt oppdatert mars-utslag. Det samme gjøres for alle måneder etter hvert som nye observasjoner kommer inn.

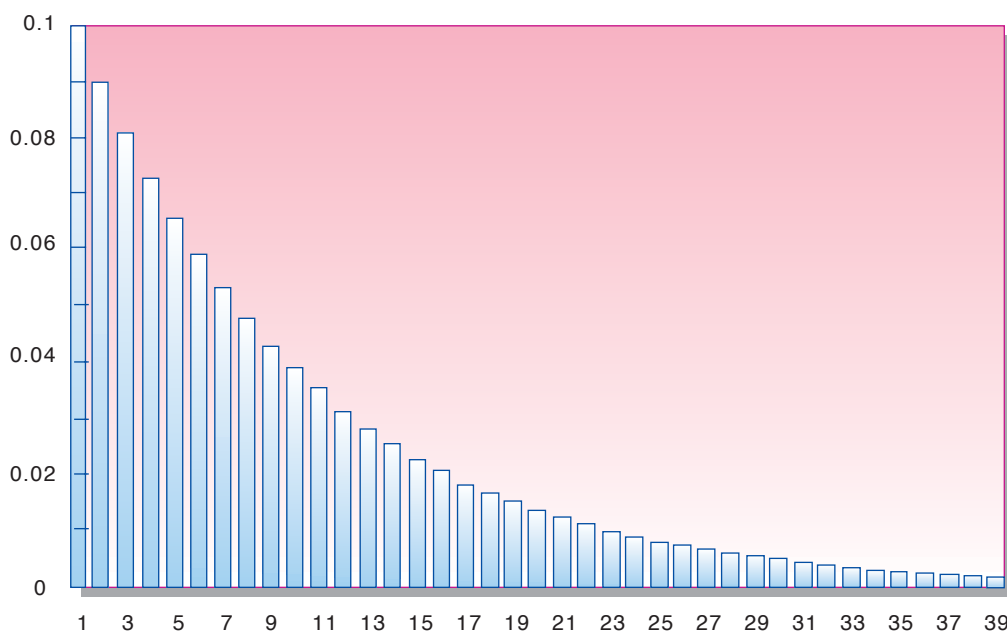
Analogt oppdateres nivået og stigningen til tidsrekken. Formlene blir noe utvidet i forhold likningene (10.6) og (10.7). I tillegg er det en randbetingelse om at summen av sesongutslagene i den additive modellen tilnærmet skal være lik 0, og at multiplikasjonen av sesongutslagene i den multiplikative modellen tilnærmet skal være lik 1.

Beregning av optimale verdier på glattingsparametrene a , b og c gjøres ved bruk av en ikke-lineær optimeringsmetode. Ikke alle programpakker har denne muligheten. Televerket har en egenutviklet programpakke med prognosemetoder som har denne muligheten.

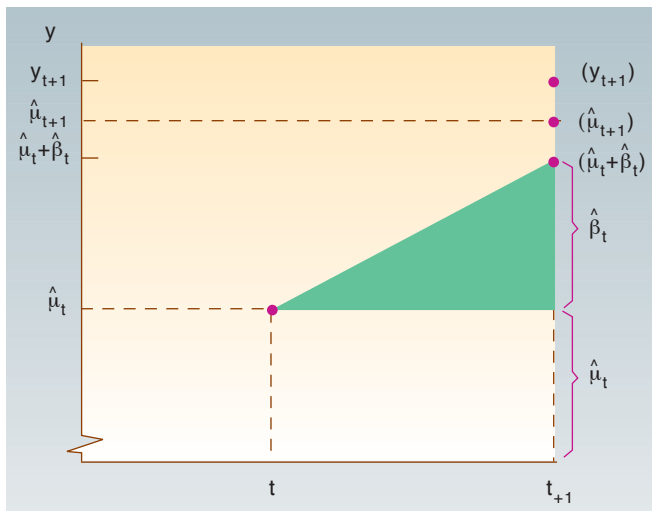
På samme måte som i Holts modeller kan nå nivået, stigningen og tilhørende sesongkomponenter ved siste observerte tidspunkt n beregnes. Deretter lages prognoser for framtidige tidspunkter basert på de beregnede størrelser.

11 Tidsrekkemodeller

De tidsrekkemodeller som omtales her kalles ARIMA-modeller, som står for AutoRegressive Integrated Moving Average modeller (på norsk: autoregress-



Figur 10.2 Eksempel på vektorer i en eksponentiell glattingsmodell $a = 0,9$



Figur 10.3 Relasjon mellom størrelser i Holts modeller

ive integrerte glidende gjennomsnittsmo-
deller).

I 1920-årene utviklet G U Yule de auto-
regressive modeller. Senere kom glidende
gjennomsnittsmo-
deller. Det å sette
disse modellene sammen til en mer fleksibel
modelltype ble så gjort, men det
viste seg vanskelig å anvende modell-
typen. Gjennombruddet kom på slutten
av 1960-tallet med arbeidene til Box og
Jenkins som i boken *Time Series Ana-
lysis, Forecasting and Control* [6] gav en
enhetlig matematisk/ statistisk ramme for
modellene. Metoden de utviklet er kalt
Box-Jenkins metode, og den er en effektiv
og praktisk prosedyre for å anvende
ARIMA-modellene i prognosesammenheng.
Et av de viktigste elementene i
arbeidet til Box og Jenkins er utviklingen
av metoder til å identifisere den ARIMA-
modell som gir best tilpasning til de gitte
observasjoner, samt tester for å evaluere
hvorvidt tilpasningen er tilfredsstillende.

Ved bruk av ARIMA-modeller kan for-
holdvis kompliserte tidsrekker modell-
eres ved bruk av få parametere. Dette er
også et viktig prinsipp i statistisk teori
(The principle of parcimony): å benytte
så få parametere som mulig, men ha en
akseptabel tilpasning. Med andre ord,
ingen overparametrisering i modellene.

ARIMA-modellene, som kun er basert på
tidsrekken tidligere observasjoner, er
spesielt velegnet til å modellere sesong-
messige variasjoner. Modellene brukes til
å lage prognoser på kort og middels sikt.

Dersom det skal lages langsiktige prog-
noser, kan det brukes andre modelltyper;
eventuelt kan det brukes Transfermod-

eller. Disse modellene er en
utvidelse av ARIMA-modellene,
hvor det inkluderes forklarings-
variable i tillegg til selve tids-
rekken.

Av tidligere arbeid på dette feltet
i Televerket henvises det til:
[27], [28], [29] og [30].

Av annen litteratur på området
henvises det til: [6], [10], [31],
[32].

I det følgende ses det på
hovedelementene i en ARIMA-
modell.

11.1 Autoregressiv mod- ell

Autoregressiv er sammensatt av
auto, som betyr selv, og
regressiv, som betyr tilbakegående. En
autoregressiv modell er en beskrivelse av
en tidsrekke som uttrykker den siste
observasjonen som en veiet sum av de
forutgående observasjonene. Vi antar her
at antall forutgående observasjoner som
det er nødvendig å ta med er p .

Den autoregressive modellen er gitt ved:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + a_t \quad (11.1)$$

Her er:

- $\{\varphi\}$ parametere i modellen
- $\{a\}$ støyledd i modellen.

Her er det parametrene i modellen som
angir hvor stor vekt som skal legges på
de enkelte observasjoner. Denne mod-
ellen betegnes også som $AR(p)$ modell
fordi det er en autoregressiv modell med
 p parametere.

Støyleddet betegnes her med a og skiller
seg noe fra støyleddbetegnelse i andre
modeller. Det forutsettes at støyleddet har
forventning lik 0, at støyleddene er uav-
hengige og har konstant varians og at de
gjerner er normalfordelt. Dette er de
samme forutsetninger som er angitt i
kapittel 7. Støyleddene kan da betegnes
som hvit støy.

11.2 Glidende gjennomsnitts- modell

Dette er en modell som er dual til den
autoregressive modell. I en glidende
gjennomsnittsmo-
dell uttrykkes siste
observasjon som en veiet sum av tidlige-
re støyledd. Det forutsettes at antall
støyledd som det er nødvendig å ta med,

er q . Modellen betegnes da som $MA(q)$
modell og er gitt ved:

$$y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (11.2)$$

Her er:

- $\{\theta\}$ parametere i modellen.

Etter at antall parametere i modellen er
bestemt og når verdien på parametrene er
estimert, er modellen bestemt.

11.3 Den sammensatte modell

Den sammensatte modell kombinerer den
autoregressive modell og glidende gjenn-
omsnittsmo-
dell. Fordelen ved denne
modellen er at en med få parametere kan
beskrive situasjonen der modell med
autoregressive eller med glidende gjenn-
omsnittsmo-
deller krever et stort antall
parametere. Dermed begrenses antall
parametere i modellen. Den sammensatte
med p autoregressive parametere og q
glidende gjennomsnittspare-
metere betegnes som $ARMA(p,q)$ og er gitt ved:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (11.3)$$

Vanligvis vil størrelsen på p og q ikke
overstige 2.

11.4 Sesongmodeller

Mange tidsrekker har sesongvariasjoner.
En tidsrekke som eksempelvis angir
månedlig etterspørsel, vil naturlig ha
sesongvariasjoner, og da med en periode
på 12. Ved å studere etterspørselen ved
tidspunkt t , $t-12$, $t-24$, ..., etc. Dette betyr
eksempelvis at sesongutslaget i februar
ett år vil være korrelert med sesongut-
slaget februar året før etc.

En autoregressiv modell som reflekterer
noe av denne avhengighetsstrukturen, er:

$$y_t = \varphi_{12} y_{t-12} + \varphi_{24} y_{t-24} + \dots + a_t \quad (11.4)$$

Hvis det er nødvendig å ta med mange
parametere i den autoregressive mod-
ellen, kan det være aktuelt å ta med noen
glidende gjennomsnittsledd for å redusere
antall parametere. Vi vil da ha en
sammensatt modell på sesongbasis.
Denne modellen vil da kunne beskrives
på analog måte med modellen i likning
(11.3), men med tidsavstand 12.

I de fleste tilfelle vil det ikke bare være
avhengighet mellom etterspørselen i
samme måned i påfølgende år, men også

i etterspørselen i påfølgende måneder samme år. Ved å kombinere disse avhengighetsstrukturene får vi sesongmodellene. Likningen for modellene settes ikke opp her.

11.5 Ikke-stasjonære modeller

De modeller som hittil er satt opp, kan benyttes på tidsrekker som er stasjonære – det vil si tidsrekker som fluktuierer rundt en fast gjennomsnittsverdi og som for øvrig har egenskaper som holder den konstant over tid. De fleste tidsrekker er ikke-stasjonære, og det er fortrinnsvis de som er interessante i en prognosesituasjon.

For å kunne modellere ikke-stasjonære tidsrekker ved bruk av Box-Jenkins metode må tidsrekkene transformeres. Dette gjøres ved å differensiere tidsrekken. Her differensieres tidsrekken $\{y_t\}$ og det lages en ny tidsrekke som er gitt ved:

$$z_t = y_t - y_{t-1} \quad (11.5)$$

Anta at y er totalt antall hovedabonnement ved tidspunkt t . Denne tidsrekken vil selvsagt alltid stige og vil ikke være stasjonær. Den tidsrekken som er definert ved likning (11.5) angir differansen mellom akkumulert antall abonnement i tidspunkt t og tidspunkt $t - 1$ – det vil si tilveksten i abonnement i løpet av denne perioden. Dersom den tidsrekken som angir tilveksten i abonnement pr tidsenhet hele tiden ligger på samme nivå, har vi en stasjonær tidsrekke. Erfaring viser imidlertid at det vanligvis vil være nødvendig å differensiere mer for å få stasjonærhet. Dersom det er tidsrekker med sesongvariasjoner, kan en kombinasjon med differensiering på sesongbasis og på observasjonsbasis være det rette.

Når tidsrekken er differensiert slik at den er stasjonær, starter modellering av tidsrekken. Når modellering er foretatt av den differensierte tidsrekke, integreres den slik at vi kommer tilbake til den opprinnelige tidsrekken, og vi har da en prognosemodell for den opprinnelige tidsrekken. Det er denne prosessen som står for "Integrated" i ARIMA modellene. Hele navnet er AutoRegressive Integrated Moving Average modeller.

11.6 Modellidentifisering

Ved å studere avhengighetsstrukturen i tidsrekken kan en finne fram til den modellen som passer best for tidsrekken. Opplegget som brukes her er noe av det

genuine i Box-Jenkins metode. Dette kalles modellidentifikasjon.

Avhengighetsstrukturen i tidsrekken beskrives ved autokorrelasjoner og partielle autokorrelasjoner. Autokorrelasjoner ble omtalt i kapittel 7. De angir avhengigheter mellom observasjoner som ligger i en gitt lengde fra hverandre. De partielle autokorrelasjoner angir avhengigheter mellom observasjoner som ligger i en gitt lengde fra hverandre, gitt at de mellomliggende observasjoner er kjente.

Enhver ARIMA har et entydig distinkt mønster i autokorrelasjoner og partielle autokorrelasjoner. Det er disse mønstrene som hjelper oss til å bestemme hva slags modell vi har og hvor mange autoregressive og glidende gjennomsnittsparemetere vi skal ha i modellen.

11.7 Estimering av parametrene

Når modelltype og antall parametre i modellen er bestemt, må verdien på parametrene bestemmes. Dette gjøres ved bruk av en ikke-lineær estimeringsmetode som minimerer variansen til støyleddene. Dette er ekvivalent med at det finnes verdier på de ukjente parametrene som gir best mulig tilpasning mellom modellen og de faktiske observasjonene.

Etter at modellen er estimert, undersøkes det om den framkomne modellen beskriver tidsrekken tilfredsstillende. Her finnes det en rekke ulike tester som benyttes. Dersom testene ikke gir tilfredsstillende resultat, gjentas prosessen med modellidentifikasjon og estimering av parametrene eventuelt flere ganger.

11.8 Prognoser

Etter at modelleringen er slutført, lages det prognoser. Prognosene er basert på den modellen som er utviklet. Framskrivningen skjer ved estimering av støyledd og tilpasningen til de aktuelle observasjoner.

Som nevnt kan ARIMA-modellen bygges ut ved inklusjon av ekstra forklaringsvariable. Disse modellene kalles transfermodeller. Det henvises til faglitteratur på området for en mer detaljert innføring.

12 Kalmanfiltermodeller

12.1 Generell beskrivelse

I 1960 publiserte Kalman dynamiske lineære modeller og Kalmanfilteret som viste seg å være grunnleggende for hele regulerings- og kontrollteorien [33]. Det førte nærmest til en revolusjon innen området.

Det har også vist seg at de samme type modeller kan brukes innen prognostisering. Et grunnleggende arbeid her er artikkelen til Harrison og Stevens, *Bayesian forecasting*, i 1976 [34].

Det er vanskelig å kunne beskrive denne teknikken godt uten å gi et sett med formler. Generelt kan det sies at de dynamisk lineære modeller er en meget vid klasse av modeller. Modellklassen inneholder eksempelvis både regresjonsmodeller, glattingsmodeller og tidsrekke-modeller.

I modellklassen forutsettes det at verdien på de parametre som inngår vil kunne endre seg med tiden (derav ordet dynamisk). I glattingsmodellene som er presentert tidligere, ser vi at parametrene endrer seg med tiden – etter hvert som vi foretar oppdateringen. Dermed vil en glattingsmodell kunne beskrive en utvikling der veksten etter hvert endrer seg. I regresjonsmodeller kan vi ikke det. Hvis vi har en modell $y = a + bt$, vil verdien på a og b være de samme uavhengig av tiden. Dette er da en lineær modell som ikke er dynamisk.

Populært kan Kalmanfilteret forklares på følgende måte: Vi har et sett med parametre som angir tilstanden til tidsrekken (eller til prosessen). Ved tidspunkt t har vi et sett med verdier på parametrene. Når vi så mottar ny informasjon om tidsrekken ved tidspunkt $t + 1$, skal verdien på alle parametrene oppdateres. Legg merke til at det samme gjøres også i andre prognosemodeller når vi får inn ny informasjon. Her brukes da Kalmanfilteret til oppdateringen.

Vi kaller innovasjonen for differansen mellom de observerte verdier og de predikerte verdier. Disse kan også kalles prognosefeil, og de vil være et korrektiv til prognosen (predikeringen). Er differansen positiv, har prognosen vært for lav og vice versa. Kalmanfilteret vil være et tall eller et sett med tall (matrise) som brukes til å gi innovasjonene vekt. Den fundamentale formel for oppdatering av parameterverdiene er gitt ved:

Oppdaterte parameterverdier
 = Gamle parameterverdier
 + Kalmanfilteret * Innovasjonen

Prinsippet her likner på det som ble beskrevet for glattingsmodellene. Der ble det imidlertid brukt glattingsparametere istedenfor Kalmanfilter.

Kalmanfilteret filtrerer ut støyen og sørger samtidig for å minimalisere kvadratteilen til de anslåtte parametrene i modellen.

Oppdateringen eller estimeringen av parameterverdiene foretas rekursivt ved oppdatering av likninger både for parameterverdier og for varianser og kovarianser. Som for glattingsmodellene er det også nødvendig å ha startverdier for å starte opp de rekursive beregningene.

12.2 Detaljert beskrivelse

De dynamisk lineære modeller består av to sett med likninger. Det er systemlikning(er) og målelikning(er). Systemlikningen er gitt ved:

$$b_t = Ab_{t-1} + a_t \quad (12.1)$$

b angir parameterverdiene. A er en matrise som angir relasjonen mellom de oppdaterte parameterverdier. a er støyleddene som har kovariansmatrise Q .

Målelikningen er gitt ved:

$$y_t = Hb_t + e_t \quad (12.2)$$

y angir observasjonene – det vil si tidsrekken. H angir relasjonen mellom observasjonen(e) og parametrene. e er støyleddene som har kovariansmatrise R .

For å konkretisere dette, tar vi utgangspunkt i Holts modell. Den vil kunne representeres ved følgende tre likninger hvorav den siste er målelikningen:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + a_{1t} \quad (12.3)$$

$$\beta_t = \beta_{t-1} + a_{2t} \quad (12.4)$$

$$y_t = \mu_t + e_t \quad (12.5)$$

Likning (12.3) angir at nivået ved tidspunkt t er lik nivået ved tidspunkt $t - 1$ pluss veksten ved tidspunkt $t - 1$ pluss tilfeldig støy. Likning (12.4) viser at veksten ved tidspunkt t er lik veksten ved tidspunkt $t - 1$ pluss tilfeldig støy. Likning (12.5) er i dette tilfellet svært enkel. Den forteller at den observerte verdi på tidsrekken ved tidspunkt t er gitt ved nivået på tidsrekken ved tidspunkt t pluss en tilfeldig støykomponent. Legg merke til at y er den eneste størrelse som

kan observeres. Det kan ikke verdien på parametrene eller tilstandene til prosessen, som de også kalles.

De dynamisk lineære modellene oppdateres ved bruk av Kalmanfilter på følgende måte:

La $b_{t/s}$ være forventet verdi på parametrene ved tidspunkt t gitt at vi har målinger fram til tidspunkt s . Ett skritts prediksjon er da:

$$b_{t/t-1} = A_t b_{t-1/t-1} \quad (12.6)$$

$$P_{t/t-1} = A_t P_{t-1/t-1} A_t' + Q_t \quad (12.7)$$

Her er $P_{t/s}$ kovariansmatriser til $b_{t/s}$.

Oppdateringslikningene er:

$$b_{t/t} = b_{t/t-1} + K_t (y_t - H_t b_{t/t-1}) \quad (12.8)$$

$$P_{t/t} = (1 - K_t H_t) P_{t/t-1} \quad (12.9)$$

der Kalman-gevinsten, K_t , finnes ved å minimere $E(b_{t/t} - b_t)(b_{t/t} - b_t)$, hvilket gir:

$$K_t = P_{t/t-1} H_t' (H_t P_{t/t-1} H_t' + R_t)^{-1} \quad (12.10)$$

De oppdaterte parameterverdier finnes av likning (12.9). Dermed kan det lages prognoser ut fra beregnede verdier på parametrene i modellen.

Televerket har i samarbeid med Norsk Regnesentral utarbeidet en meget god programpakke basert på dynamisk lineære modeller og Kalmanfilter. Programpakken inneholder ikke bare generell håndtering av de gitte modellene, men det ligger også inne prosedyrer som kan brukes for å identifisere plutselige trendskift og plutselige nivåhopp i tidsrekken som da vil føre til raffinering av de gitte estimeringsprosedyrer. I tillegg ligger det inne automatisk outlierdeteksjon – hvilket vil si at observasjoner som ligger unormalt langt bort fra forventede verdier vil bli justert slik at de ikke ødelegger estimeringen slik som nevnt i kapittel 7 [35], [36].

Endelig har programpakken prosedyrer for håndtering av manglende observasjoner. Som tidligere nevnt forutsetter en rekke av prognosemodellene at det er like lang tidsavstand mellom observasjonene – det vil si at det foretas observasjoner eksempelvis på uke-, måneds- eller årsbasis. Selv om dette gjøres, vil det kunne være situasjoner som gjør at enkelte eller et sett av påfølgende observasjoner mangler. I programpakken med Kalmanfilter er det utviklet optimale prosedyrer for Kalmanfilteroppdateringen når observasjoner mangler [37], [38].

Annen dokumentasjon om bruk av disse modellene på etterspørselsdata i Televerket er [5], [39], [40], [41] og [42].

For øvrig henvises det til [13], [34] og [43].

13 Lokale og aggregerte prognoser

13.1 Problemstilling

I mange sammenhenger lages det lokale og aggregerte prognoser. Det kan eksempelvis lages prognoser for trafikken i hvert sentralområde og i hver region og prognoser for landet som helhet. Sentralprognosene og regionsprognosene vil da være lokale, mens landsprognosen vil være den aggregerte. For så vidt vil også regionsprognosen kunne betraktes som en aggregert prognose for sentralområdene.

Spørsmålet er så hvorledes vi best kan nyttiggjøre oss den informasjon som ligger i de lokale prognosene og den informasjon som ligger i den aggregerte prognosen.

Det er klart at summen av de lokale prognosene sjelden blir lik den aggregerte prognosen. Hvorledes skal vi da justere prognosene slik at de blir konsistente? Den aggregerte prognosen er basert på større datamengder enn de lokale prognosene. I mange tilfeller er det her også mulig å benytte mer avanserte prognosemodeller. Positive og negative usikkerheter kan på et aggregert nivå til en hvis grad nøytralisere hverandre ("Store tall lov"). Den relative usikkerheten i den aggregerte prognosen vil derfor være mindre enn for de lokale prognosene. Samtidig er det viktig å være oppmerksom på at mange lokale forhold kan være modellert inn i de lokale prognosene.

13.2 Bottom-up-metoden

Denne metoden neglisjerer mulighet til å benytte en aggregert prognosemodell som korrektiv til de lokale prognosene. Derfor anbefales ikke denne metoden. Bottom-up-metoden er basert på de lokale prognosene. Den aggregerte prognosen er definert som summen av de lokale prognosene.

13.3 Top-down-metoden

Denne metoden er basert på å bryte ned den aggregerte prognosen i lokale prognoser. Det forutsettes her at den

aggregerte prognosen er riktig. Nedbrytningen gjøres da ved å justere de opprinnelige lokale prognoser slik at summen av dem blir lik den aggregerte prognosen.

Anta eksempelvis at summen av de lokale prognoser er 7 % større enn den aggregerte prognosen. Ved å dividere hver av de lokale prognosene med 1,07, vil summen av de lokale prognosene bli lik den aggregerte prognosen. Dermed er alle prognosene konsistente.

Svakheten ved denne metoden er for det første at den aggregerte prognosen antas som sikker. Samtidig forutsettes det at alle de lokale prognosene er relativt like usikre.

13.4 Veiet minste kvadraters metode

Denne metoden er basert på at det tas hensyn til prognoseusikkerheten både i den aggregerte prognosen og i de lokale prognosene.

Et naturlig mål vil være å minimalisere kvadratavviket mellom de opprinnelige prognosene og de justerte prognosene. I tillegg bør det tas hensyn til usikkerheten i de ulike prognosene som er laget. Hver kvadratsum vil ut fra dette tillegges en vekt som representerer det inverse av prognoseusikkerheten. Med andre ord: dersom det er en stor usikkerhet i en prognose, så vil det under justeringen ikke bli lagt så stor vekt på denne prognosen som på en prognose der det er liten usikkerhet.

Prognoseusikkerheten er her definert som variansen til prognosen. Denne størrelsen kan beregnes direkte ved bruk av enkelte prognosemodeller. Ved bruk av andre prognosemodeller, må et anslag for prognoseusikkerheten beregnes på annen måte. Det kan eksempelvis gjøres ved å holde tilbake noen observasjoner og så beregne ett-skritts prognosefeilen og så kvadrere denne størrelse.

Vi lar a være de opprinnelige prognosene og b være de justerte prognosene. Det er laget prognoser for s lokale områder nummerert fra 1 til s . Den aggregerte prognosen har prefiks T . Variansen i de respektive prognosene er angitt ved σ^2 opphøyd i andre.

Likning (13.1) angir da det uttrykket som skal minimeres. Og det skal gjøres under betingelse av at den justerte aggregerte prognosen er lik summen av de justerte lokale prognosene. Dette er uttrykt i likning (13.2).

$$Q = \frac{1}{\sigma_T^2} (a_T - b_T)^2 + \sum_{i=1}^s \frac{1}{\sigma_i^2} (a_i - b_i)^2 \quad (13.1)$$

gitt at:

$$b_T = \sum_{i=1}^s b_i \quad (13.2)$$

Her blir altså både totalprognosen og de lokale prognosene justert. Dette gjøres ved bruk av Lagranges multiplikatorregel. Resultatet er vist i likning (13.3) som angir de justerte prognosene.

$$b_i = a_i - \sigma_i^2 \frac{\sum_{i=1}^s a_i - a_T}{\sum_{i=1}^s \sigma_i^2 - \sigma_T^2} \quad (13.3)$$

Likning (13.3) forteller at den justerte prognosen er lik den opprinnelige prognosen minus et lite tillegg som er positivt eller negativt avhengig av om differansen mellom den opprinnelige aggregerte prognosen og summen av de opprinnelige lokale prognosene er positiv eller negativ. Dette tillegget justeres ut ifra størrelsesforskjell på variansene til de respektive prognoser.

Denne metoden er etter bidrag fra Norge tatt inn i CCITTs rekommendasjon E 506 for prognostisering av internasjonal trafikk [44].

14 Prognostisering av trafikkmatriser

14.1 Trafikkmatriser

Så langt har vi sett på tilfeller der vi har en aggregert prognose og et sett med lokale prognoser. Vi vil imidlertid ha en mer komplisert situasjon dersom vi skal prognostisere en trafikkmatrise.

En trafikkmatrise kjennetegnes ved at den i hvert element angir trafikken mellom to punkter eller områder. Eksempel på trafikkmatrise er trafikken mellom alle fjernsentralene i Norge – altså i landsnettet. For nettkonfigurasjon se figur 2.5. I rad nr 1 i denne trafikkmatrisen vil vi ha utgående trafikk fra sentral nr 1 til de øvrige sentraler. I kolonne nr 1 vil vi ha innkommende trafikk fra alle øvrige sentraler til sentral nr 1, etc. Det element i trafikkmatrisen som er på rad nr i og i kolonne nr j angir den

trafikken som går fra sentral nr i til sentral nr j .

I tillegg kan radene og kolonnene i trafikkmatrisen summeres. Radsummene angir totaltrafikken ut fra de respektive sentraler, mens kolonnesummene angir totaltrafikken inn til de respektive sentralene. Summen av radsummene eventuelt kolonnesummene angir totaltrafikken i trafikkmatrisen. Se figur 14.1.

Etter å ha observert trafikken i trafikkmatrisen over en periode, er det mulig å lage prognoser for

- enkeltelementene i trafikkmatrisen
- radsummene – den utgående trafikken fra sentralene
- kolonnesummene – den innkommende trafikken til sentralene
- totalsummen – totaltrafikken i trafikkmatrisen.

Naturlig nok vil de ulike prognosene her ikke være konsistente. Spørsmålet er så hvorledes skal prognosene justeres slik at de blir konsistente.

14.2 Kruthofs metode

Kruthofs metode ble publisert i 1937 [45]. Den forutsetter at det ikke lages prognoser for enkeltelementene i trafikkmatrisen, men kun for radsummer og kolonnesummer. Ut fra den tids forhold er ikke det så urimelig, da det på den tiden ikke var så mange trafikkmålinger.

Metoden baserte seg på siste observerte trafikkmatrise og på prognoser for rad- og kolonnesummer.

Ved å summere rad nr 1 i trafikkmatrisen ses det at denne summen som vi kaller S , er forskjellig fra prognosen som vi benevner P . For å få alle elementene på rad nr 1 konsistente med prognosen multipliseres de med forholdet P/S . Det samme gjøres analogt for alle de øvrige radene. Da vil alle elementene i trafikkmatrisen være konsistente med radsumprognosene.

Summen av trafikkelementene i de respektive kolonnene i trafikkmatrisen vil imidlertid ikke stemme med prognosene for kolonnesummene. Dermed gjentas den samme justeringsprosessen her. Alle elementene i trafikkmatrisen vil da være konsistente med prognosene for kolonnesummene.

Denne iterasjonsprosessen gjentas på henholdsvis rader og kolonner inntil det skjer minimale endringer. Iterasjonspro-

sessen har da konverget, og vi har fått et sett med elementer i trafikkmatrisen som er konsistente med prognosene for rad- og kolonnesummer.

14.3 Kruthofs utvidede metode

En utvidelse av Kruthofs metode vil være å også lage prognoser for enkeltelementene i trafikkmatrisen og bruke disse som input istedenfor sist observerte trafikkmatrise. Dette vil gi bedre resultater. En svakhet med metoden er dog at all vekt tillegges rad- og kolonnesumprognosene. Disse endres ikke under iterasjonen – kun de enkelte elementer i trafikkmatrisen.

14.4 Veiet minste kvadraters metode

Disse modellene er utviklet av Stordahl [41]. Som i kapittel 13.4 settes det opp et uttrykk der kvadratsummer mellom opprinnelige prognoser og de justerte prognoser skal minimaliseres. Dette skal gjøres under randbetingelse om at prognosene for alle trafikkelementer er konsistente med prognosene for alle radsummer og med prognosen for alle kolonnesummer og med totalsumprognosen.

Argumentasjonen er at det er naturlig at alle prognosene bør justeres noe – ikke bare prognosene for de enkelte trafikkelementer. Det er rimelig at de ulike typer prognoser bør justeres relativt avhengig av hvor stor prognoseusikkerheten er. Det vil ikke her bli presentert formler som rent matematisk beskriver optimeringssituasjonen. Den vil imidlertid være forholdsvis lik den som er beskrevet ved likning (13.1) og (13.2) i kapittel 13.4, men med flere likninger som randbetingelser. Til hver kvadratsum allokeres det en vekt. Et aktuelt valg av vekter kan være å velge dem lik den inverse av variansen til prognosen.

Løsningen på optimeringen finnes ved å benytte Lagranges multiplikator metode. For en trafikkmatrise med dimensjon n fører dette til at det må løses et likningssystem med $n(n+3)+2$ ukjente. Dette kan ikke gjøres for hånd, men må gjøres maskinelt.

I samarbeid med Norsk Regnesentral er det utviklet en programpakke som inneholder de nevnte metoder for prognostisering av trafikkmatriser. Her løses det nevnte likningssystemet numerisk i selve programpakken.

I enkelte situasjoner kan det være slik at vi har spesiell kunnskap om enkelte av elementene i trafikkmatrisen. Det kan være elementer der vi vet at trafikken er 0 og det kan være elementer der vi med stor sikkerhet kjenner den framtidige utviklingen av trafikken. Disse elementene bør da utelates fra selve justeringsprosessen. Det kan vises matematisk at den optimeringen som er beskrevet også kan gjøres på trafikkmatriser der enkeltelementer ikke skal tas med eller på trafikkmatriser der deler av matrisen er borte.

Med andre ord: vi kan kjøre den samme justeringen på en redusert matrise. Det er derfor laget en programpakke kalt REDUSER som angir hvilke data som skal brukes i en redusert trafikkmatrise. Programpakken summerer også rad- og kolonnesommene.

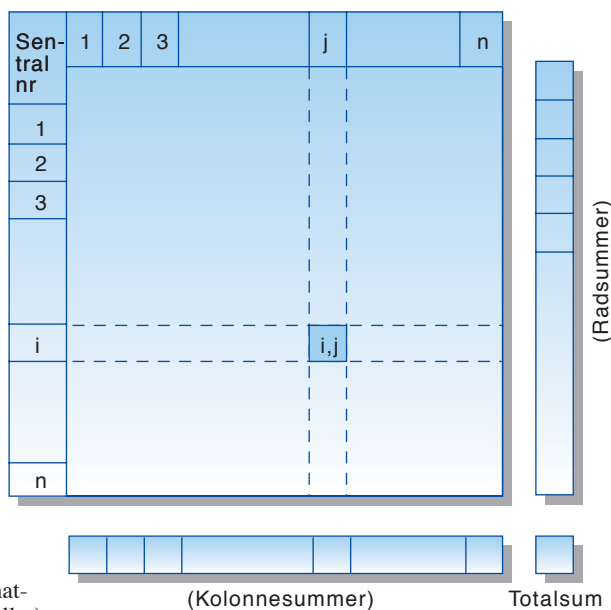
14.5 Prognosesystem for prognostisering av trafikkmatriser

For prognostisering av trafikkmatriser er det utviklet et sett med programpakker som kan brukes. Figur 14.2 gir en oversikt over prosessen og hvorledes de ulike programpakker brukes til prognostiseringen. Dette gjøres ved følgende trinn:

- 1) Først samles trafikkmålingene inn og lagres i en database (Trafikkmålinger).
- 2) Deretter behandles trafikkmålingene, om nødvendig. Dette gjøres hvis trafikkmatrisen skal reduseres til en mindre matrise eller hvis radsummer og kolonnesummer ikke tidligere er beregnet. Dette gjøres ved bruk av programpakken Reduser.
- 3) Etter denne operasjonen framkommer Reduserte trafikkmatriser.
- 4) Ut fra et sett med trafikkmålinger skal det nå lages prognoser for de enkelte elementer i trafikkmatrisen. Dette kan gjøres automatisk (Automatiske prognosemodeller) ved bruk av dynamisk

lineære modeller og Kalmanfilter. En må være klar over at det her er tale om et stort antall trafikkelementer. I fjernettet er det nå omlag 15 fjernsentraler, slik at det her må lages rundt 200 enkeltprognoser. Programpakken med Kalmanfilter er utviklet slik at prognosemodellene er robuste overfor endringer i tidsrekken. Som nevnt i kapittel 12 er det lagt inn automatisk outlier-deteksjon, hvilket justerer unormale observasjoner. I tillegg er det lagt inn prosedyrer for deteksjon av trendsift og nivåhopp i tidsrekken og endelig prosedyre for håndtering av manglende observasjoner. Dermed vil disse prognosemodellene kunne brukes automatisk.

- 5) I spesielle tilfeller kan det være aktuelt å benytte andre prognosemodeller på elementer i trafikkmatrisen. Dette kan eventuelt gjøres også ved bruk av andre tilgjengelige programpakker. For prognoser av rad- og kolonnesummer og for totalsummen vil dette være relevant. På aggregert nivå vil det være aktuelt å bruke mer kompliserte prognosemodeller (Manuelle prognosemodeller).
- 6) Etter at prognostiseringen er slutført, legges prognosene for de enkelte elementer, for rad- og kolonnesummer og for totalsummen ut på separate filer (Prognoser for matriser; rad/kol:sum og totalsum). Disse



Figur 14.1 Trafikkmatrise

prognosene vil med sikkerhet ikke være konsistente. Det må derfor foretas justeringer av prognosene ut fra de prinsipper som er presentert.

- 7) Dersom det er ønskelig å foreta endringer ut fra subjektive vurderinger, markedsmessige vurderinger, etc, er det mulig å gå direkte inn i filene og endre prognosene.
- 8) Mange prognosemodeller angir prognoseusikkerheten automatisk i sin output. Dermed vil standardavvik eller varians til prognosen være angitt. De beregnede verdier legges på en egen fil (Varians i prognoser for matrise; rad/kol: sum og totalsum).
- 9) Det er også mulig å justere de angitte verdier for varians ved bruk av subjektivt skjønn (Eventuell subjektiv vektning).
- 10) De endelige vekter for de enkelte prognoser legges så ut på fil (Vekter for prognostisering).
- 11) Filer med trafikkmatriser og filer med vekter danner nå inngangsdata for trafikkmatriseprognostisering.

Programpakken med trafikkmatriseprognostisering benyttes så til beregning av konsistente trafikkmatriseprognoser (Trafikkmatrise-modeller).

- 12) Etter at prognostiseringen av trafikkmatrisen er gjennomført, framkommer så endelig den prognostiserte trafikkmatrise. Vanligvis lages det ikke bare prognoser for et tidspunkt men for en tidsperiode. Vi får da et sett med prognostiserte trafikkmatriser (Prognostiserte trafikkmatriser).

Ved å bruke dette systemet for prognostisering av trafikkmatriser kan det spares mye tid, fordi det er lagt mye automatikk i selve prognostiseringen. Det er også lagt opp effektive rutiner for behandling av dataene. Systemet er fleksibelt utformet slik at det er mulig å gå inn og korrigere mellomliggende resultater ved bruk av supplerende metoder og eventuelt subjektive vurderinger. Det er imidlertid viktig at det har vært en viss stabilitet i utviklingen av det nettområdet som ses på. Store omrutinger og endringer i nettområdet må det justeres for.

For nærmere detaljer henvises til [36], [40], [41], [46], [47], [48], [49], [50] og [51].

å undersøke i hvilken grad de ulike prognosemodellene tilfredsstill forutsetningene. En svært viktig forutsetning er at residualene – det vil si de estimerte støyleddene – tilfredsstill de fire forutsetninger som er gitt i kapittel 7. Samtidig bør standardavviket (og variansen) være liten og naturlig nok det beregnede konfidensintervall.

Det brukes også spesielle tester/observatorer for å vurdere tilpasningen og prognosene for ulike typer modeller. Eksempel på slike tester/observatorer er multipl korrelasjonskoeffisient i regresjonsmodeller, Portmanteau observator i tidsrekke modeller og Durbin-Watson observator og autokorrelasjonsfunksjon for ulike modeller.

På bakgrunn av disse størrelsene samt vurdering av egenskaper og størrelse på residualene vil det ofte være mulig å eliminere flere av modellene som er foreslått fordi de ikke oppfører seg tilfredsstillende.

De modellene som vi til slutt står tilbake med, bør så evalueres ut fra selve prognosene som lages.

15.2 Metode for sammenlikning av prognosemodeller

For ikke å vente på nye observasjoner for å kunne sammenlikne de ulike prognosemodellene, er det vanlig å holde tilbake noen observasjoner. Anta at vi totalt har $m + n$ observasjoner. Vi bruker da n observasjoner på selve modellbyggingen. Deretter lages det prognoser m tidsenheter framover. Vi får da m prognoserverdier som direkte kan sammenliknes med de m observasjoner som vi har holdt tilbake.

Residualet e er som illustrert i figur 7.2 lik differansen mellom observasjonen og tilpasningen. Vi får her en helt analog definisjon av prognosefeilen som er lik differansen mellom den observasjon som er holdt tilbake og selve prognosen. Vi er da i den situasjon at vi for hver prognosemodell som vi skal sammenlikne, har et sett med m prognosefeil. Det vil da være størrelsen på prognosefeilene som er avgjørende for hvilken prognosemodell som kommer best ut.

I det følgende er det satt opp ulike mål eller kriterier for residualene som kan brukes til å sammenlikne ulike prognosemodeller.

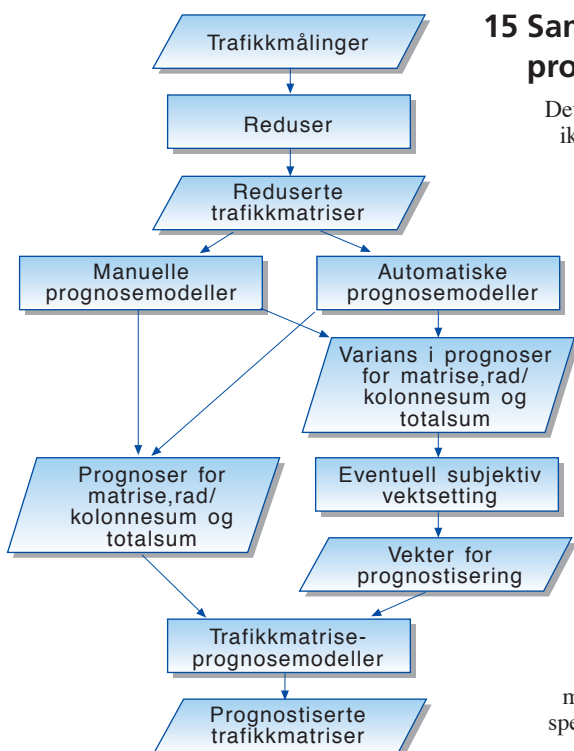
15 Sammenlikning av prognosemetoder

Det vil i en rekke tilfeller ikke være selvsagt hva slags prognosemetode som skal velges. Det kan da være riktig å prøve flere av metodene for å finne fram til hvilke eller hvilken som fungerer best.

Det vil her bli sett på en del kriterier som kan benyttes ved valg av prognosemetode.

15.1 Evaluering av prognosemetoder

Til de ulike prognosemetoder knytter det seg spesielle forutsetninger. Første del av evalueringen er



Figur 14.2 Prognosesystem for prognostisering av trafikkmatriser

15.3 Gjennomsnittlig feil

Denne observatoren kalles på engelsk "Mean Error". I programutskriftene forkortes den til *ME*.

Gjennomsnittlig prognosefeil er definert som summen av alle prognosefeilene dividert med antall observasjoner som er holdt tilbake (*m*). Navnet er svært dekkende.

Denne observatoren viser om det er skjevhet i prognosene. Det kan være grunn til å stille spørsmålstegn ved prognosemodellen dersom gjennomsnittlig prognosefeil er stor. Det er viktig å være oppmerksom på at selv om gjennomsnittlig prognosefeil for en modell er liten, trenger dette ikke å bety at prognosemodellen er god. Det kan tenkes at det er mange store positive og mange store negative prognosefeil, men at de gjensidig delvis opphever hverandre. Av den grunn kan ikke dette kriteriet brukes alene, men det må brukes sammen med andre kriterier.

15.4 Gjennomsnittlig prosentfeil

Denne observatoren kalles på engelsk "Mean Percent Error" som forkortes *MPE*.

Gjennomsnittlig prosentfeil knytter prognosefeilen relativt til størrelsen på prognosen. Dette gjøres ved at hver enkelt prognosefeil divideres med observasjonen på samme tidspunkt.

Anta eksempelvis at observasjonen ved et tidspunkt er 500 og at prognosen er lik 550. Prognosefeilen blir da $500 - 550 = -50$. Den relative prognosefeilen er $-50/500 = -0,10$. Gjennomsnittlig prosentfeil er definert ved å summere de relative prognosefeilene for alle *m* tidspunkter og så dividere på *m* og deretter multiplisere med 100 %. I eksempelet over var den relative prognosefeilen -0,10. Når dette multipliseres med 100 %, får vi -10 % som da er prosentfeilen. Deretter beregnes et gjennomsnitt av alle prosentfeil, og vi får gjennomsnittlig prosentfeil.

Også denne observatoren angir eventuell skjevhet i prognosene. Fordelen med gjennomsnittlig prosentfeil i forhold til gjennomsnittlig prognosefeil er at den er noe lettere å forholde seg til. Hvis gjennomsnittlige prosentfeil er 5 %, vet vi med en gang mye om skjevheten. Hvis vi får oppgitt at gjennomsnittlig prognosefeil er 34, er det ikke lett å ta stilling til dette uten å få oppgitt noe om størrelsen på prognosene.

15.5 Rot av gjennomsnittlig kvadratfeil

Denne observatoren er svært hyppig brukt som prognosekriterium. Den er et godt mål på prognoseusikkerheten. På engelsk betegnes den som "Root Mean Square Error" og forkortes *RMSE*.

Rot av gjennomsnittlig kvadratfeil er definert som summen av kvadratet av prognosefeilene dividert med antall observasjoner *m*. Til slutt tas kvadratroten av dette.

Ulempen ved gjennomsnittlig prognosefeil, *ME*, var at positive og negative feil kunne oppheve hverandre. Ved å ta kvadratet til prognosefeil – det vil si opphøye det i andre – vil vi unngå dette problemet. Da vil alle både store positive og negative prognosefeil gi et stort positivt bidrag. Det beregnes så et gjennomsnitt av alle kvadratfeilene, og det tas kvadratrot av dette uttrykket for at det skal være av samme størrelsesorden som prognosefeilene.

15.6 Gjennomsnittlig absolutt feil

Denne observatoren betegnes på engelsk "Mean Absolute Error" som forkortes *MAE*.

Gjennomsnittlig absolutt feil er definert identisk med gjennomsnittlig feil, *ME*, kun med den forskjell at istedenfor prognosefeil settes det inn absoluttverdien av prognosefeilen. Med andre ord er *MAE* definert som summen av absoluttverdien til prognosene dividert med antall observasjoner *m*.

15.7 Gjennomsnittlig absolutt prosentfeil

Denne observatoren kalles på engelsk "Mean Absolute Percent Error" som forkortes *MAPE*.

Gjennomsnittlig absolutt prosentfeil er definert identisk med gjennomsnittlig prosentfeil, *MPE*, kun med den forskjell at istedenfor prognosefeil settes det inn tallverdien av prognosefeil i uttrykket.

I eksempelet fra kapittel 15.4 så vi at den relative prognosefeilen var -10 %. For absolutt prognosefeil får vi tallverdien av -10 % som er 10 %. Den gjennomsnittlige absolutte prognosefeilen finnes ved å ta gjennomsnittet av de *m* absolutte prognosefeilene.

Fordelen med *MAPE* framfor *MPE* er at *MAPE*, fordi den er basert på absolutt-

verdier, kan identifisere dårlige prognoser også der positive og negative prognosefeil kompenserer hverandre. Dette kan som sagt ikke *MPE*. En annen fordel med *MAPE* er at den forteller hvor stor prognoseusikkerheten er i forhold til prognosen. Dette gjør *MAPE* mer informativ enn *RMSE*. Imidlertid brukes *RMSE* mer, men da er det nødvendig å vurdere prognosetallene samtidig med de oppgitte resultater på *RMSE*.

15.8 Valg av kriterium

Det er ikke nødvendig å bruke alle observatorene som er satt opp her. Det kan være aktuelt å bruke *ME* eller *MPE* for å undersøke om det er noen skjevhet i prognosene.

I tillegg kan en av de øvrige, *RMSE*, *MAE* eller *MAPE* brukes for å finne anlag for prognoseavviket.

Mål for sammenlikning av ulike prognosemodeller vil da være at prognose-skjevheten skal være liten og prognoseavviket skal være lite. I tillegg bør de respektive prognosemodellene tilfredsstillende vanlige krav til residualer, etc.

Det må også benyttes en del sunt skjønn under evalueringen. Anta for eksempel at tidsrekken får en helt annen utvikling gjennom de *m* siste observasjonene i forhold til de *n* første observasjonene. Det vil da være vanskelig å benytte *ME*, *MPE*, *RMSE*, *MAE* og *MAPE* til vurdering av de ulike prognosemodellene. Vi må være klar over at en prognosemodell ikke vil kunne lage unormale prognoser. En prognosemodell vil alltid lage prognoser basert på de inngangsdata som modellen har. Dette betyr at prognosemodellen vil lage prognoser som forlenger den utviklingen som vi til nå har sett. Dette er basert på at forutsetningene er de samme for den framtidige utvikling som det har vært på utviklingen inntil nå. Dersom dette ikke er tilfelle og vi vet det, må prognosemodellen endres slik at det inkluderes nye forklaringsvariable som er i stand til å forklare en markant endring i tidsrekkens forløp. Med andre ord: dersom vi får et markant endret forløp på tidsrekken på de *m* siste observasjonene i forhold til de *n* første, er det liten vits i å bruke de nevnte kriterier for å finne den beste blant en rekke dårlige prognosemodeller. Det som da må gjøres, er å utvikle bedre prognosemodeller – og da bruke flere av observasjonene til dette.

Dersom antallet observasjoner *m* som vi holder tilbake for sammenlikningens

skyld, er lite, vil det også være en klar statistisk usikkerhet i en konklusjon med rangering av de beste prognosemodeller ut fra de kriterier vi har angitt her. Dette er jo også noe av sjarmen med å lage prognoser. Det er ingen bestående sannheter. Etter hvert som nye observasjoner kommer inn, skal den som har ansvaret for prognosene evaluere resultatene og være åpen for å skifte prognosemodell dersom dette er riktig. Uansett skal parametrene estimeres på nytt når ny informasjon er tilgjengelig.

16 Programpakker og verktøy i prognosearbeidet

Bruk av stormaskiner eller PC-er er en naturlig del av prognosearbeidet. Det henger sammen med spesielt to faktorer: Datahåndteringen og beregningene. Med de datamengder det her er tale om er det ikke mulig å foreta manuelle sorteringer og justeringer av data. De prognosemodeller som er omtalt i denne artikkelen, er også så beregningsteknisk kompliserte at det ikke er mulig – det vil si umulig – å foreta dette manuelt.

I det følgende vil vi se litt på de programpakker og verktøy som er tilgjengelig på markedet, men legge mest vekt på det som Televerket selv bruker.

16.1 Regneark

Bruk av regneark har øket nærmest eksponentielt parallelt med utbredelsen av PC-ene. Regneark benyttes ofte i tilknytning til databaser og tekstbehandling. I Televerket ser utviklingen av regneark til å gå fra Enable til Excel og til Lotus. Policyen til Televerket er å kjøpe kombinerte pakker som kombinerer tekstbehandling å regneark.

Regnearket har den senere tid tatt inn en del instruksjoner fra høynivåspråk som gjør regnearket mer slagkraftig. Det er nå mulig å sette opp programstrukturer med IF, THEN, ELSE, etc.

Ved bruk av regneark er det nå lettvis å foreta justeringer og håndtering av data. I kapittel 4 ble det gitt mange eksempler på arbeid som det var nødvendig å foreta på dataene for å få dem på en hensiktsmessig form for å kunne lage prognoser. En meget stor andel av prognosearbeidet går faktisk med til denne type håndtering av dataene. Til dette arbeidet er regnearket et meget hensiktsmessig verktøy.

Det er viktig å være svært strukturert når det lages regnearkmodeller. Det bør være gode ledetekster og oversiktlige skjermbilder og en god logisk oppbygging av hele strukturen.

16.2 Tilgjengelige programpakker

Det finnes en rekke aktuelle programpakker – i tillegg til den som Televerket selv har utviklet – som inneholder relevante prognosemetoder. Av de mest kjente statistiske programpakker nevnes Statgraphics, SPSS, BMDP, SAS, MINITAB og SYSTAT.

Programpakkene inneholder stort sett glattingsmodeller, regresjonsmodeller og tidsrekkemodeller. Ikke alle programpakker har Kalmanfilter.

I Televerket har AS vært et brukerorientert databehandlingsverktøy knyttet til Televerkets IBM stormaskin. AS har hatt en egen statistikkdel som har inneholdt noen prognosemodeller. Televerket går imidlertid over til SAS. Dette databehandlingsverktøyet inneholder en meget omfattende statistikkdel hvor også de relevante prognosemetoder finnes.

Televerket kjører også et prognosekurs [52]. Til dette kurset benyttes Statgraphics som også inneholder de fleste relevante prognosemetoder med unntak av Kalmanfilter. Statgraphics brukes til kurset fordi det er menystyrt og svært lett å lære seg og dermed gir pedagogiske fordeler.

Programpakkene Statgraphics, SPSS, BMDP, SAS, MINITAB og SYSTAT kan benyttes både på stormaskiner og på PC-er.

I tillegg til dette finnes det en rekke spesialpakker for både tidsrekkeanalyse og Kalmanfilter.

16.3 Televerkets egenutviklede programpakker for prognostisering

Som nevnt tidligere har Televerket utviklet egne programpakker i samarbeid med Norsk Regnesentral for prognostisering. Arbeidet har bestått i et samarbeid mellom Norsk Regnesentral, TF, det tidligere Trafikkplankontoret (TNP) og Nettplanleggingsgruppen i Øst, tidligere Ø.LNP. Programpakken som er utviklet heter Forecast, og den består av tre mindre programpakker:

- Programpakke for reduksjon og summering av trafikkmatriser

- Programpakke for beregning av trafikkmatriseprognoser
- Programpakke med ulike prognosemodeller.

De to første programpakkene er relativt grundig omtalt i kapittel 14. Den siste programpakken inneholder følgende prognosemetoder:

- Kalmanfilter med lokalt nivå
- Kalmanfilter med lineær vekst
- Kalmanfilter med additive sesongkomponenter
- Holts metode
- Holt-Winters additive metode
- Holt-Winters multiplikative metode.

I tillegg finnes det som nevnt i kapittel 12 om dynamisk lineære modeller og Kalmanfilter, en rekke prosedyrer som kan benyttes i tilknytning til metodene. Det er lagt inn følgende opsjoner:

- Outlierdeteksjon
- Detektere unormale observasjoner og fjerne dem slik at de ikke forårsaker estimeringsfeil og dermed feil prognoser
- Identifikasjon av trendskift
- Identifisere tidspunkt for trendskift og redifinere modellen
- Identifikasjon av nivåhopp
- Identifisere tidspunkt for nivåhopp og redefinere modellen
- Robusthet mot manglende observasjoner
- Håndtering av manglende observasjoner. Foreta identifisering av modellen og estimering av parametere når det mangler observasjoner.

Disse prosedyrene ligger inne i programpakken og kan om ønskelig automatisk brukes som opsjoner, slik at prognosene kjøres automatisk ut når en gitt modell er valgt. Det vil kun være aktuelt å kjøre ut prognoser automatisk dersom det er et spesielt stort antall prognoser som skal lages – eksempelvis i en stor trafikkmatrise.

For øvrig brukes denne programpakken til å lage prognoser for en rekke viktige tidsrekker i Televerket. Den brukes blant annet til å lage prognoser for tellerskritt-

volumet både regionalt og for husstands- og forretningsmarkedet i Televerket.

For mer detaljert informasjon om programpakken henvises det til [36].

Referanser

- 1 Trælnes, T, Stordahl, K. *Behovsstudie for nye teletjenester i hjemmene*. Kjeller, Televerkets Forskningsinstitutt, 1976. (TF-rapport 19/76.)
- 2 Trælnes, T, Stordahl, K. Bruk av markedsundersøkelse ved prognostisering av nye teletjenester. *Bedriftsøkonomen*, 9, 406–409, 1976.
- 3 Stordahl, K. En markedsundersøkelse om knappsatsapparater og nye særtjenester. *Teletronikk*, 73(3), 230–241, 1977.
- 4 Bøe, J, Hjelkrem, C. Utarbeiding av prognoser for etterspørsel etter telefonabonnement i Oslo teledistrikt. *Teletronikk*, 87(1), 93–109, 1991.
- 5 Bøe, J, Stordahl, K. Tellerskrittene – Televerkets viktigste inntektskilde. *Teletronikk*, 87(1), 78–92, 1991.
- 6 Box, G E P, Jenkins, G M. *Time series analysis: forecasting and control*. San Fransisco, Holden Day, 1969.
- 7 Gilchrist, W. *Statistical forecasting*. New York, Wiley, 1976.
- 8 Makridakis, S, Wheelwright, S, McGee. *Forecasting: methods and application*. New York, Wiley, 1978.
- 9 Wonnacott, T H, Wonnacott, R J. *Regression*. New York, Wiley, 1981.
- 10 Abraham, B, Ledolter, J. *Statistical methods for forecasting*. New York, Wiley, 1983.
- 11 Levenbach, H, Cleary, J P. *The modern forecaster*. Belmont, Calif., Lifetime Learning Publications, 1984.
- 12 CCITT. *GAS 10 Handbook, Data acquisition and forecasting methods for telecommunication planning*. Genève, 1987.
- 13 Harvey, A C. *Forecasting, structural time series models and the Kalman filter*. Cambridge, Mass., 1989.
- 14 Theil, H. *Principles of econometrics*. New York, Wiley, 1971.
- 15 Mosteller, F, Tukey. *Data analyses and regression*. Reading, Mass., Addison Wesley, 1977.
- 16 Draper, N R, Smith, H. *Applied regression analysis, 2nd ed.* New York, Wiley, 1981.
- 17 Chow. *Econometrics*. New York, McGraw Hill, 1983.
- 18 Lekval, P, Wahlbin, C. *A study of some assumptions underlying innovation diffusion functions*. Swedish Journal of Economics, 1973.
- 19 Hernes, G. Diffusion and growth – the non-homogeneous case. *Scandinavian Journal of Economics*, 1976.
- 20 Blindheim, E, Rygg, M. *En modell for å utarbeide prognoser for nye teletjenester*. Kjeller, Televerkets Forskningsinstitutt, 1982. (TF-rapport 13/82.)
- 21 Log, S. *Diffusjonsmodeller for nye teletjenester*. Kjeller, Televerkets Forskningsinstitutt, 1982. (TF-rapport 14/82.)
- 22 Bøe, J, Hjelkrem, C. Metodebeskrivelse for utarbeidelse av prognoser for ISDN-aksesser i Oslo teledistrikt. *Teletronikk*, 87(1), 110–123, 1991.
- 23 Montgomery, D C, Johnson, L A. *Forecasting and time series analyses*. New York, McGraw-Hill, 1976.
- 24 Makridakis, S, Wheelwright, S. *Interactive forecasting*. San Fransisco, Calif., Holden Day, 1977.
- 25 Brown, R G. *Smoothing, forecasting and prediction*. Englewood Cliffs, N J, Prentice Hall, 1963.
- 26 Winters, P R. Forecasting sales by exponentially weighted moving averages. *Management service*, 6(3), 324–342, 1960.
- 27 Stordahl, K et al. *Korttidsprognosemodeller basert på tidsrekkeanalyse, Del I og II*. Kjeller, Televerkets Forskningsinstitutt, 1979. (TF-rapport 10/79.)
- 28 Stordahl, K. Analysing telecommunication data by use of ARIMA models. *Forecasting Public Utility*.

- Ed: O D Anderson. Amsterdam, North Holland, 1980.
- 29 Damsleth, E, Stordahl, K. Etterspør- sel etter hovedabonnement for tele- fonen – en prognosemodell. *Bedrifts- økonomen*, 3, 164–170, 1981.
- 30 Hjelkrem, C, Bøe, J. Box-Jenkins metode. *Teletronikk* 90(1), 110–124, 1994 (denne utgave).
- 31 Jenkins, G M, McLeod, G. *Case studies in time series analysis*. Lan- caster, G M Jenkins, 1982.
- 32 McLeod, G. *Box and Jenkins in Practice*. Lancaster, G M Jenkins, 1982.
- 33 Kalman, R E. A new approach to linear filtering and prediction prob- lems. *Journal of Basic Engineering*, 82, 34–45, 1960.
- 34 Harrison, P J, Stevens, C F. Bayesian forecasting. *J Roy. Stat. Soc.*, B 38, 189–203, 1976.
- 35 Aldrin, M, Bølviken, E. *Automatiske prognosemetoder basert på Kalman- filteret: muligheter og problemer*. Oslo, Norsk Regnesentral, 1987. (NR STAT 26/87.)
- 36 Aldrin, M, Volden, R. *Brukermanual for Forecast*. Oslo, Norsk Regne- sentral, 1987. (NR STAT 22–24/87.)
- *Brukermanual I for prognostiser- ing av trafikkmatriser. Program- pakke for reduksjon og summering av trafikkmatriser*.
- *Brukermanual II for prognostiser- ing av trafikk*.
- *Brukermanual III for prognostiser- ing av trafikkmatriser. Program- pakke for konsistente trafikk- matriser*.
- 37 Aldrin, M. *Prognosemetoders robusthet mot manglende data*. Oslo, Norsk Regnesentral, 1986. (NR STAT 6/86.)
- 38 Wright, D H. Forecasting data published at irregular time intervals using an extension of Holt's method. *Management Science*, 32, 1986.
- 39 Bølviken, E. *Forecasting telephone traffic using Kalman filtering; theoretical considerations*. Oslo, Norsk Regnesentral, 1986. (NR STAT 5/86.)
- 40 Stordahl, K, Damsleth, E. Traffic forecasting methods for network planning. *Proceedings 6th interna- tional conference on forecasting and analyses for business in the information age*. 1986.
- 41 Stordahl, K. *Utvidet veiet minste kvadraters metode: en metode for trafikkmatriseprognostisering*. Kjeller, Televerkets Forskningsinsti- tutt, 1987. (TF-rapport 38/87.)
- 42 Bøe, J, Hjelkrem, C. Glattingsmod- eller. *Teletronikk* 90(1), 76–86, 1994 (denne utgave).
- 43 Brown, R G. *Random signal analysis and Kalman filtering*. New York, Wiley, 1983.
- 44 CCITT. SG II Rekommandasjon E 506, *Traffic Forecasting*. Genève, 1988.
- 45 Kruithof, J. Telfoonverkeersrekning. *De Ingenieur*, 52(8), 1937.
- 46 Stordahl, K, Holden. Traffic forecast- ing models based on top down and bottom up procedures. *Proceedings 11th International Teletraffic Con- gress 3.4B*. 1985.
- 47 Bølviken, Damsleth. *Forecasting tra- ffic matrices by Kalman filter*. Oslo, Norsk Regnesentral, 1986. (NR STAT 10/86.)
- 48 Stordahl, K. Methods for traffic matrix forecasting. *Proceedings 12th international teletraffic congress*, 5.3A, 1988.
- 49 Cole, A C. Progress in forecasting and traffic measurement techniques as seen at recent teletraffic con- gresses. *Proceedings 12th international teletraffic congress*, 1988.
- 50 Bakken, A, Sigstad, H. *Resultater for uttesting av EDB baserte prognose- metoder for Televerket*. Oslo, 1990.
- 51 Stordahl, K. *Prognosemetoder for nettplanlegging*. Kjeller, Televerkets Forskningsinstitutt, 1986. (TF-rap- port 85/86.)
- 52 Bøe, J, Hjelkrem, C, Stordahl, K. *Prognosemetoder*. Kurskompendium for Teleskolen. Grimstad, 1991.
- 53 Stordahl, K, Hjelkrem, C, Bøe, J. Regresjonsmodeller. *Teletronikk* 90(1), 87–102, 1994 (denne utgave).
- 54 Stordahl, K. Bruk av ekspertunder- søkkelser til prognoser for nye tele- tjenester. *Teletronikk* 90(1), 53–64, 1994 (denne utgave).

Bruk av ekspertundersøkelser til prognoser for nye teletjenester i privatmarkedet

AV KJELL STORDAHL

1 Bakgrunn

Det å kunne lage prognoser for en ny teletjeneste er i seg selv problematisk. Det vil være lettere å lage prognoser for en teletjeneste som har vært på markedet en periode. Da vil utformingen og funksjonaliteten til tjenesten være kjent. Kostnadene til investering for utbygging av tjenesten og for drift av tjenesten vil være kjente. Samtidig vil også takstene for tjenesten være gitt, og ut fra dette vil kundene over perioden ha generert abonnements- og trafikkundersøkelser.

Når tjenesten er ny og ennå ikke introdusert i markedet, vil ingen av de ovennevnte data være tilgjengelige. Dette gjør at prognostisering for etterspørsel etter tjenesten blir spesielt vanskelig og også usikker.

Tradisjonelle kvantitative prognosemetoder kan derfor ikke benyttes. Isteden må en benytte andre angrepsmåter. Aktuelle kvalitative angrepsmåter/ prognosemetoder er:

- Markedsundersøkelser
- Potensialvurderinger
- Ekspertbedømmelser
- Analogimetoden
- Scenariometoden
- Delphimetoden.

Metodene omtales nærmere i følgende kapittel.

Hvorfor er det viktig å utarbeide prognoser for tjenester som vi ennå ikke har på markedet? I en del tilfeller er det ikke tatt noen beslutning på innføring av tjenesten. Da vil ekspertvurderinger og markedsundersøkelser være et viktig underlag for å kunne anslå det potensielle behovet for tjenesten. Denne type vurderinger er helt nødvendig som en del av et beslutningsgrunnlaget for innføring av en tjeneste.

Ser vi spesielt på privatmarkedet og tjenester som krever stor kapasitet i nettet, vil en prognose for slike tjenester allerede i dag influere på hvorledes abonnentnettet skal bygges. Mange har i lengre tid arbeidet med tekniske løsninger for FTTH (Fiber To The Home), FTTB (Fiber To The Building) og FTTC (Fiber To The Curb). Den fremtidige etterspørsel etter høykapasitets tjenester både til bedriftskunder og privatkunder vil være avgjørende for de nettløsninger som etter hvert skal velges for å tilfredsstille etterspørselen.

Gravekostnadene er en vesentlig del av de totale kostnadene i abonnentnettet. Det er derfor svært viktig så tidlig som mulig å ha kjennskap til en sannsynlig utvikling i etterspørselen etter høykapasitetstjenester. Sammen med utnyttelse av ny teknikk – det vil si bruk av optiske nettkomponenter – vil etterspørsel gi grunnlag for nye strukturplaner som nå lages i hvert sentralområde.

I denne artikkelen ses det på resultatene fra en eldre Delphiundersøkelse som ble gjennomført av Televerket midt i 70-årene. Det foretas en vurdering av hvor gode prognosene var og av hvilke erfaringer som er blitt både fra gjennomføringen og fra resultatene fra undersøkelsen. Deretter omtales en ny Delphiundersøkelse som nå gjennomføres av Televerket på det samme markedet – privatmarkedet. I tillegg gis det en oversikt over ulike kvalitative metoder som kan brukes til å lage prognoser.

2 Ulike kvalitative metoder for å utarbeide prognoser

2.1 Markedsundersøkelser

Markedsundersøkelser kan også brukes som underlag for prognoser. Et eksempel på bruk av markedsundersøkelser for å lage prognoser er å spørre aktuelle abonnenter på den nye tjenesten om hva de er villig til å betale for tjenesten. Det kan også stilles et sett med prisspørsmål slik at en kan få estimert en etterspørselskurve som funksjon av prisen. Ut fra forutsetninger om kostnads- og prisutviklingen kan det så ut fra etterspørselskurven lages en prognose for tjenesten.

Markedsundersøkelser kan også brukes til å kartlegge substitutter for en tjeneste. Eksempelvis vil møtevirksomheten i norske bedrifter være et volum som kan være et substitutt for bildetelefon og konferansejernsyn.

2.2 Potensialvurderinger

Et potensial er det samme som metningsnivået for en tjeneste. Det betyr at det eksempelvis kan være det maksimale antall abonnenter som vil etterspørre en tjeneste.

Anslag for et potensial kan gjøres på ulike måter. Som nevnt i 2.1 kan et potensial finnes ved bruk av markedsundersøkelser. Ved å kartlegge møtevirksomhet i bedrifter finnes et potensial for telemøter fremfor 'face to face' møter.

Anslag for potensialer kan også finnes ved å gå inn i offisiell statistikk. Totalt antall norske husstander vil kunne være et potensial for en gitt teletjeneste – eller det kan være summen av antall leiligheter, bolighus og hytter.

En fremgangsmåte for å lage prognoser ut fra et anslag for potensialet er å benytte metningsmodeller der potensialet eller metningsnivået er en av de fundamentale parametrene i modellen.

2.3 Ekspertbedømmelser

Når det ikke eksisterer data, utviklingstrender eller annen kvantitativ informasjon som underlag for prognosene, vil ekspertbedømmelser være en aktuell metode.

En metode som er velkjent er brainstorming. Det vil si at det innkalles en gruppe eksperter som sitter sammen og inspireres og influeres av hverandre. Resultatene her kommer frem gjennom en forholdsvis ustrukturert prosess.

En annen fremgangsmåte er at en prognoseansvarlig plukker ut et sett med eksperter og at han på en systematisk måte fremlegger problemstillingen gjerne med et sett med gjennomarbeidede spørreskjemaer. Ekspertenes svar vil da være prognosegrunnlaget.

2.4 Analogimetoden

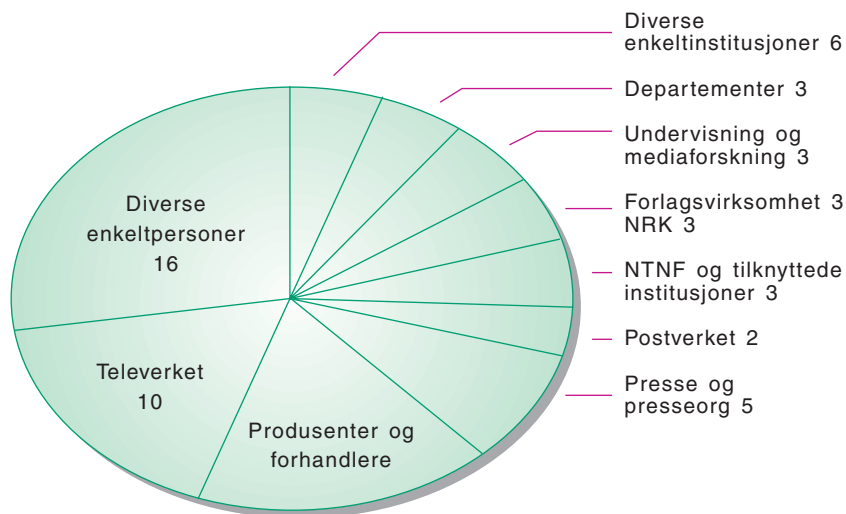
Analogimetoden går ut på å anta at utviklingen for ulike produkter er sammenliknbar eller at utviklingen av like produkter i ulike områder, eksempelvis ulike land, kan være lik, men i forskjellige tidsintervaller.

Dersom det skal lages prognoser for utviklingen av en ny teletjeneste, kan det være naturlig å se på hvorledes utviklingen har vært i et land der denne tjenesten allerede har blitt introdusert. Ved å se på utviklingen av ISDN i Frankrike der tjenesten en periode har vært kommersielt tilgjengelig og anta at utviklingen i Norge vil ha analogt forløp etter kommersiell introduksjon, er det mulig å lage prognoser for ISDN også for Norge. En må imidlertid være varsom med slike avledninger fordi forholdene i de to land som sammenliknes kan være vidt forskjellige – både demografisk, økonomisk og med hensyn til utforming og funksjonalitet av tjenesten.

2.5 Scenariometoden

Metoden går ut på å lage enkeltprognoser, men med en fremstilling av flere variable

001.18:621.39



Figur 3.1 Deltakernes sammensetning i Delphiundersøkelsen

som gjensidig varierer og som kan påvirke hverandre. Det er en konstruktiv metode som kan benyttes til å avdekke og vurdere nærmere årsakssammenhenger og påvirkninger.

Metoden går ut på å ta utgangspunkt i et spesielt år – eksempelvis 10 år frem i tid. Det skal så lages en forventet beskrivelse av situasjonen dette året. Hvis vi tar utgangspunkt i Televerket, vil det være naturlig å se på Televerkets konkurranse-situasjon dette året. Her åpnes det selv-sagt for ulike situasjonsbilder, og dette er en av fordelene med scenariometoden. Det er mulig å foreta ulike beskrivelser avhengig av de forutsetninger som legges til grunn. Deretter går metoden ut på å gi en beskrivelse av utviklingen av de til-hørende variable 10 år fremover i tid. Beskrivelsen av denne utviklingen for å komme frem til det gitte fremtidsbilde blir da scenariometodens prognoser.

2.6 Delphimetoden

Delphimetoden er en raffinering av ekspertundersøkelse. Navnet Delphi kommer fra oldtidens Hellas der oraklet ble oppsøkt og hvor det gav prognoser som var mer eller mindre vel funderte.

I Delphimetoden velges det ut et sett med eksperter. En del av ekspertene skal være spesialister innenfor det aktuelle fagom-rådet, men mange eksperter kan også være spesialister innen andre områder.

Delphimetoden går ut på å gjøre et grundig forarbeid før undersøkelsen foretas. Deretter utformes det et spørre-skjema til bruk i undersøkelsen. Aktuelle spørsmål i en Delphiundersøkelse er forventet introduksjonstidspunkt for en ny tjeneste, forventet utvikling av tjenesten og gjerne etterspørsel etter tjenesten som funksjon av pris.

Det sendes så ut spørreskjema til et sett med eksperter. Svarene returneres, hvoretter de behandles på statistisk form slik at enhver vil se hvorledes han har svart i forhold til de øvrige. Disse resultatene sendes så sammen med spørre-skjemaet tilbake til hver enkelt ekspert som igjen skal svare på de samme spørsmålene. I tillegg skal det også sendes ut spesielle opplysninger som enkelte av deltakerne har angitt som begrunnelse for sine vurderinger. På denne måten skal ekspertene etter hvert nærme seg en felles oppfatning. Vanligvis vil det være tilstrekkelig å bruke to til tre runder i en Delphiundersøkelse.

I det videre vil det bli sett mer detaljert på Delphiundersøkelser som Televerket har gjennomført.

3 Erfaringer fra tidligere Delphiundersøkelse

3.1 Målsetting og bakgrunn for undersøkelsen

Fra august 1975 til september 1976 gjennomførte Televerkets Forskningsinstitutt i samarbeid med Industrikonsulent A/S en større Delphiundersøkelse. Resultatene er dokumentert i TF-rapport 19/76 "Behovsstudie for nye teletjenester i hjemmene" [1].

Målsettingen med undersøkelsen var å finne grove anslag for den langsiktige etterspørsel etter nye teletjenester for privatmarkedet. Resultatene skulle være et underlag for langtidsplanleggingsarbeidet i Televerket. Med den lange levetiden som nettkomponentene har og ut fra både teknologisk utvikling, økonomisk utvikling og samfunnsutviklingen var det viktig for Televerket å kunne kjenne til

markedsutviklingen for nye teletjenester. Langsiktige planer for utbygging av Televerkets infrastruktur var avhengig av denne type prognoser.

I sluttrapporten fra undersøkelsen heter det "Ambisjonsnivået ved de tallmessige og kvalitativt angitte uttrykk for tjeneste-behov, har i lys av usikkerheten i problemstillingen ikke vært særlig høy. En har vært inneforstått med at de resultater som fremkom, neppe ville være direkte egnet for drastiske beslutninger om idrift-setting. Undersøkelsen må heller sees på som et første skritt i markedsanalysene og som en læreprosess om hvilke teleanvendelser som kan muliggjøres i fremtiden".

En må også være klar over at undersøkelsen ble gjennomført på en tid da de fleste tjenestene som ble beskrevet i undersøkelsen ikke var kjent eller definert i Televerket. Mange av tjenestene i undersøkelsen ble beskrevet ut fra litteratursøk og ut fra utenlandske – spesielt amerikanske – fagtidsskrifter. På den tiden var det heller ikke noen kjente markedsundersøkelser tilgjengelig internasjonalt for disse nye tjenestene.

Delphiundersøkelsen ble gjennomført blant et sett eksperter – et sett subjektivt utvalgte personer – for å vurdere en rekke nye teletjenester som først var forventet innført i det norske telenettet frem mot århundreskiftet.

I Norge hadde vi på den tiden praktisk talt ingen som arbeidet med markedsanalyser. Vi hadde bare de klassiske tjenestene telefon, teleks, datel og telegram og vi hadde svært lange ventelister på telefonabonnement. Det var ikke noe behov for verken markedsføring eller markedsanalyser for å få økt salget og dermed økt inntektene av tjenestene.

3.2 Gjennomføring av undersøkelsen

Planleggingen av gjennomføringen av Delphiundersøkelsen tok tre måneder. En svært viktig aktivitet i denne perioden var arbeid med definisjon av de aktuelle tjenester og med utforming av selve spørreskjemaet. Det er også på dette punktet Delphiundersøkelsen skiller seg fra en del andre kvalitative undersøkelser. Selve underlaget skal være godt gjennomarbeidet og det som det spørres om, skal være godt definert og strukturert. Dette gjør det enklere for ekspertene å svare på spørsmålene selv om det er kompliserte problemstillinger.

Spørsmål i Delphiundersøkelsen

Eksempler på anvendelser:

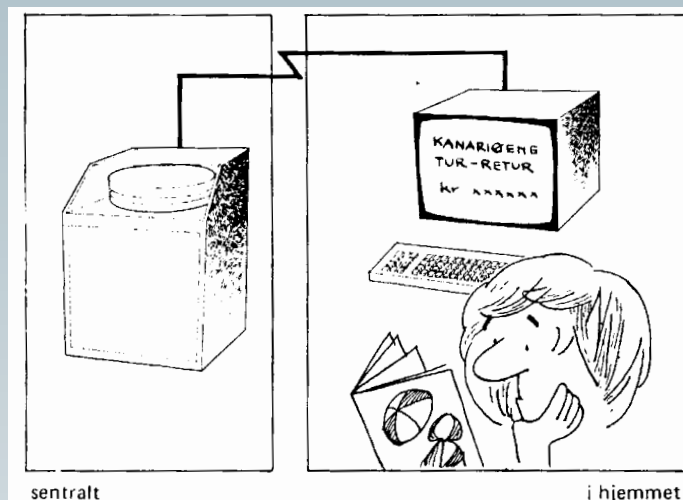
Tekniske forutsetninger:

Spørsmål		Svar							Kommentar	
Ønskelighet	1 Hvordan ser De på ønskeligheten av en slik tjeneste ut fra en samfunnsmessig helhetsvurdering?	Sterkt uønsket			Indifferent (likegyldig)			Sterkt ønsket		
		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
		-3	-2	-1	0	1	2	3		
Ved besvarelse av spørsmål 2, 3, 4 og 5 må man forsøke å tenke seg at disponibel inntekt målt i 1976-kroner er 50 % høyere enn i dag. Videre må spørsmålene ses i sammenheng med hverandre.										
Bruksvolum	2 Hvor mange ganger pr uke tror De en norsk husstand som har anskaffet nødvendig utstyr for tjenesten, vil benytte den?	<input type="text"/>	ganger pr uke							
	3 Hvor mange minutter pr gang tror De tjenesten gjennomsnittlig vil bli anvendt av dem som har anskaffet nødvendig utstyr?	<input type="text"/>	minutter pr gang							
	4 Hvor meget vil de som har anskaffet nødvendig utstyr for tjenesten, akseptere som månedlig utgift for det bruksvolum De har angitt ovenfor (målt i 1976-kroner og anskaffelsesutgiftene holdt utenfor)?	<input type="text"/>	kroner pr måned							
	5 Hvor stor prosentdel av norske husholdninger vil ønske å anskaffe og bruke utstyr for denne tjenesten dersom de må betale det beløp som De har angitt under spørsmål 4?	<input type="text"/>	% av husholdningen							
Ved besvarelse av spørsmål 6, 7 og 8 må man forsøke å trekke både markedssiden og tilbudssiden inn i vurderingen. De må derfor basere svarene på egne forutsetninger om kostnader, teknologi, priser, markedsforhold, utbyggingstakt, politiske beslutninger, etc.										
Introduksj.	6 Når omtrent tror De tjenesten vil bli introdusert for norske husstander?	Årstall:	<input type="text"/>							
Utbygging	7 Hvor stor andel av de norske husstandene tror De vil være tilbudt tjenesten henholdsvis 5 år og 15 år etter introduksjonen?	Prosentvis andel etter 5 år	<input type="text"/>	Prosentvis andel etter 15 år	<input type="text"/>					
			<input type="text"/>		<input type="text"/>					
Markedsdekning	8 Hvor stor andel av de husstandene som er tilbudt tjenesten, vil ha anskaffet den henholdsvis 5 år og 15 år etter introduksjonen?	Prosentvis andel etter 5 år	<input type="text"/>	Prosentvis andel etter 15 år	<input type="text"/>					
			<input type="text"/>		<input type="text"/>					

Figur 3.2 Spørreskjema i Delphiundersøkelsen

Innhenting av informasjon fra databank

Tilgang til tekstinformasjon og stillestående bilde på TV-skjerm. Informasjonen kan være relativt spesialisert, og derfor dekke individuelle og lokale informasjonsbehov. Databanken oppdateres fortløpende, og er tilgjengelig til vilkårlige tidspunkter. "Teleavis på bestilling" er tatt med som egen tjeneste, nr 11.



Eksempler på anvendelser:

- Som tjeneste 5 "Teletekstoverføring på TV-skjerm", men med større datamengder og større individuelle valgmuligheter
- Informasjonsinnhenting, eksempelvis forbrukeropplysninger, juridisk informasjon, etc
- Oversikt over alternative reiseruter
- Varekataloger/vareoversikt
- Oversikt over bibliotektilbud
- Kontoforespørsel

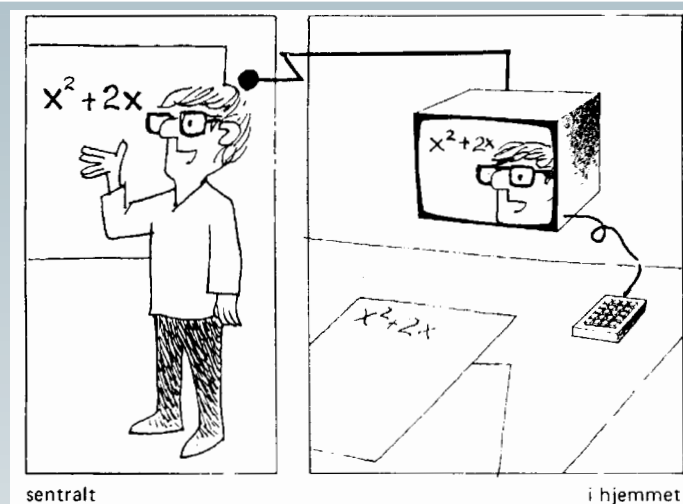
Tekniske forutsetninger:

Bestilling av informasjon skjer ved hjelp av et tastatur og en returkanal (eksempelvis telefonlinje) til databank. Informasjonen kan mottas på TV-skjerm.

Figur 3.3 Innhenting av informasjon fra databank

Teleundervisning med lærer

Tilgang til undervisningsprogram med levende bilde. Undervisningen foregår lokalt til fastsatte tidspunkter, og er basert på forelesninger hvor de hjemmевærende elevene kan gi enkle svar på spørsmål, etc. Elevene kan velge fritt mellom et begrenset antall programmer.



Eksempler på anvendelser:

Kurstilbud innen ulike emner som ikke krever fysisk oppmøte på et undervisningssted.

Tekniske forutsetninger:

TV-apparat og egen undervisningskanal. Samspill mellom lærer og elev ved hjelp av et tastatur og returkanal (eksempelvis telefonlinje).

Figur 3.4 Teleundervisning mot lærer

Deretter ble det arbeidet med å velge ut personer som skulle delta i Delphiundersøkelsen. Totalt ble det sendt ut forespørsel til 123 personer. Det ble lagt vekt på at deltakerne skulle være bredt representert for å kunne dekke flest mulig berøringspunkter mellom samfunn og nye tele-tjenester i hjemmene. Det skulle altså ikke være noen representativ markedsundersøkelse, men en undersøkelse med deltakelse av en rekke spesialister. Det deltok 54 personer i begge rundene, mens 64 personer deltok enten i første eller andre runde. Sammensetning av deltakerne er vist i figur 3.1

En Delphiundersøkelse gjennomføres vanligvis over to eller tre runder. I denne undersøkelsen ble det vurdert at to runder var tilstrekkelig. Til gjengjeld ble deltakerne invitert til et avsluttende møte hvor sluttresultatene fra runde to ble presentert og hvor det var mulig å komme med utfyllende kommentarer.

3.3 Definerte tjenester

De tjenestene som ble definert i Delphiundersøkelsen, var:

- Flere TV-kanaler – større programtilbud
- TV-program mot særskilt betaling
- TV-program fra programbibliotek
- Videospiller og -opptaker
- Teletekstoverføring på TV-skjerm
- Innhenting av informasjon fra databank
- Selvbetjening over telenettet
- Teleundervisning mot lærer
- Teleundervisning mot datamaskin
- Teleavis på abonnement
- Teleavis på bestilling
- Telepost
- Bildetelefon
- Tele-hjemmearbeid
- Datarespons og telekontroll
- Nye telefontjenester
- Telealarm
- Fjernkontroll
- Radioprogram på bestilling
- Stillbilde telefon.

Det registreres at det her er tjenester som krever både smal- og bredbåndskapasitet.

I Delphiundersøkelsen ble det for hver tjeneste spurt om:

- Ønskelighet for tjenesten
- Hvor ofte og lenge tjenesten ville bli brukt (bruksfrekvens)
- Pris for bruk av tjenesten
- Etterspørsel gitt pris
- Introduksjonstidspunkt for tjenesten
- Etterspørsel hhv 5 og 10 år etter introduksjon
- Utbyggingsgrad hhv 5 og 10 år etter introduksjon.

Figur 3.2 viser hvorledes spørreskjemaet var utformet.

Det var et spørreskjema for hver tjeneste. På spørreskjemaet ble det også gitt en illustrasjon av tjenesten. I tillegg ble det gitt eksempler på hvorledes tjenesten kunne anvendes, og endelig ble det gitt en teknisk beskrivelse av tjenesten.

Figur 3.3 og 3.4 viser hvorledes tjenestene "Innhenting av informasjon fra databank" og "Teleundervisning mot lærer" er illustrert.

3.4 Ønskeligheten av tjenestene

Figur 3.5 angir hvorledes en i 1976 på en skala fra -3 til +3 vurderte samfunnsmessig ønskelighet av de ulike tjenester. Det ses at vanlig TV og telefon, som var tatt med som kontrollspørsmål, lå høyest. Deretter kommer nye telefontjenester, telealarm og teleundervisning mot lærer og først deretter flere TV-kanaler. Alle tjenestene, unntatt teleavis som lå på rundt 0, får en positiv vurdering.

Det ble parallelt i prosjektet foretatt en markedsundersøkelse blant vanlige husstander. På spørsmål om ønskelighet for et utvalg av tjenestene, var den vanlige bruker noe mer tilbakeholden enn ekspertene.

Hvorledes ønskeligheten for disse tjenestene er i dag, er ikke så lett å svare på, men det er grunn til å tro at den ut fra forholdet mellom tjenestene ikke er så forskjellig.

3.5 Bruk av tjenestene

På grunnlag av spørsmål om forventet bruksfrekvens og brukslengde var det mulig å beregne gjennomsnittlig tidsforbruk på enkelte tjenester. Det forutsettes ikke her at en abonnent abonnerer på

samtliges tjenester. Resultatet av beregningen er vist i figur 3.6.

Tidsforbruket som er angitt er relatert til en husstand. I 1976 var det gjennomsnittlig 2.9 personer i en husstand. Dette tallet er nå betraktelig lavere – noe som ut fra en 1993 betraktning vil redusere på det angitte tidsforbruk.

3.6 Introduksjonstidspunkt for tjenestene

Et av de mest interessante spørsmålene i undersøkelsen var introduksjonstidspunkt for de ulike tjenester. Resultatene fra undersøkelsen er angitt i figur 3.7.

I figuren er det oppgitt median og kvartiler. Dette er en vanlig fremstillingsform i Delphiundersøkelser. Årsaken til at det vanligvis brukes median og kvartiler fremfor gjennomsnittsverdi og standardavvik er at spesielt grove misforståelser og feilvurderinger kan føre til kvantifiserte svar som kan påvirke en gjennomsnittsverdi betraktelig. En median som er det samme som midtobservasjonen, er robust i forhold til slike outliers. Det samme er kvartilene som er det samme som 25 prosent- og 75 prosentpersentilen. Det ses av figuren at det er disse kvartilene som brukes. Det betyr at 50 % av alle svarene ligger mellom 25 % og 75 % persentilen som i en Delphiundersøkelse kan betraktes som et konfidensintervall.

Flere TV-kanaler – større programtilbud
Det har vært en kontinuerlig utvikling på dette feltet. Kabelfjernsynsnettene har med sine store antenner tatt inn et stadig større programtilbud. Prisen på parabolantennen for den enkelte bruker har gått ned, noe som førte til en stadig større etterspørsel i 80-årene. På slutten av 80-årene ble også de første kringkastingssatellittene skutt opp, hvilket betinger enda mindre diameter på parabolantennene.

TV-program mot særskilt betaling
Innført i 1985. Kryptering og adgangskontroll blir nå i stigende grad benyttet.

TV-program fra programbibliotek
Er ennå ikke realisert.

Teletekstoverføring på TV-skjerm
Denne tjenesten ble kalt tekst-TV og kom omtrent midt på 80-tallet.

Innhenting av informasjon fra databank
Denne tjenesten ble kalt teledata. Etter en lengre prøveperiode ble den kommersiell et stykke ut på 80-tallet.

Gjennomsnittsverdier fra
Delphi-runde 2

Selvbetjening over telenettet

Tjenesten ble definert ved at en ut fra stillbildeinformasjon skal kunne foreta bestillinger. Vi har siden starten på 80-tallet hatt elektronisk betalingsformidling. Dette har imidlertid ikke vært fra betalingsterminaler i hjemmet. Slike transaksjoner kan nå gjøres ved postdata. Her er det ikke inkludert levende bilde. I Delphiundersøkelsen antas introduksjonstidspunkt mellom 1990 og 2000. En mer avansert tjeneste basert på sekvenser av levende bilde er nå under utredning. Den krever imidlertid et mer avansert nett enn det vi har i dag.

Teleundervisning mot lærer

Tjenesten har vært i utprøving i lang tid og må kunne sies å være introdusert.

Teleundervisning mot datamaskin

Denne tjenesten krever kun PC og datatilknytning og har i lengre tid vært realiserbar på datanettene. Datex ble innført i 1980 og datapak i 1985. Prisen vil i 1994 gå kraftig ned ved introduksjon av ISDN.

Teleavis på abonnement

Forløpig ikke realisert.

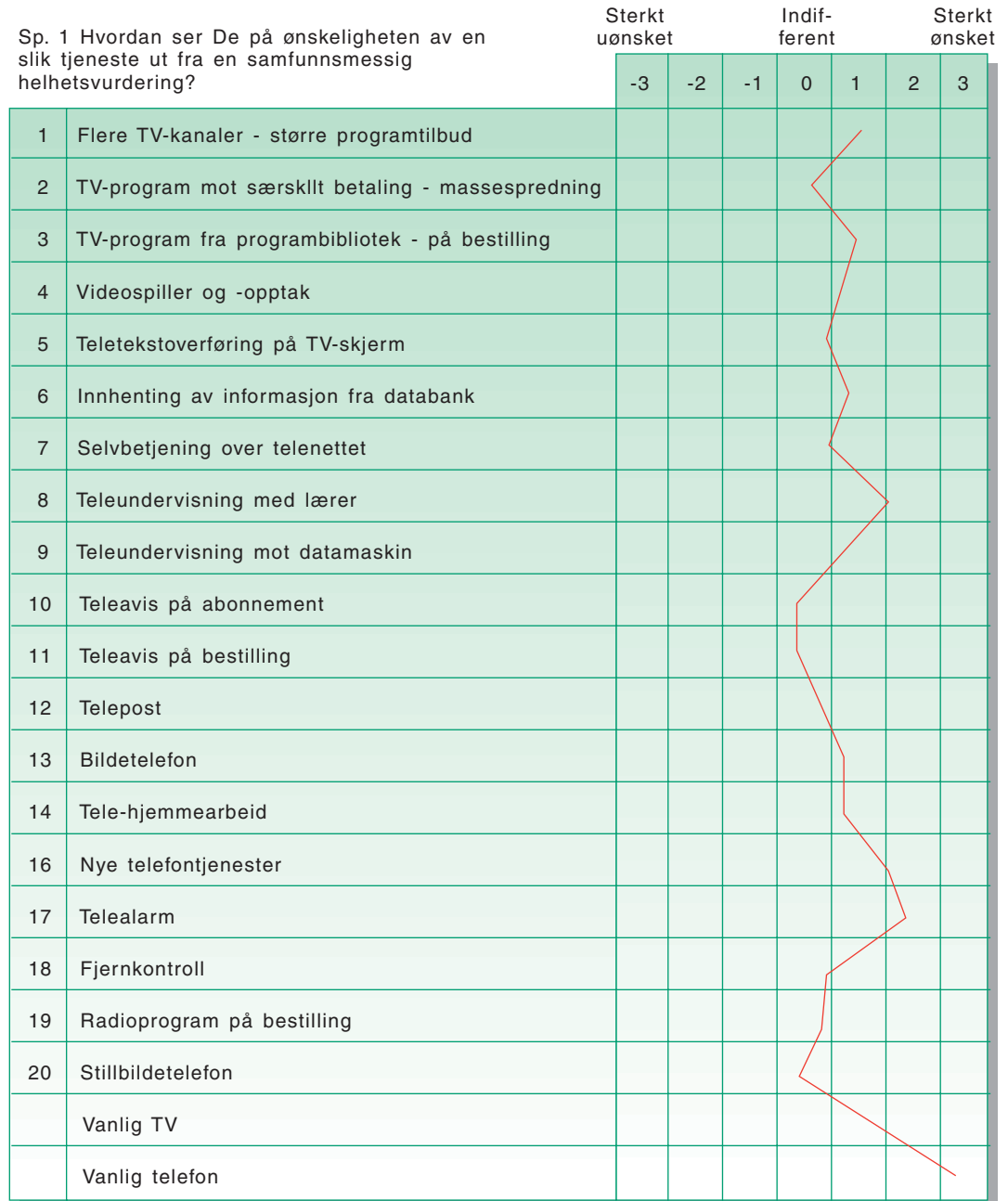
Teleavis på bestilling (spesialisert stoff fra flere aviser etter eget ønske)

Foreløpig ikke realisert.

Telepost

Har i prinsippet hatt tjenesten i mange år. Nå begynner tjenesten å bli attraktiv for hjemmene fordi faksmaskinene har sunket i pris og fordi ISDN gir mulighet

Sp. 1 Hvordan ser De på ønskeligheten av en slik tjeneste ut fra en samfunnsmessig helhetsvurdering?



Figur 3.5 Ønskelighet for teletjenestene

for flere tilknytninger på samme nummer.

Bildetelefon

Kommersielt tilgjengelig i 1994 ved innføring av ISDN.

Telehjemmearbeid (Fjernarbeid)

Har vært tilgjengelig etter innføring av datanettene på første halvdel av 80-tallet.

Telealarm

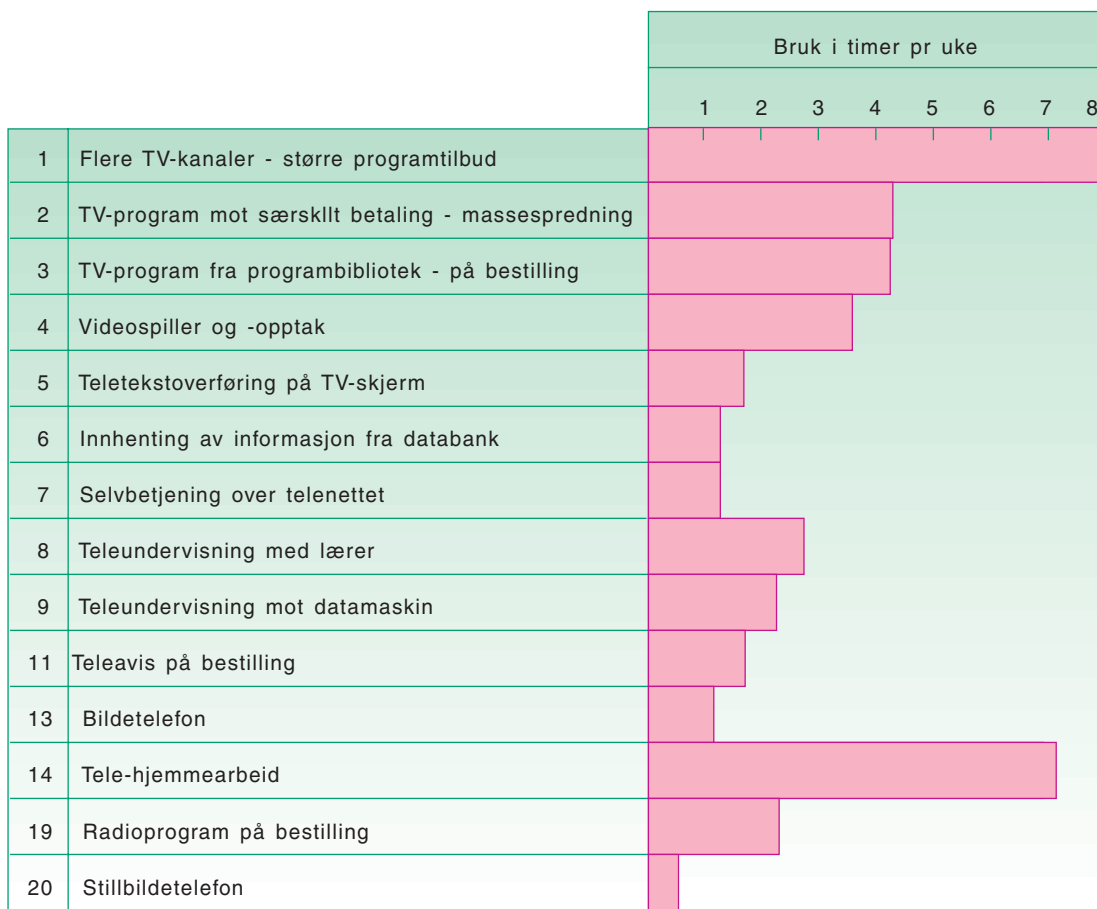
Tilgjengelig fra 1980.

Fjernkontroll

Tjenesten har ikke noe med telenettets funksjonalitet å gjøre. Ble forholdsvis tidlig mulig ved bruk av telenettet.

Radioprogram på bestilling

Ennå ikke realisert.



datatorget som blant annet tilbyr næringslivsinformasjon, offentlig registre, elektronisk telefonkatalog, databaser, etc.

Televerket drev de siste år på 80-tallet forsøksvirksomhet med teleundervisning med lærer. Utprøving av tjenesten fortsetter også i samarbeid med andre institusjoner. Det ses også på undervisningsformer i desentrale "studioer". Ved kommersiell introduksjon av ISDN i 1994 vil tekniske muligheter for å få lyd-, data- og bildefor-

Figur 3.6 Tidsforbruk for ulike tjenester

Stillbildetelefon

Dette navnet brukes ikke. Men funksjonen er fullt realiserbar eksempelvis i ISDN, som har vært i prøvedrift i flere år og som kommersielt innføres i 1994.

Resultatene fra Delphiundersøkelsen når det gjelder introduksjonsår, må sies å være overraskende bra, når det tas hensyn til at denne undersøkelsen ble utført i 1976 med det utgangspunktet en hadde på den tiden.

3.7 Utbyggingsgrad og etterspørsel

Vanligvis vil vi ha en prøveperiode i tilknytning til innføring av en ny tjeneste. Deretter settes den i kommersiell drift. Et godt eksempel på dette er ISDN som er et nett som muliggjør nye tjenester. Det har nå i lengre tid vært i prøvedrift. Et utsagn om når en tjeneste introduseres kan derfor tolkes på flere måter.

På figur 3.8 til 3.15 er det vist hvorledes ekspertene for et utvalg av tjenestene i Delphiundersøkelsen vurderte utviklingen av etterspørselen (markedsdekning) og utbyggingsgraden etter introduksjon av respektive tjenester.

Flere TV-kanaler – større programtilbud ble i Delphiundersøkelsen definert som mulighet til å få tilgang til utenlandske sendinger samt et eventuelt norsk TV2. Det med å ta inn utenlandske sendinger dekkes nå opp av kabelfjernsynsnettene samt med de som har egne parabolantenner. For øvrig vil vi nå om kort tid ha praktisk talt full dekning for et norsk TV2. Figur 3.8 gir en god beskrivelse av denne utviklingen.

Så langt har vi ikke noe tilbud på TV-program fra programbibliotek eller video on demand, som tjenesten også kalles. Det finnes imidlertid tekniske løsninger ved bruk av IN-noden som muliggjør tjenesten. I USA er tjenesten på forsøksstadiet. Anslagene i figur 3.9 er derfor noe optimistiske.

Innhenting av informasjon fra databank kom relativt tidlig til Norge. Ideen ble hentet fra Prestel-tjenesten i England. Tjenesten fikk navnet Teledata og forsøk ble startet tidlig på 80-tallet. Tjenesten var basert på en kombinasjon av bruk av telefon og fjernsynsskjerm. Denne tjenesten tok imidlertid aldri skikkelig av. Vi har imidlertid i dag teletorgtjenester som gir ekspertinformasjon muntlig og vi har

bindelse til en lærer være mulig. Dette krever en ISDN grunntilknytning samt et opplegg med lærer. Figur 3.11 gir en god beskrivelse av utviklingen. Vi vil imidlertid raskere komme opp i 100 % utbyggingsgrad.

Alle er kjent med den svært raske utviklingen i etterspørsel etter og bruk av faks i næringslivet. Det vil etter hvert skje en parallell utvikling i privatmarkedet som så vidt har startet opp. Muligheten til å bruke faksmaskin hjemme kom tidligere enn det som er angitt på figur 3.12, men start av en signifikant etterspørsel har det ikke vært tidligere. Nettet har imidlertid generelt vært utbygd slik at vi i lengre tid har hatt 100 % utbyggingsgrad.

Som nevnt blir ISDN etter en pilotfase på noen år kommersielt tilgjengelig i 1994. På forholdsvis kort tid vil infrastrukturen bli bygget ut slik at alle som ønsker ISDN skal kunne få abonnement. Det betyr at også bildetelefon blir tilgjengelig. Figur 3.13 treffer rimelig godt på denne utviklingen. Utbyggingsgraden vil raskt gå opp i 100 %, men det er grunn til å tro at etterspørselen de første årene før masseproduksjon av bildetelefon vil være noe lavere.

Det kan være vanskelig med noen entydig definisjon av hvilke tekniske og kommunikasjonsmessige krav som stilles for at arbeid i hjemmet skal kunne defineres som fjernarbeid. Ulike arbeidsfunksjoner vil eksempelvis kreve ulike typer tjenester. Derfor er det også vanskelig å vurdere kurvene i figur 3.14. I Delphiundersøkelsen ble telehjemmearbeid (fjernarbeid) definert som arbeidsplass med mulighet for overføring av tekst, lyd og stillestående bilde. Ved introduksjon av ISDN i 1994 vil vi få mer avanserte muligheter til en forholdsvis akseptabel pris.

Nye telefontjenester ble i Delphiundersøkelsen definert som en pakke med tjenester som blant annet inneholdt kortnummervalg, ventekopling, medflytting og viderekopling i tillegg til et knappsatsapparat. Figur 3.15 gir en god beskrivelse av utviklingen. Tjenestene kom med den første 10C sentralen som ble installert i 1976. Etter en forsiktig startfase har det vært en jevn utvikling av disse tjenester ved innføring av 10C sentraler og digitale sentraler og ved bruk av CHATS utstyr.

3.8 Evaluering av resultatene fra Delphiundersøkelsen

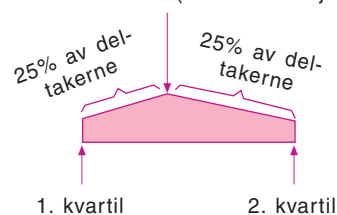
Alt i alt må det sies at de prognosene som ble laget, har vært gode og kanskje bemerkelsesverdige gode når det tas hensyn til på hvilket tidspunkt prognosene ble laget og de forutsetninger og den bakgrunnsinformasjon som en hadde på den tiden. Det var tross alt i en periode hvor Televerket kun hadde tjenestene telefon, teleks, datel og telegram og samtidig ingen markedsaktiviteter. Storparten av tjenestene ble definert og beskrevet i prosjektet, da denne type informasjon ikke var kjent i Norge.

4 Delphiundersøkelse i TITAN-prosjektet

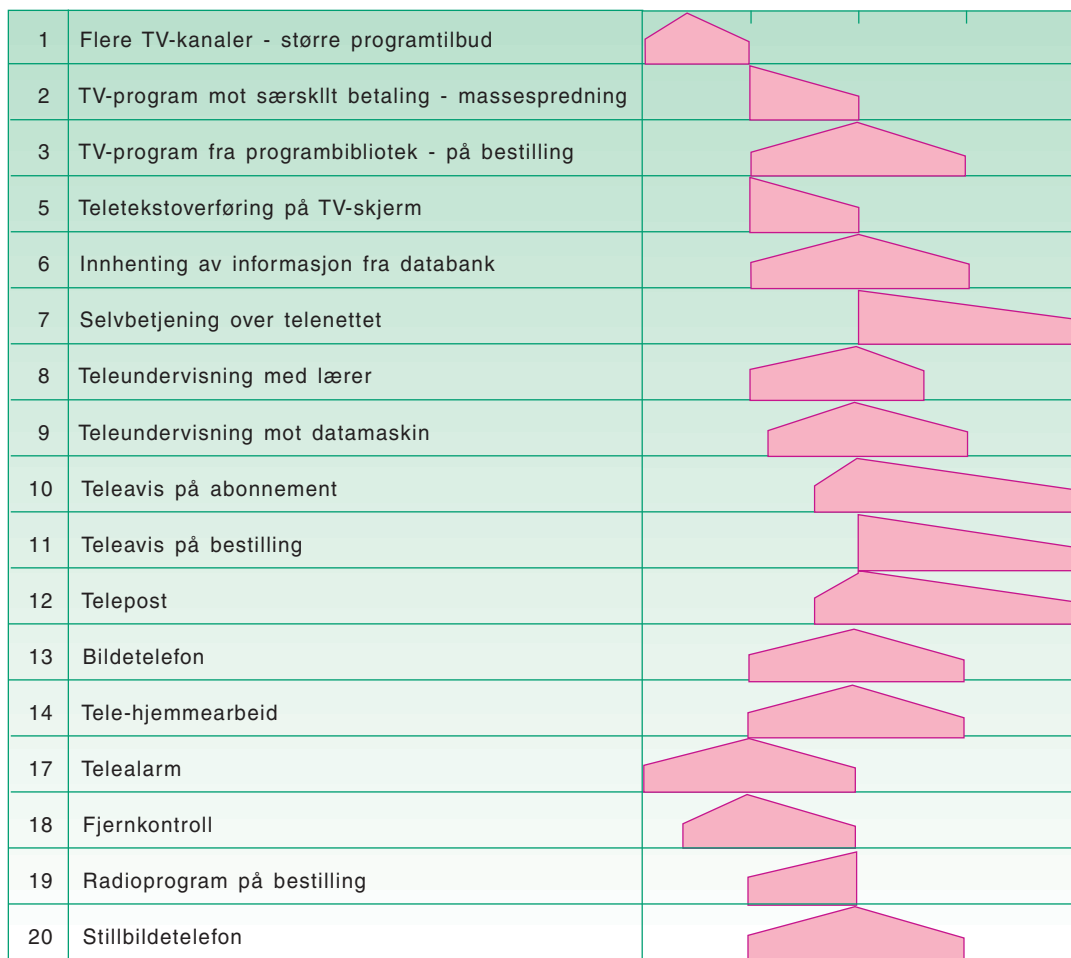
4.1 TITAN-prosjektet

TITAN-prosjektet er et RACE-prosjekt [3]. Målsettingen med prosjektet er å utviklet et verktøy for å beregne utbyggingskostnader i abonnentnettet basert på innføring og bruk av ny teknologi i form av optiske nettkomponenter. Som underlag er det nødvendig å gå igjennom og vurdere ulike nettarkitekturer – både nåværende basert på eldre teknologi som et referansegrunnlag og ny arkitektur ved bruk av optiske nettkomponenter.

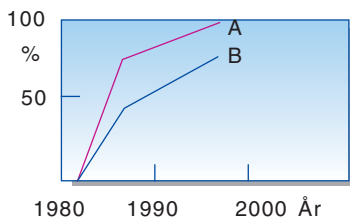
årstall for median (midtobservasjonen)



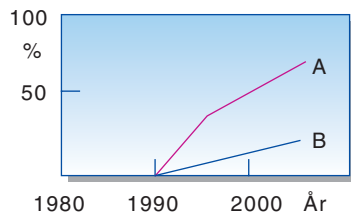
1980 1985 1990 1995 2000



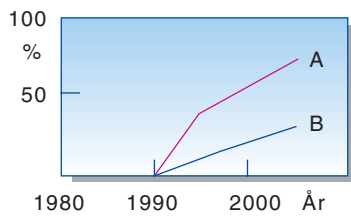
Figur 3.7 Introduksjonsår – median og kvartiler for tjenestene



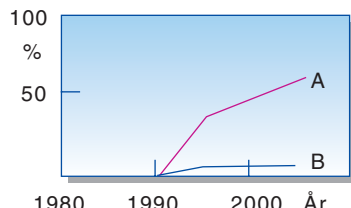
Figur 3.8 Flere TV-kanaler – større programtilbud
A: Utbyggingsgrad,
B: Etterspørsel



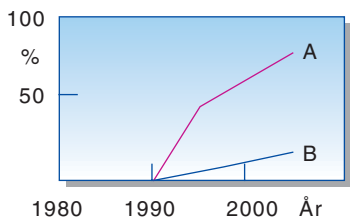
Figur 3.13 Bildetelefon
A: Utbyggingsgrad,
B: Etterspørsel



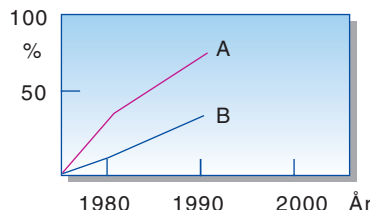
Figur 3.9 TV-program fra programbibliotek
A: Utbyggingsgrad,
B: Etterspørsel



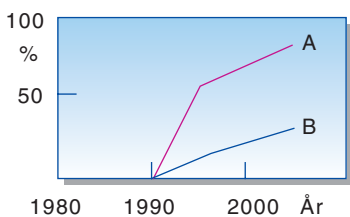
Figur 3.14 Telehjemmearbeid (Fjernarbeid)
A: Utbyggingsgrad,
B: Etterspørsel



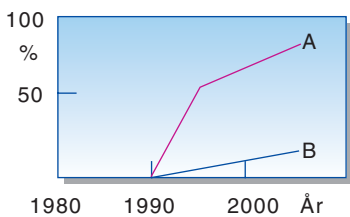
Figur 3.10 Innhenting av informasjon fra databank
A: Utbyggingsgrad,
B: Etterspørsel



Figur 3.15 Nye telefontjenester
A: Utbyggingsgrad,
B: Etterspørsel



Figur 3.11 Teleundervisning mot lærer
A: Utbyggingsgrad,
B: Etterspørsel



Figur 3.12 Telepost
A: Utbyggingsgrad,
B: Etterspørsel

I prosjektet bygges det også opp en database med priser og prisutvikling for aktuelle nettkomponenter. Ut fra denne databasen lages det også prognoser for fremtidige priser på nettkomponentene [2].

For å kunne vurdere ulike arkitekturer og utbyggingsstrategi og utbyggingsvolum er det helt nødvendig å ha kjennskap til de tjenester og applikasjoner som vil bli etterspurt og til hvor stor etterspørselen vil være. Med andre ord, det må lages prognoser for etterspørsel etter de ulike tjenestene. I TITAN-prosjektet har en kun mandat for å lage prognoser for privatmarkedet. Dette er for hver enkelt tjeneste gjort ved bruk av potensialbetraktninger og metningsmodeller, og det er gjort ved bruk av en Delphiundersøkelse. Det er den siste undersøkelsen som vil bli omtalt her.

4.2 Delphiundersøkelse

Delphiundersøkelsen gjennomføres nå ved at det er sendt ut 100 spørreskjemaer til 10 eksperter i hver av de 10 deltak-

ende landene i TITAN-prosjektet. I denne artikkelen trekkes ikke de nye resultatene frem, men det ses isteden på den angrepsmåten som er benyttet, samtidig som det er naturlig å sammenlikne de to Delphiundersøkelsene fra henholdsvis 1976 og 1993.

Spørreskjemaet i Delphiundersøkelsen består av:

- Generell informasjon om undersøkelsen
- Tekniske forutsetninger og opplysninger
- Økonomiske forutsetninger og opplysninger
- Oversikt over de ulike tjenestene
- Selve spørreskjemaene – ett skjema for hver tjeneste.

De tjenestene som på sikt er aktuelle for privatmarkedet og som er definert i prosjektet, er følgende:

- Teleinteraksjon, underholdning
 - Video on demand
 - Multimedia telespill
 - Bildetelefon
 - Telemedisin
- Telekjøp, elektronisk marked
 - Vare- og billettbestilling
 - Avertering og marketing
- Temakanaler og informasjon
 - Interaktivt TV og spesialiserte kanaler
 - Elektronisk avis
- Fjernarbeid og fjernundervisning
 - Fjernarbeid – enkelt kontor
 - Fjernarbeid – avansert kontor
 - Fjernundervisning i hjemmet
 - Fjernundervisning i studio

Et meget avgjørende punkt med hensyn til kostnader er den kapasiteten som tjenestene vil kreve. Det er delt opp i følgende aktuelle hastighetsklasser:

- 64 kbit/s
- 2 Mbit/s (Vidbånd)
- 8 Mbit/s (Bredbånd).

Det forventes at også høykvalitetsfjernsyn på sikt ved bruk av optimal koding skal kunne overføres på 8 Mbit/s.

I TITAN-prosjektet ble det foretatt beregninger for de investeringer som må gjøres ut fra etablering av tilpasset nettarkitektur og ut fra antakelse om en gitt fyllingsgrad, om avskrivningstid på investeringskostnader, om driftsutgifter og om et gitt overskudd. Deretter ble det formulert spørsmål om betalingsvillighet basert på prisintervaller der det også tas hensyn til prognostisert prisnedgang på sikt.

I figur 4.1 er det som eksempel vist hvorledes en tjeneste beskrives i Delphiundersøkelsen. Det er her tatt utgangspunkt i video on demand. Først er det gitt en generell beskrivelse av tjenesten. Deretter informeres det om de tekniske forutsetningene på en forholdsvis ukomplisert måte. Her skilles det mellom vidbånd og bredbånd. Stort sett kan de ulike tjenestene overføres på begge hastighetsklasser. Det vil hovedsakelig være kvaliteten som endres.

Deretter gis det en illustrasjon av tjenesten og til slutt noen eksempler på bruk av tjenesten.

Deretter kommer det en rekke spørsmål om hver tjeneste. Disse spørsmål er for enkelthets skyld nesten identiske for hver tjeneste. Spørsmålene er som følger:

- Angivelse av grunner for bruk av tjenesten
- Karakteristikk av typiske brukere av tjenesten
- Angivelse av forventet etterspørsel som funksjon av prisalternativer
- Angivelse av forventet etterspørsel som funksjon av tiden
- Anslag for hvor stor andel av disponibel inntekt en husholdning er villig til å betale for tjenesten som funksjon av husholdningens disponible inntekt
- Angivelse for alternativer til bruk av tjenesten.

Det er også skilt på vidbånd- og bredbåndskvalitet på tjenestene.

For øvrig er det gitt mulighet for å kunne gi spesielle kommentarer og argumentasjon i tilknytning til de enkelte spørsmål. I figur 4.2 er det gitt et eksempel på et utsnitt av spørsmålene for video on demand tjenesten.

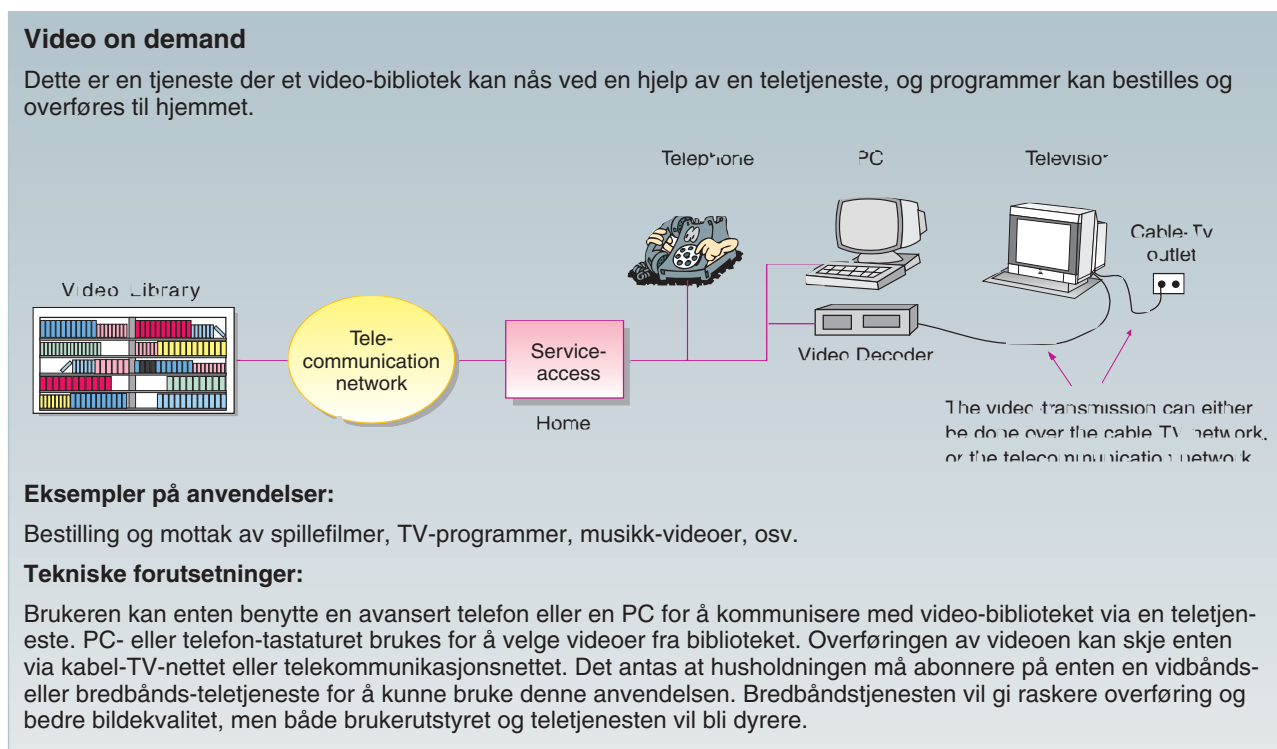
Til slutt i spørreskjemaet er det også laget "samlespørsmål" som kartlegger etterspørsel for den enkelte hastighetsklasse: 64 kbit/s, 2 Mbit/s og 8 Mbit/s.

4.3 Bruk av resultatene i TITAN-prosjektet

I Delphiundersøkelsen kartlegges etterspørselen etter de ulike tjenestene som en funksjon av pris. Samtidig finnes det også en direkte prognose for tjenestene for årene 1995, 2000, 2005 og 2010. Et metningsnivå for tjenestene angis også. I denne sammenheng er prognosespørsmålet det mest kompliserte å svare på. Det blir derfor lagt mest vekt på hva ekspertene mener de enkelte abonnenter er villig til å betale for tjenesten

I figur 4.3 og figur 4.4 er det gitt et eksempel på hvorledes resultater fra undersøkelsen kan brukes. Dette er foreløpige resultater fra første runde i Delphiundersøkelsen, og de er kun basert på 10 norske eksperter og kan i denne sammenheng ikke betraktes som signifikante resultater. De er imidlertid tatt med for å illustrere metodikken med bruk av denne type resultater. Tjenesten som er valgt ut her er video on demand.

Når det gjelder priser og kostnader, er det viktig å være klar over at trafikk-kostnader og trafikkvolum ikke er tatt med i dette prosjektet. Dette passer bra med hvorledes kostnadene allokeres i det norske telenettet. Alle kostnader som relateres til trafikkmaskinen – det vil si



Figur 4.1 Beskrivelse av video on demand

- 2) Of all households that might end up with a home office, what percentage of the user terminals and teleservice do you believe will be paid by the household, and what percentage will be paid by the employer?

Percentage paid by the household:	Percentage paid by the employer:
%	%

(Note: The two columns must add up to 100 %)

Comments (if any):

- 3) Given that the annual costs (1993 NOK) are as given in the table. What percentage of the residential market do you believe will demand a telecommunication service for this application?

(Note: Assume overall prices and income to be constant over time, and that there is a universal knowledge of the application. For explanation of annual costs, see page 6.)

In the case of wideband teleservices					
Annual costs:	5800 NOK	6600 NOK	8200 NOK	9900 NOK	19800 NOK
Answer (percentage):	%	%	%	%	%
In the case of broadband teleservices					
Annual costs:	15000 NOK	20000 NOK	25000 NOK	30000 NOK	60000 NOK
Answer (percentage):	%	%	%	%	%

Comments (if any):

- 4) What percentage of the residential market do you believe over time will end up with a telecommunications service for this application?

(Note: Assume a cost and a cost-trend which will be implicit in your demand predictions.)

In the case of wideband teleservices					
Year:	1995	2000	2005	2010	Maximum demand (Saturation)
Answer (percentage):	%	%	%	%	%
In the case of broadband teleservices					
Year:	1995	2000	2005	2010	Maximum demand (Saturation)
Answer (percentage):	%	%	%	%	%

Comments (if any):

- 5) Given that the household pays, and given the following set of annual disposable income of the household. How much of the income do you believe a family will be willing to spend on this application?

(Note: For explanation of annual disposable income, see page 6.)

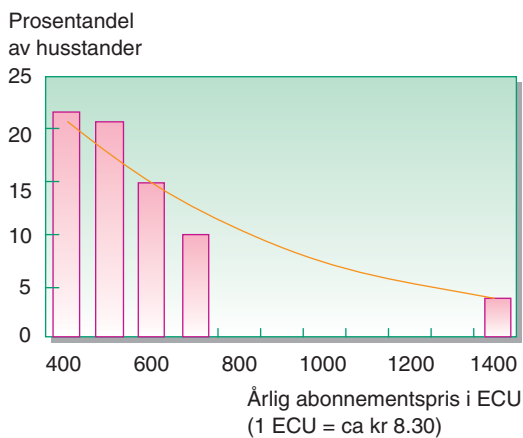
Disposable income:	66 000 NOK	100 000 NOK	165 000 NOK	290 000 NOK	410 000 NOK
Answer (NOK):	NOK	NOK	NOK	NOK	NOK

Comments (if any):

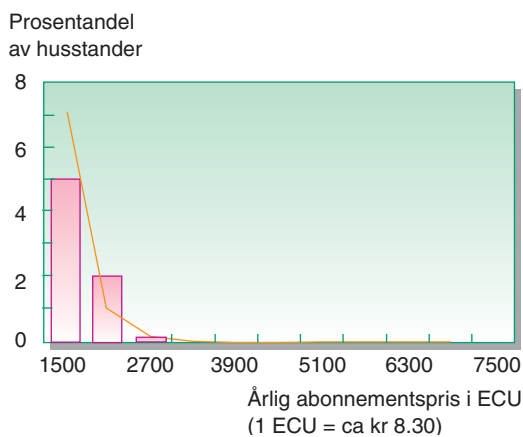
- 6) What kind of alternatives might exist for this application?

Answer:

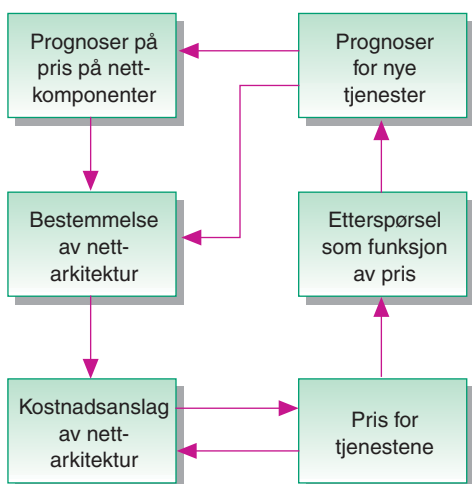
Figur 4.2 Spørsmål knyttet til video on demand tjenesten



Figur 4.3 Eksempel på estimert etterspørselskurve for video on demand med vidbåndstilknytning (2 Mbit/s)



Figur 4.4 Eksempel på estimert etterspørselskurve for video on demand med bredbåndstilknytning (8 Mbit/s)



Figur 4.5 Prosess for beregning av prognoser, kostnader og priser for ny infrastruktur i abonnentnettet

den delen av telenettet som ligger over abonnentmodulen i endesentralen – skal refunderes med trafikkinntekter. De øvrige kostnader – altså i abonnentnettet skal refunderes med abonnementsinntekter. Dermed holder det i TITAN-prosjektet å begrense seg til abonnementssetterspørsel, abonnementskostnader og abonnementspriser.

Som tidligere nevnt er det i TITAN-prosjektet bygd opp en database for pris på nettkomponenter. Ut fra denne databasen lages det også prognoser for prisnedgang på de ulike nettkomponenter. Ser vi på utbyggingsområder bestående av private abonnenter, vil det ut fra en gitt pris for en tjeneste kunne anslås en gitt etterspørsel.

Videre arbeid er ut fra etterspørsel etter ulike tjenester å bestemme en hensiktsmessig nettarkitektur. Dermed kan investeringskostnader og øvrige kostnader beregnes ved bruk av det utviklede beregningsverktøy. Spørsmålet er så om de kostnader som er beregnet harmonerer med de priser som opprinnelig er antatt for tjenestene. Dersom det er betydelig uoverensstemmelse her, må en gå inn et annet sted på etterspørselskurven og gjenta hele prosessen. Samme prosess benyttes også ved ulike tidsintervaller etter hvert som prisprognosene avtar. På den måten lages det også prognoser for etterspørselen etter tjenestene.

Det lages altså ikke direkte prognoser, men det lages indirekte prognoser for de ulike tjenester basert på etterspørselskurven fra Delphiundersøkelsen og prisprognoser sammen med kalkulasjon om kostnader ved oppbygging og drift av nettet. Figur 4.5 gir en oppsummering av dette. Det samme opplegget kan også brukes i blandede områder med forretningsabonnenter og private abonnenter og i områder med bare forretningsabonnenter, men det betinger informasjon om etterspørsel i bedriftsmarkedet på samme form som for privatmarkedet.

5 Sammenlikning av de to Delphiundersøkelser

Formen på de to undersøkelsene er svært like. Spørreskjemaene er laget nesten på identisk måte. Dette skyldes ikke minst den positive erfaring en hadde med den første Delphiundersøkelsen.

Ellers kan det konstateres at tjenestespekteret i to undersøkelser er overraskende likt. Ses det bort fra telegame og telecommunity (telemedisin) så er tjen-

estene fra 1993-undersøkelsen også definert i Delphiundersøkelsen fra 1976.

Konklusjonen er at i situasjoner der en ikke har historiske data om etterspørsel etter nye teletjenester, vil en Delphiundersøkelse være en av de aktuelle kvalitative prognosemetoder.

Referanser

- 1 Trælnes, T, Stordahl, K. *Behovsstudie for nye teletjenester i hjemmene*. Kjeller, Televerkets Forskningsinstitutt, 1976. (TF-rapport 19/76.)
- 2 Stordahl, K, Olsen, B T. Prognoser for pris på nettkomponenter basert på utvidelse av lærekurve modellen. *Elektronikk* 90(1), 166–172, 1994 (denne utgave).
- 3 RACE 2087 TITAN. Deliverable #2.

Underlag for prognoser for telefonabonnement: Data og analyse

AV CARLO HJELKREM OG JOHANNES BØE

1 Data og definisjoner

Det mest naturlige utgangspunkt for en analyse som skal munne ut i prognoser, er definisjonen av den (de) variable som skal framskrives. I Televerket er det TELSIS/AB, databasen som inneholder opplysninger om abonnentene og telefonabonnementene, som er statistikkgrunnlaget. Der er abonnementstypene definert slik:

- *Hovedabonnement*: Sum av eneabonnement, partsabonnement¹, abonnement på bylinje til hussentraler og telefonapparatsystemer, abonnement på innvalgsnummer og tjenesteboligabonnement. Forkortes HA.
- *Telefontilknyttinger*: Sum av HA, korttidsabonnement, samtale-automater, talestasjoner¹, tjenestebolig- og tjenestearbeidstelefoner, grønne nummer, teletorg, åpen linje, telekonto og ekstranummer til viderekopling (CHATS). Forkortes TTK.
- *Totalt antall tilknyttinger*: Sum av telefontilknyttinger, start-, stopp- og samlenummer m m og prøvenummer. Forkortes TAT.

Status for antall innkoblede abonnement fordelt på abonnementstyper finner vi i TELSIS/AB Rapport 35 (R35). Et eksempel på denne rapporten er vist i tabell 1. Den viser status for Region Oslo pr 01.01.93.

Som vi ser i tabell 1 tilsvarer altså "Totalsum" det vi ovenfor har kalt "Totalt antall tilknyttinger". Siden Totalsum inneholder alle typer tilknyttinger, har vi valgt å bruke begrepet "Totalt antall tilknyttinger" (TAT).

Hvilke av disse abonnementstypene det bør lages prognoser for, vil avhenge av hva prognosene skal brukes til.

- I *økonomi-/budsjettssammenheng* er det viktig å ha oversikt over HA, da det er disse som i hovedsak genererer inntekter.
- For *økonomisk analyse* (inkl investeringsanalyser) vil det også være av interesse å fokusere på TTK og TAT, da disse strengt tatt også representerer belagt kapasitet i sentraler/nett som bør forrentes. TTK er for øvrig også det

begrep som benyttes i *internasjonale sammenlikninger*.

- I *teknisk planlegging/oppfølging* er det TAT som er viktigst, fordi sentralene må dimensjoneres etter alle tilknyttingstyper.

Tradisjonelt har det vært to måter å måle abonnementene på i TELSIS/ AB: Tilgang og etterspørsel.

- *Tilgang* måler antall abonnement/tilknyttinger som i løpet av en gitt periode koples til sentralene/nettet.
- *Etterspørsel* måler antall abonnement/tilknyttinger som på et gitt tidspunkt, eller i løpet av en gitt periode, etterspørres fra abonnentene/Televerket.

Generelt kan en si at avviket mellom etterspørsel og tilgang er representert ved "den flytende ordremasse". Denne består

av ueffektuerte ordrer og ordrer som er effektuert men ikke ajourført i TELSIS/AB, og disse slås sammen til "ueffektuerte og ikke oppdaterte".

Nedenfor er de (gjeldende) presise definisjoner for tilgang og etterspørsel gjengitt. Som vi ser, er det også relevant å snakke om brutto og netto i sammenheng med tilgang. I forhold til etterspørsel er ikke bruttobegrepet klart definert.

- *Brutto tilgang*: Antall nye abonnement/tilknyttinger som i løpet av en periode er koplet til nettet.
- *Netto tilgang*: Brutto tilgang minus frakoplinger grunnet oppsigelser pluss tilkoplinger grunnet tilflyttinger minus frakoplinger grunnet fraflyttinger i samme periode.
- *Total etterspørsel*:

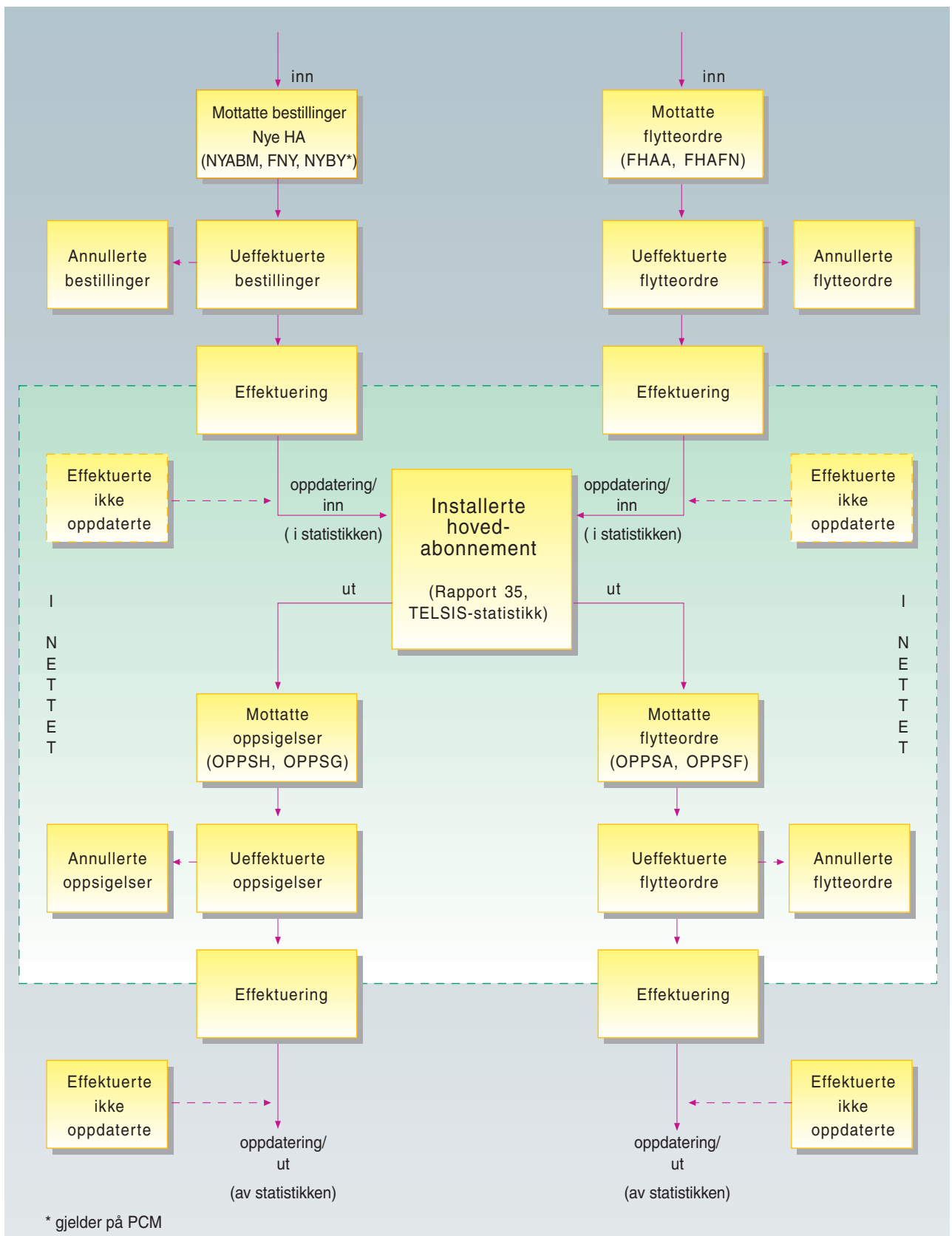
001.18:621.39

Tabell 1 TELSIS/AB Rapport 35

Region Oslo Abonnementskontoret Rapport nr: 35		Televerket	Dato: 04.01.93 kl: 11.26	
Antall hovedabonnement og andre abonnement pr 01.01.93 for hele regionen				
		I alt	Av disse Bylj. hussl/ Appsys	Innv. nr
Abonnementskategori				
1	Hovedabonnement	456677	67815	10359
2	Korttidsabonnement	471		
3	TVs samtaleautomater	2706		
4	Talestasjoner	3		
5	Tjenesteboligtelefoner	1118		
6	Tjenestearbeidstelefoner	4383	505	106
7	Grønne nummer	647		
8	Teletorg	918		
9	Åpen linje	72		
10	Ekstranr. til viderekopl. (CHATS)	697		
11	Telekonto*	2090		
12	Start, stopp- og samlenummer m m	28		
13	Prøvenummer	3803		
14	Sum telefontilknyttinger (1–6)	469782		
15	Sum andre abonnement (2–8)	16936		
16	Totalsum (1–13)	473613	68320	10465

* Senere tatt ut av statistikkene. Se for øvrig kapittel 3.

¹ Markerer abonnement som ikke eksisterer i dag. Disse er tatt med for sammenlikning av historiske data.



Figur 1 TELSIS/AB flyttdiagram

Antall installerte abonnement/tilknyttinger
+ antall ueffektuerte og ikke oppdaterte nybestillinger
- antall ueffektuerte og ikke oppdaterte oppsigelser
+ antall ueffektuerte og ikke oppdaterte tilflyttinger
- antall ueffektuerte og ikke oppdaterte fraflyttinger
<hr/>
= Total etterspørsel etter abonnement
<hr/>
- (Netto) årlig/månedlig etterspørsel: Endring i total etterspørsel etter abonnement i løpet av et kalenderår/måned.

Legg merke til at total etterspørsel er et beholdningsbegrep (gir status for/ måler alle abonnement som etterspørres på et gitt tidspunkt), mens tilgang og årlig/månedlig etterspørsel begge er strømningsbegreper (sier noe om endringer i antall abonnement i løpet av en gitt periode). I tillegg kan bemerkes at total etterspørsel er et nettobegrep, siden abonnement som “går ut” trekkes fra de som “kommer inn”.

Begrepene etterspørsel og tilgang kan grafisk framstilles ved figur 1. Figuren viser hvordan HA “flyter” gjennom TEL-SIS/AB. Det samme gjelder selvsagt i prinsippet også for TTK og TAT. Figuren gir i tillegg en oversikt over enkelte ordrekoder i TEL-SIS/AB.

Den stiplede store boksen representerer nettet (eller sentralen) som abonnementet er tilknyttet. Som vi ser, vil en bestilling på nytt abonnement som regel først bli registrert som ueffektuert. Montøren vil så (forhåpentlig på avtalt tidspunkt) utføre tilkopling til nettet/sentralen. Da er ordren utført og abonnenten tilkoplek nettet. I TEL-SIS/AB skal da den såkalte “statuskoden” for denne ordren oppdateres, slik at det framgår at abonnementet er innkoplek. Imidlertid vil en rekke av de utførte ordrene ikke bli oppdatert/ registrert som innkoplek umiddelbart etter utførelse. Abonnementet blir altså ikke å finne i statistikken over innkoplede abonnement (R35) før etter oppdatering i TEL-SIS/AB.

Disse ordrene er i figuren representert ved den stiplede boksen med “Effektuerte, ikke oppdaterte”. Tilsvarende gjelder for ordrer på tilflyttinger (flytting til region, sentralområde e l).

En oppsigelsesordre (fra abonnenten selv eller gjort av Televerket pga gjeld) eller en ordre på fraflytting følger i prinsippet samme mønster: Ordren er etter effektivering frakoplek nettet, men i mange tilfeller ikke registrert i R35 (effektuerte, ikke oppdaterte).

Tilgang og etterspørsel vil i prinsippet utvikle seg noe forskjellig fra periode til periode. Spesielt gjelder dette månedsdata. Dette henger sammen med at tilgangsdata er sterkt påvirket av Televerkets arbeidsrutiner for inn-/ utkoplinger (som igjen blir påvirket av etterspørselen): Dimensjonering av ordremottak, montørstab, etterbehandling etc påvirker den takt Televerket makter å utføre ordrer og etterpå registrere at så er gjort. Etterspørsel korrigerer for (mye av) variasjoner som skyldes Televerkets rutiner, idet ueffektuerte og ikke oppdaterte også er inkludert. Dermed måler etterspørsel i langt større grad aktiviteten hos abonnentene (interne og eksterne).

Det er rapportene over registrerte innkoplede abonnement, dvs R35, som gir grunnlaget for statistikken for netto tilgang av abonnement. Som figur 1 viser, gir m a o ikke statistikken det fulle bildet av netto tilgang av abonnement, idet en ikke har lagt til de abonnement som er tilkoplek men ikke oppdatert, og tilsvarende ikke har trukket fra de abonnement som er frakoplek men ikke oppdatert. Tilsvarende resonnement kan gjøres for brutto tilgang av abonnement.

Figur 1 gjelder i prinsippet for region, sentralområde og endefordeler. På landsnivå antas at alle flyttinger til og fra regionene utlikner hverandre (flytting fra en region genererer flytting til en annen region), og dermed gjelder figuren også for landet totalt. Praksis viser imidlertid at flyttinger innen f eks et sentralområde på kort sikt ikke utlikner hverandre: Hvis abonnenten skifter adresse innen sentralområdet og “flytte-fra-ordren” behandles senere en “flytte-til-ordren”, risikerer vi at “flytte-fra-ordren” blir liggende som ueffektuert eller effektuert men ikke oppdatert (og er dermed fremdeles i R35), mens “flytte-til-ordren” er blitt utført og oppdatert. I dette tilfellet har vi i en kort periode et abonnement for mye i tilgangsdata. På sikt jevner dette seg selvsagt ut, mens vi likevel har dette problemet når vi skriver ut månedsstatistikken.

2 Definisjonsendringer

I prinsippet er det slik at data i utgangspunktet skal være sammenliknbare i alle perioder når en tidsserie anvendes i en statistisk prognosemodell. Hvis det har skjedd endringer i hvordan vi teller f eks HA eller etterspørsel (dvs definisjonene er endret), vil dette medføre at data ikke er direkte sammenliknbare i periodene før og etter en slik endring. Det kan da bli vanskelig å tolke hvor stor f eks veksten egentlig har vært i perioden før og etter definisjonsendringen. Likevel er det et faktum at det de siste årene har vært flere endringer i definisjonene, både mht abonnemestypene TTK og TAT, samt to endringer i definisjonen på total etterspørsel etter abonnement.

Avhengig av nettoeffekten av slike *definisjonsendringer*, vil de statistikk-messig være representert ved *hopp opp eller ned i data*. Region Oslo “mistet” f eks 911 HA som følge av definisjonsendringen for etterspørsel 01.04.90. Hvis data etter ny definisjon blindt etterfølger data etter gammel definisjon, vil en prognosemodell tolke dette som endring i etterspørselen, selv om dette i realiteten ikke er tilfellet. I tillegg vil den få problemer med å “føyte” seg (dvs beskrive den historiske utvikling) rundt de datapunkter som representerer definisjonsendringen, fordi det ikke er noe i data for tilsvarende periode årene før (eller etter) som tilsier at det skal komme en definisjonsendring. Dette øker usikkerheten i modellen og dermed også usikkerheten i prognosene.

2.1 Endringer i definisjonen for netto etterspørsel

La oss se på de ulike definisjonene som har eksistert for total etterspørsel etter abonnement/tilknyttinger de siste årene:

- <i>Perioden fram til og med 07.88:</i>
Antall installerte
+ antall ueffektuerte nybestillinger
+ antall ueffektuerte tilflyttinger
<hr/>
= Total etterspørsel
<hr/>
- <i>Perioden fra og med 08.88 til og med 03.90²:</i>
Antall installerte
+ antall ueffektuerte nybestillinger
+ antall ueffektuerte tilflyttinger
- antall ueffektuerte fraflyttinger

- + antall effektuerte, ikke oppdaterte nybestillinger
 - antall effektuerte, ikke oppdaterte oppsigelser³
 - + antall effektuerte, ikke oppdaterte tilflyttinger
 - antall effektuerte, ikke oppdaterte fraflyttinger³
-
- = Total etterspørsel

- Perioden fra og med 04.90:

Se definisjonen i kapittel 1.

Bakgrunnen for endringene i definisjonen finner vi i analyser av ordreflyten i TELSIS/AB og av diskusjoner om hva vi mener med begrepet netto etterspørsel. Ved overgang fra manuelt abonnementsarkiv til TELSIS/AB "forsvant" mellom 2000 og 3000 abonnement fra regionens abonnementsstatistikker. Ved nærmere analyser fant en at disse abonnementene befant seg i en "gråson" mellom to statistikkrapporter:

De var trukket ut av rapporten for ueffektuerte ordrer fordi de var blitt effektuert (dvs avhengig av ordretype var de fysisk inn- eller utkoplet), men var ikke kommet inn i statistikken over inn-/utkoplete abonnement (dvs statistikken for antall installerte abonnement) pga manglende oppdatering i TELSIS/AB. Denne "gråsonen" er representert ved de effektuerte, ikke oppdaterte i definisjonen over.

Analysen av "gråsonen" førte til definisjonsendring 01.08.88. Den videre analyse av ordreflyten i TELSIS/AB, samt klargjøring av begrepet netto etterspørsel, resulterte i figuren som representerer ordreflyten i TELSIS/AB, dvs figur 1, og definisjonsendring 01.04.90.

Som et eksempel på "hopp i data" som følge av definisjonsendring, viser figur 2 den månedlige etterspørsel etter HA i Region Oslo for perioden 1987–1990. Tidspunktene for definisjonsendringene er merket av.

Ser vi på datapunktet for august 1988 er det ikke mulig å se noe "hopp" her (i forhold til måneden før). Dette til tross for at definisjonen ble utvidet med netto effektuerte, ikke oppdaterte ordre. Grunnen til dette finner vi i regionens behandling av definisjonsendringen: For det første var dette den første kjente definisjonsendring som hadde funnet sted. Det hersket dermed tvil om hvordan denne endringen skulle håndteres statistikk-messig. For det andre ble definisjonsendringen ansett som en korreksjon av en feil som var oppstått som følge av innføringen av TELSIS/AB: Når netto effektuerte, ikke oppdaterte ordrer ble tatt inn i definisjonen, var det fordi TELSIS/AB ikke talte opp disse ordrene, mens de tidligere hadde vært inkludert i endringstallene for de installerte abonnement. Resultatet ble at kun endringen i de effektuerte, ikke oppdaterte ordrene (netto) ble tatt med i månedsstatistikken, og ikke før i oktober (netto antall effektuerte, ikke oppdaterte ordrer i oktober minus netto antall i september), siden månedlig etterspørsel jo likevel måler endringen i totaletterspørselen fra en måned til en annen! På denne måten ble definisjonsendringen "kamouflert" og en unngikk hopp i data. Likefullt er det en endring i hva som måles fra måned til måned i perioden etter oktober 1988, og dette er det viktig å være klar over når prognosemodeller skal bygges.

Går vi til april 1990 har vi derimot ingen problemer med å se ett "hopp" i data (fra perioden før). På det tidspunktet ble ueffektuerte oppsigelsene ble tatt inn i definisjonen utgjorde de 911 abonnement, og dermed "mistet" regionen tilsvarende antall: Statistikk-messig medførte dette at regionen fikk negativ etterspørsel denne måneden.

Av de mest vanlige kvantitative prognosemetodene har regresjonsanalyse utviklet metodikk som gjør

det mulig å takle problemer av typen nevnt ovenfor (ved bruk av såkalte dummyvariable). Derimot kan glattingsmodeller (som f eks Holt-Winters metode) og regresjonsanalyse hvor en "glemmer" problematikken gi dårlige resultater. I de siste tilfeller må en justere data, f eks ved å reprodusere historikken etter ny definisjon. Dette kan likevel vise seg å være vanskelig, da en som regel ikke har data som viser hvordan slike justeringer kan gjøres i tidligere perioder.

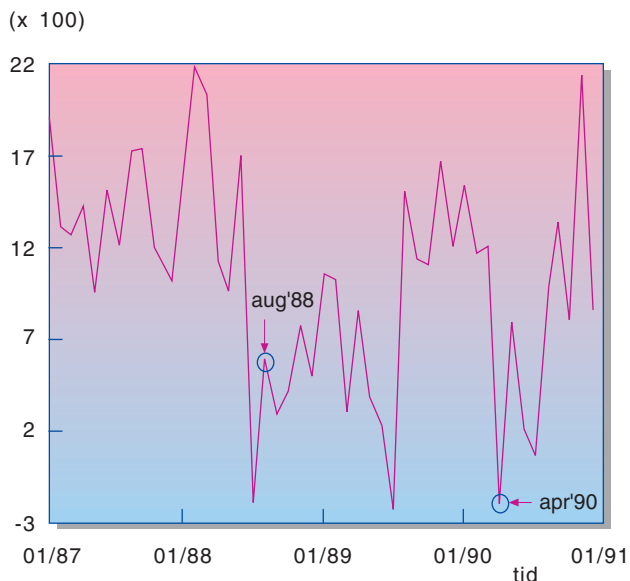
2.2 Andre definisjonsendringer

Ovenfor er virkninger av endringer i måten vi måler etterspørsel utdypet, uavhengig av hvilken abonnementsstype dette måtte gjelde. Endringer i hva som skal telles som HA, TTK eller TAT vil imidlertid kunne gi liknende effekter. Vi skal her kort nevne noen slike endringer.

I 1990 startet ISDN prøvetjeneste i Oslo, Bergen, Stavanger og Trondheim. Registrering av ISDN-aksessene ble da gjort i TELSIS/AB og innlemmet i telefonstatistikkene ved at både en 2B+d aksess og en 30B+d aksess ble talt som ett abonnement/tilknytting. Disse ble imidlertid tatt ut av statistikkene i september 1991⁴.

Går vi tilbake til figur 1 inneholdt ikke definisjonen for TTK t o m 1991 teletorg, åpen linje, grønne nummer, telekonto og ekstranummer til viderekopling (CHATS).

Telekonto ble innlemmet i september 1992, men er f o m april 1993 tatt ut av rapportene over telefonabonnement/-tilknyttinger. At telekonto ble innlemmet i statistikkene skyldes en feiltakelse: Telekonto-numrene representerer ikke telefonnumre i tradisjonell forstand, men representerer en "fiktiv nummerserie" som er å sammenlikne med et kontonummer.



Figur 2 Månedlig netto etterspørsel, etter HA, 1987–1990

² Markerer perioden hvor det var avvik mellom Region Oslo og Teledirektoratets definisjon.

³ Markerer elementer som ikke var spesifisert i Teledirektoratets definisjon. Dette avviker er i denne omgang ikke spesielt interessant, men er uthevet fordi det er Region Oslos data som er brukt i senere eksempler.

⁴ Region Oslo hadde på dette tidspunkt 250 2B+d aksesser og 21 30B+d aksesser. I forhold til den totale abonnementsmassen utgjorde m a o disse aksessene en beskjeden andel.

I januar 1992 ble teletorg, åpen linje, grønne nummer, telekonto og ekstranummer til viderekopling (CHATS) flyttet fra Start-, stopp-, og samlenummer m m til egne poster, se figur 1. I perioden før januar 1992 var disse m a o ikke inkludert i TTK, mens de både før og etter dette var inkludert i TAT.

Også andre endringer av enkelte statistikkdefinisjoner har vært foretatt i den senere tid, men effektene av disse har vært små⁵.

3 Netto etterspørsel – en variabel det er vanskelig å lage gode prognoser for

Definisjonen av total etterspørsel etter abonnement består av antall innkoplede abonnement og fire hovedstrømmer av ordre på abonnement. Av ordrestrømmene er to på vei inn i nettet: Nye bestillinger og ordrer på tilflyttinger. De andre to er på vei ut av nettet: Oppsigelser og ordrer på fraflyttinger. Ser vi bort fra de innkoplede abonnementene, er det nettoen av status for disse fire strømmene som pr definisjon utgjør begrepet total etterspørsel etter abonnement.

I perioden fra januar 1988, da TELSIS/AB ble grunnlag for offisiell statistikk i Region Oslo, og fram til den siste definisjonsendring i april 1990, er det klart at mangler ved definisjonen for etterspørsel gjorde at en ikke målte etterspørselen på riktig måte. Dette kan i og for seg forsvares hvis en vet hva en måler, og en i tillegg måler det samme hele tiden. Likevel gav definisjonen i denne perioden opphav til problemer fordi den gjorde begrepet netto etterspørsel avhengig av faktorer som ikke har med etterspørsel fra markedet å gjøre.

⁵ Et eksempel på dette kan være MAKROTEL-abonnement (en 30B+d aksess med redusert tjenestespekter) som ble trukket ut av statistikkene i april 1993. Disse utgjorde imidlertid da bare 11 aksesser i Region Oslo, og det var få andre regioner som tilbød MAKROTEL på dette tidspunktet. Et annet eksempel kan være statistikken over ueffektuerte og ikke oppdaterte ordrer (TELSIS/AB Rapport 34), som f o m september 1991 gikk over fra å telle ordrene til å telle abonnementene som ordrene inneholder. Dette medførte f eks en økning på 81 HA for september i Region Oslo.

Hva som menes med dette, forklares i det følgende.

Når definisjonen på netto etterspørsel mangler en eller flere av ordrestrømmene nevnt ovenfor, se figur 1, eller mangler en av "boksene" i en/ flere av ordrestrømmene, vil etterspørselsdata være avhengige av administrative rutiner i TELSIS/AB.

Dette er i utgangspunktet uheldig fordi vi med data for etterspørsel ønsker å måle variasjoner i etterspørselen fra markedet, og da må ikke data være påvirket av Televerkets rutiner. Ser vi f eks definisjonen som gjaldt fram til og med 07.88 (se kapittel 2), så manglet denne begge utstrømmene. Oppsigelser og fraflyttinger ble, i henhold til denne definisjonen, først regnet som etterspørsel etter at de var effektuert og oppdatert (i R35). Med andre ord kunne oppsigelser og fraflyttinger bli liggende lenge som ueffektuerte eller som effektuerte, ikke oppdaterte før de under oppdatering av statuskode kom inn i statistikken som fratrekk i netto etterspørsel. I slike tilfeller "lyver" altså statistikken for flere måneder, ja kanskje over flere år, fordi den ikke viser etterspørselen slik den virkelig har vært:

Etterspørselstallene blir for høye når oppsigelses-/ fraflyttingsordrer hopper seg opp og for lave når "slakken" tas inn. Som regel kan ikke data eller noen prognosemodell korrigerer for dette, og vi får derfor et skjevt bilde av historisk utvikling. Dette påvirker deretter en prognosemodells evne til å beskrive framtidig utvikling. Tilsvarende resonnerment kan også gjøres for mangelen i definisjonen i perioden 08.88 – 04.90.

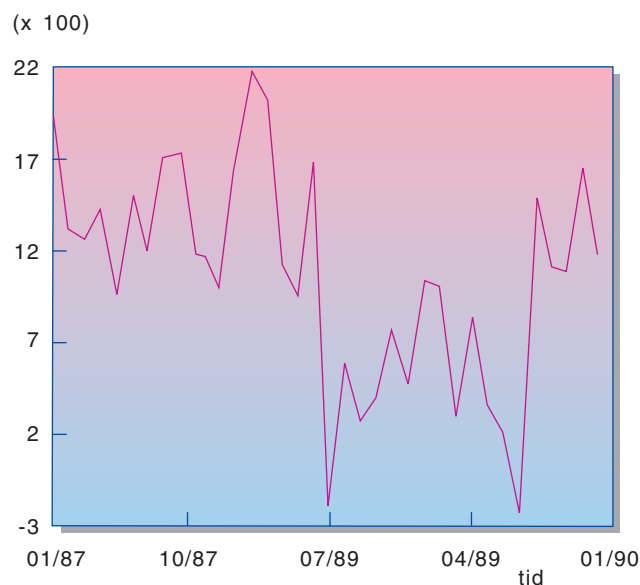
Selv om vi har oppnådd å rydde opp i definisjonen for netto etterspørsel, så er definisjonen fremdeles ikke helt uavhengig av arbeidsrutiner. Et uromoment som er uavhengig av den gjeldende definisjon, er inkassorutinene. Et eksempel som belyser denne problemstillingen er utviklingen i netto månedlig etterspørsel etter HA for perioden 1987–1989 i Region Oslo. Data for denne perioden er grafisk framstilt i figur 3.

I denne perioden hadde vi som før nevnt et skifte av definisjon i august 1988, som det ikke er lett å avdekke i denne figuren. Det vi imidlertid raskt blir slått av, er juli-observasjonene for 1988 og 1989. Disse

viser negativ netto etterspørsel for disse månedene. Kan dette være mulig?

Siden vi måler en nettostrøm er dette selvfølgelig teoretisk mulig hvis vi i perioden får flere oppsigelser og fraflyttinger enn nye abonnement og tilflyttinger. Ved nærmere ettersyn viste det seg at det var Region Oslos opprydding i 3–4 år gamle inkassosaker som var grunnen. I 1988 ble slike saker behandlet sentralt i regionen og deretter sendt til inkasso. Flere av disse sakene endte i oppsigelser og ble først fanget opp av definisjonen for netto etterspørsel i månedene juli– desember 1988, noe som gav utslag i lave tall for netto etterspørsel i disse månedene (observasjonene for tilsvarende periode i 1987 kan til sammenlikning anses for normale). Juli er feriemåned og har normalt få nybestillinger, slik at "inkasso-oppsigelsene" får netto-tallet til å tippe over på minussiden. I januar 1989 ble strategien med sentral behandling av slike saker oppgitt til fordel for en desentral fordeling. Oppryddingsarbeidet ble intensivert, og dette kommer spesielt til syne i data for juni og juli 1989. Sistnevnte observasjon blir derfor negativ av samme årsak som juli 1988.

Figur 4 viser månedlige (effektuerte og oppdaterte) oppsigelser i perioden 1987–1989. Tallene for 1987 kan anses som "normale". Vi ser helt klart hoppet i oppsigelser i juli 1988. Disse holder seg på et, inntil da, høyt nivå ut året. I 1989 forsterkes dette i perioden juni–november. At dette i figur 3 ikke kommer like godt til syne for alle disse



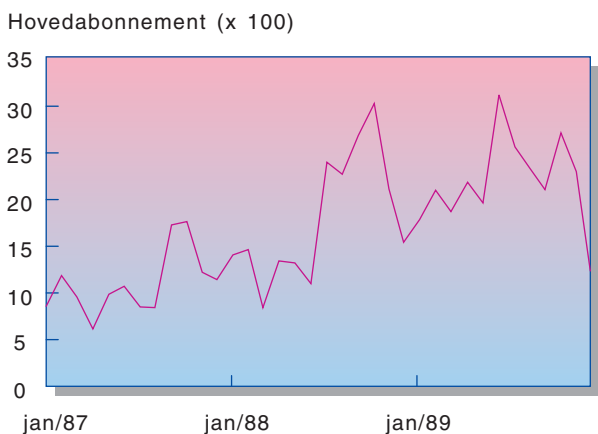
Figur 3 Utviklingen i netto månedlig etterspørsel etter HA for perioden 1987–1989 i Region Oslo

periodene, skyldes at dette skjules i nettoresultatet av alle de fire ordrestrømmene.

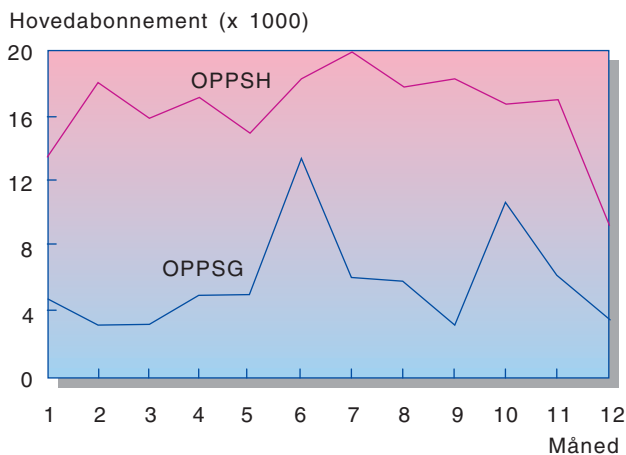
I 1989 fikk vi splittet statistikken for oppsigelser. På denne måten kan vi nå følge oppsigelser gjort av abonnentene selv (OPPSH) og oppsigelser grunnet gjeld (OPPSG). Sistnevnte tilsvarer inkassosakene nevnt ovenfor.

Figur 5 viser denne oppsplittede utviklingen pr måned i 1989.

Som vi ser, er det OPPSH som utgjør den største andelen av oppsigelsene i 1989. OPPSG viser på sin side et hopp i juni og et nytt i oktober. Går vi tilbake til figur 3, ser vi at dette klart gir seg utslag i juni (en måned hvor netto etterspørsel etter HA vanligvis øker i forhold til mai), mens oktober er overskygget av andre effekter.



Figur 4 Månedlige (effektuerte og oppdaterte) oppsigelser i perioden 1987-1989



Figur 5 Oppsplittet utvikling i oppsigelser pr måned i 1989. OPPSH = oppsigelser gjort av abonnentene selv, OPPSG = oppsigelser gjort av Televerket, grunnet gjeld

Konklusjonen for tilfellet med inkassorutinene blir som i tilfellet med sviktende definisjon for netto etterspørsel: Abonnement som skulle vært oppsagt i tidligere perioder, forkludrer data på en slik måte at vi som analyserer dem tror at etterspørselen er høyere enn den i virkeligheten er. Når det så ryddes opp, ser det ut som om etterspørselen er lavere enn den virkelige er.

Oppsummert kan vi altså si at det å lage prognoser for netto etterspørsel etter abonnementet er noe komplisert:

For det første består begrepet netto etterspørsel av 4 hovedstrømmer som i prinsippet kan utvikle seg helt forskjellig. Teoretisk burde en derfor lage separate prognoser for hver av disse strømmene. I praksis kan dette, av grunner nevnt ovenfor, vise seg vanskelig for oppsigelsene, men hovedstrømmene bør likevel følges nøye opp fra måned til måned slik at vi

har kontroll med hvorfor nettoen er blitt slik som den er. For det andre har oppsigelser grunnet gjeld vist seg å være påvirket av inkassorutinene. Rutinene for inkassofunksjonen bør derfor studeres slik at vi ikke opplever uheldige overraskelser i data.

Data for stengninger av abonnementet hjelper oss til å forutsi slike effekter: Inkassorutinene for Televerket sier at abonnementene skal stenges før de går til inkasso. En splittet statistikk for stengning, slik at antall stengte abonnement grunnet gjeld kan hentes ut som separat statistikk, vil med å være til stor hjelp som "ledende indikator" for abonnement som går til oppsigelse grunnet gjeld.

4 Klargjøring av data

I Televerket har det tradisjonelt vært laget etterspørselsprognoser for abonnementsutviklingen.

Arbeidet med forslag til etterspørselsprognosene begynner med et betydelig forarbeid i form av behandling av grunnlagsdata på sentralnivå. Status for total etterspørsel i hvert sentralområde finnes ved å sette sammen statistikk fra R35 (antall installerte abonnement) og R34 (ordrestatistikken, dvs antall ueffektuerte og ikke oppdaterte ordrer, gitt i TELSIS/AB Rapport 34). Dette korresponderer med definisjonen for etterspørsel gitt

ovenfor. Imidlertid er det ikke slik at disse rapportene gir den fulle og hele sannhet om etterspørselen i et gitt sentralområde. Visse justeringer og korrigeringer må også foretas før data er på en slik form at de viser hva etterspørselen i et bestemt år eller måned har vært og kan sammenliknes med de andre periodene.

Vi skal her vise hvordan vi arbeider oss fram til data på det vi kan kalle "riktig etterspørselsform". Med "riktig" menes her at data er ordnet slik at de på en mest mulig korrekt måte gjenspeiler den utvikling som etterspørselen virkelig har hatt. Når vi har oppnådd dette, eller kommet så nær opp til dette som mulig, kan vi benytte oss av de metoder som vi måtte ha til rådighet for å lage prognosemodeller som framskriver den historiske utvikling. De mest brukte og anerkjente metoder for kvantitative prognosemodeller er:

- Eksponentiell glatting (enkel, dobbel, trippel)
- Holts metode
- Holt-Winters metode
- Lineære regresjonsmodeller (enkel, multipl)
- Logistiske vekstkurver (Metningsmodeller/S-kurver)
- Box-Jenkins metode
- State Space-modeller/Kalmanfilter
- Transferfunksjon-modeller.

Valg av metode avhenger bl a av inngående kjennskap og erfaring med disse, prognosehorisont, antall variable som det skal lages prognoser for, datatilgang for forklaringsfaktorer, tid til disposisjon m m.

En annen klasse av prognosemetoder er de såkalte kvalitative metodene. De mest kjente er:

- Ekspertundersøkelser
- Delphimetoden
- Scenariometoden
- Analogimetoden
- Markedsundersøkelser.

De fleste av disse metodene er omtalt andre steder i denne utgaven av Teletronikk. [1] er en lærebok som gir relativt enkle introduksjoner til flere av disse metodene, men med størst vekt på de kvantitative.

Uansett hvilken prognosemetode eller modell som benyttes, må annen "modellekstern" informasjon brukes som kor-

rektiv til modell-prognosene. Dette kan være generelle forhold som vi vet at prognosemodellene ikke fanger opp, spesielle forhold ved teknikk og marked i framtidige perioder som ikke reflekteres i den historiske utvikling, informasjon fra (andre) markedsundersøkelser, lokale markedsforhold, osv.

4.1 Sammensetting av årsdata for total etterspørsel

Som en illustrasjon på det arbeidet som utføres med datagrunnlaget, har vi delt prosessen opp i 4 steg og benyttet arbeidet med status for total etterspørsel pr 31.12.1990 i Region Oslo.

Steg 1: Ukorrigert etterspørsel

Steg 1 er allerede nevnt ovenfor: *Vi legger sammen antall installerte abonnemnt og antall ueffektuerte og ikke oppdaterte abonnemnt fra henholdsvis R35 og R34.* Tabell 2 viser et eksempel på dette for TAT i Oslo kommune, fordelt på sentralområder.

Netto ueffektuerte abonnemnt inneholder i denne tabellen også effektuerte, ikke oppdaterte abonnemnt.

Steg 1 viser hvordan vi kommer fram til total etterspørsel pr 31.12.90. Ønsker vi i neste omgang å ha data for hva den årlige etterspørselen var i 1990, må vi gå til tilsvarende totalletterspørselstall for 1989 og trekke disse fra de nye tallene i tabellen. Vi skal imidlertid se bort fra dette i den videre gjennomgang.

Steg 2: Rene sentralgrenser

Sammenstillingen av tabell 2 gir som nevnt likevel ikke det fulle eller sanne bildet av totalletterspørselen i hvert sentralområde: *Av ulike grunner er det slik at enkelte abonnemnt med adresse i en sentrals geografiske dekningsområde (det er dette som kalles sentralområde) er tilknyttet en sentral med annet geografisk dekningsområde. Slike tilfeller kalles nummerlån (eller fremmedtilknyttinger) mellom sentraler⁶.*

Når vi skal finne de korrekte tall for status for totalletterspørselen i hvert sentralområde, må vi derfor "tilbakeføre" alle nummerlån til det sentralområdet som abonnementsadressen faktisk (geo-

grafisk) tilhører. Oversikt over dette finner vi i "nummerlån-statistikken".

Når denne korrigeringen er utført, sier vi at statistikken viser etterspørsel fordelt på sentraler med "rene" sentralområdegrenser. Etterpå kan disse dataene summeres opp til f eks kommune-, teledistrikt- eller regiondata med rene sentralområdegrenser.

Et eksempel på korrigering for nummerlån er gitt i tabell 3.

Legg merke til den store variasjonen i nummerlån mellom de ulike sentralene. Mens Sørkedalen sentral verken låner nummer til eller fra noen andre sentraler, har f eks Centrum 6/7/8/10/FREE/3 utlånt hele 836 nummer (tilknyttinger) til andre sentraler, samtidig som sentralene (dette er flere sentraler som i sum dekker et sentralområde) har lånt 594 nummer fra andre sentraler. Netto utgjør imidlertid ikke dette mer enn -242 abonnemnt, selv om dette er et høyt nok tall på sentralområdenivå.

Eksemplet illustrerer også et annet viktig poeng:

Fra tid til annen foregår det "opprydding" i nummerlån, dvs at abonnemntene blir tilbakekoplet til de sentraler som de geografisk tilhører. I den regulære statistikken fra TELSIS framkommer dette som endringer i tilgang eller etterspørsel for de aktuelle sentraler. Likevel er ikke dette tilfellet: Vi har bare hatt en nummeromkopling. Dette viser at det er viktig å ha oversikt over nummerlånene sentralene imellom og at det er viktig å vite når slike nummeromkoplinger skjer.

Siden vi fokuserer på etterspørselen fra et geografisk avgrenset område er det m a o viktig at historikken viser hvordan denne utviklingen virkelig har vært. Det faktum at Televerket velger å kople enkelte abonnemnt opp mot en annen sentral enn den hvor abonnemntet naturlig (geografisk) hører hjemme, må altså ikke

bidra til at statistikken gir et annet bilde av etterspørselsutviklingen enn den vi faktisk har hatt!

Etterspørselsprognosene er altså prognoser for etterspørsel etter abonnemnt i geografiske områder som er uavhengige av hvor abonnemntene fysisk blir tilkople⁷.

Steg 3: Justering for flytting av sentralgrenser

Et annet viktig moment som må takles, er eventuell *flytting av sentralgrenser*. Når en sentral gir fra seg en del av sitt dekningsområde til en annen sentral, blir numrene/abonnemntene i det aktuelle geografiske området koplet om fra den

Tabell 2 Steg 1: Beregning av total etterspørsel etter TAT i Oslo kommune 31.12.90

Sentralområde i Oslo kommune	Antall innkoplede TAT (R35) (1)	Netto ueff. TAT (R34) (2)	Total ettersp. TAT (1)+(2)
Furuset	14947	17	14964
Grorud	14611	42	14653
Stovner	12857	35	12892
Økern	22066	133	22199
Bryn	18588	34	18622
Holmlia	10489	52	10541
Kastellet	21531	68	21599
Øst	25646	92	25738
Fagerborg 5	187	68	255
Fagerborg	31301	322	31623
Frogner	25533	96	25629
Slemdal	15229	118	15347
Sørkedalen	228	0	228
Tåsen	15307	58	15365
Ullern	17141	27	17168
Åsen	20364	117	20481
Centrum 4/5/9	13765	126	13891
Centrum 6/7/m fl	18131	118	18249
Maridalen	0	0	0
Nord	23333	180	23513
Vaterland	4514	56	4570
Vika	5249	87	5336
Oslo kommune sum	331017	1846	332863

⁶ I mange regioner er ikke nummerlån mellom sentralområdene praktisert. Hvis så er tilfellet, faller dette steget bort.

“givende” sentralen til den “mottakende” sentralen. Da vil statistikkene isolert sett gi et kunstig bilde av etterspørselsutviklingen: På tidspunktet for nummeromkoplingen får vi et hopp ned i statistikken for det givende sentralområdet og et tilsvarende hopp opp i det mottakende sentralområdet.

Ideelt sett burde en i slike tilfeller ha korrekte historiske data over utviklingen i den delen av dekningsområdet som ble flyttet, trekke disse ut av historikken for det givende sentralområdet og legge dem til historikken for det mottakende sentralområdet. I mangel av slike data har vi i praksis to måter å takle dette problemet på:

- 1 Vi kan enten la prognosemodellen estimere nummeromkoplingen som en del av prognosemodellen. Dette kan gjøres ved å modellere nivåskift (f.eks. ved hjelp av såkalte “dummyvariable”) i det tidspunkt hvor nummeromkoplingen faktisk skjedde.
- 2 Eller vi kan, etter beste skjønn, forsøke å rekonstruere den historiske utviklingen etter de nye sentralgrensene (manuell justering). Dette kan gjøres ved at en antar at den abonnementsmassen som omkoples til ny sentral, har hatt den samme utvikling som det totale sentralområdet den kom fra. Da har vi estimer for den historiske utvikling i den delen av dekningsområdet som ble flyttet. Når denne tidsserien så trekkes ut fra det gamle

sentralområdet og legges til det nye, har vi rekonstruert (estimert) historikken for de to sentralområdene etter de nye sentralgrensene. Dette er trolig det nærmeste vi kommer sammenliknbare historiske data for sentralområder som har fått endret sine dekningsområder (sentralgrenser).

Resonnementene ovenfor gjelder selvfølgelig også for det tilfellet at ett eller flere sentralområder avgir abonnement til et helt nytt sentralområde.

I tabell 4 er det vist et eksempel på slike manuelle justeringer for et utvalg av sentraler i Region Oslo.

Følger vi sentralen i den første linjen (Centrum 4/5/9), så gir denne utviklingen slik den framkommer etter at vi har gjennomgått Steg 1–2. Vi ser imidlertid at utviklingen viser et fall i total- etterspørselen fra 30593 TAT i 1984 til 12211 TAT i 1985. Da fikk Centrum 6/7/m fl, som var nyopprettet dette året, ca 20000 abonnement fra Centrum 4/5/9.

I linje 2 er det for årene 1982, 1983 og 1984 estimert hvordan de 20000 TAT som ble flyttet, utviklet seg. Det er antatt at de abonnemene som ble flyttet, har utviklet seg fra henholdsvis 17891 i 1982 til 18319 i 1983 og 19254 i 1984. Dette er under forutsetning av at denne abonnementsmassen har hatt samme utvikling som det totale dekningsområdet for Centrum 4/5/9.

I linje 3 (C4/5/9 ny 1985) er så den antatte utvikling for de flyttede abonnemene trukket ut av linje 1. Dermed står den beregnede utviklingen for Centrum 4/5/9 tilbake, slik vi antar at den har utviklet seg om Centrum 6/7/m fl også hadde eksistert i perioden 1982–1984.

I linje 4 ser vi statistikken for Centrum 6/7/m fl. Den starter naturligvis i 1985, det året som den ble opprettet. Imidlertid har vi beregnet hva utviklingen i årene 1982, 1983 og 1984 var (se linje 2). I linje 5 (C6/7/m fl ny 1985) er disse årene lagt til linje 4, og vi har nå en historikk for Centrum 6/7/m fl for hele perioden 1982–1990.

⁷ Et unntak er det som kalles “forlenget hovedlinje”, men da har abonnenten faktisk etterspurt et abonnement i et annet sentralområde. Statistikkene skal altså ikke korrigeres for slike tilfeller på tilsvarende måte som for nummerlån.

Tabell 3 Steg 2: Beregning av total etterspørsel etter TAT korrigeret for nummerlån i Oslo kommune 31.12.90

Sentralområde i Oslo kommune	Etterspørsel ukorrigeret	Nr. lån til annen sentral (+)	Nr. lån fra annen sentral (-)	Etterspørsel korrigeret
Furuset	14964	1	5	14960
Grorud	14653	17	3	14667
Stovner	12892	1	15	12878
Økern	22199	124	2	22321
Bryn	18622	0	378	18244
Holmlia	10541	379	79	10841
Kastellet	21599	90	4	21685
Øst	25738	60	73	25725
Fagerborg 5	255	0	130	125
Fagerborg	31623	134	112	31645
Frogner	25629	41	22	25648
Slemdal	15347	10	2	15355
Sørkedalen	228	0	0	228
Tåsen	15365	63	15	15413
Ullern	17168	15	3	17180
Åsen	20481	24	69	20436
Centrum 4/5/9	13891	583	577	13897
Centrum 6/7/m fl	18249	594	836	18007
Maridalen	0	247	0	247
Nord	23513	14	125	23402
Vaterland	4570	2	0	4572
Vika	5336	6	0	5342
Oslo kommune sum	332863	2405	2450	332818

Følger vi linje 5, ser vi imidlertid at vi også i 1989 fikk en markant nedgang i totalletterspørselen for dette sentralområdet: Fra 1988 til 1989 gikk totalletterspørselen ned fra 23855 til 17926. Dette skyldes at to nye sentralområder ble opprettet i 1989 og at disse fikk tildelt hver sin abonnementsmasse fra Centrum 6/7/m fl:

Vaterland fikk vel 2000 abonnement fra Centrum 6/7/m fl i 1989. Disse abonnementsutvikling er i perioden 1982 til 1988 antatt å ha vokst fra 1633 til 2177, og dette er satt opp i linje 6. Denne utviklingen er også beregnet ut fra en forutsetning om at abonnementsene har hatt den samme utvikling som det totale sentralområdet for Centrum 6/7/m fl, ref linje 3. Resultatet, etter at abonnementsene i linje 6 er trukket ut av linje 5, finner vi i linje 7 (C6/7/m fl ny 1989, (A)).

Tilsvarende som i tilfellet for Vaterland sentral, fikk Vika sentral vel 4500 abonnement i 1989. Utviklingen for denne abonnementsmassen er i perioden 1982–1988 beregnet på tilsvarende måte som i tilfellet for Vaterland sentral, og den er presentert i linje 8. Resultatet for Centrum 6/7/m fl, etter at den beregnede abonnementsutvikling for perioden 1982–1988 er trukket ut av linje 7, er vist i linje 9 (C6/7/m fl ny 1989, (B)).

Beregningene, slik som de er forklart ovenfor, fortøner seg som det rene puslespill. I praksis er det viktig å ta for seg “generasjon” for “generasjon” i datasettene, ettersom de ulike grensejusteringene forekommer.

Fordelen med denne kontinuerlige “rekonstruering” av data er at vi til enhver tid har data for alle sentralområder i alle perioder. Hadde vi ikke gjort dette, ville vi hatt flere sentralområder med så kort historikk at det ikke ville være mulig å trekke noen brukbare konklusjoner om historisk og framtidig utvikling.

Ulempen med denne form for rekonstruksjon er, som før nevnt, at vi ikke kan være sikre på at den tidsrekken vi kommer fram til virkelig reflekterer den sanne historiske utvikling for de nye og endrede sentralområdene. Mest sannsynlig gjør de ikke det. Likevel er dette muligens den beste løsning vi kan finne, gitt fordelene som er nevnt i avsnittet ovenfor.

Steg 4: Endelig oppsett

Når data er korrigert for nummerlån og justert for grensejusteringer, er vi klare til å lage prognoser: Vi har data på “riktig” etterspørselsform. Vi kan da sette opp de endelige oversikter over den historiske utvikling på sentralområde, teleområde, kommune og region.

Selve prognostiseringen kan enten skje ved at tidsrekken først analyseres med tanke på å bygge en prognosemodell, eller ved at en allerede innarbeidet prognosemodell benyttes.

4.2 Behandling av månedsdata

Steg 1 – Steg 4 ovenfor viser hvordan vi ved hjelp av ulike statistikkrapporter kommer fram til status for total etterspørsel ved utgangen av 1990.

Når det gjelder månedsdata, har det før 1991 ikke vært mulig å gjennomføre en slik oppfølging: Månedssstatistikkene har kun vært sammensatt av R35 og R34, dvs *månedsdata har ikke vært korrigert for nummerlån*. Fra og med januar 1991 har det vært mulig å hente ut nummerlånstatistikkene for hver måned.

Resultatet av denne mangel på månedlig oppfølging av nummerlån-statistikk har vært at Region Oslo tidligere fikk spesialkjøringer to ganger i løpet av året, for å “logge” situasjonen i forhold til sentralområdeprognosene og eventuelt revidere disse.

Når vi kommer opp på teledistrikt- og kommunenivå spiller nummerlån i de fleste tilfeller en mindre rolle, slik at vi i disse tilfellene (med forbehold om de tilfeller der vi vet det er spesielt mange) har valgt å se bort fra disse i den månedlige oppfølging.

På regionsnivå blir problemet med nummerlån enda mindre enn på teledistrikts- og kommunenivå.

5 Segmentering av markedet

Ett av de viktigste målene med etterspørselsstudier er forklaring av etterspørselens historiske utvikling. Som regel starter slikt arbeid med teorier for beskrivelse av hvordan markedet fungerer. Her spiller økonomisk teori en sentral rolle og kan langt på vei hjelpe oss fram til modeller som beskriver sammenhenger mellom årsak og virkning. I økonomiske analyser, som

Tabell 4 Steg 3: Eksempel på beregning av ny historikk etterspørsel etter TAT for noen sentraler som har hatt grensejusteringer i perioden 1982–1990

Sentralområde	1982	1983	1984	1985	1986	1987	1988	1989	1990
1 Centrum 4/5/9	28427	29107	30593	12211	12856	12704	13363	13258	13897
2 C4/5/9–C6/7 m fl 1985	17891	18319	19254						
3 C4/5/9 ny 1985	10536	10788	11339	12211	12856	12704	13363	13258	13897
4 Centrum 6/7/m fl	0	0	0	20160	21225	23625	23855	17926	18007
5 C6/7/m fl ny 1985	17891	18319	19254	20160	21225	23625	23855	17926	18007
6 C6/7 m fl–VAC 1989	1633	1672	1757	1840	1937	2156	2177		
7 C6/7 m fl ny 1989 (A)	16258	16647	17497	18320	19288	21469	21678	17926	18007
8 C6/7 m fl–VIC 1989	3406	3487	3665	3838	4040	4497	4541		
9 C6/7 m fl ny 1989 (B)	12853	13160	13832	14483	15248	16972	17137	17926	18007

spesielt er opptatt av forklaringsfaktorer, er det nettopp slike "forklaringsmodeller" som beregnes. Selve beregningen foregår ved hjelp av statistiske metoder, vanligvis en eller annen variant av regresjon. Modellene blir i neste omgang anvendt til å gi anslag for den framtidige utvikling, dvs prognoser.

5.1 Private abonnenters etterspørsel

La oss vende tilbake til økonomisk teori. I teorien for konsumentenes tilpasning forutsettes det at konsumentene maksimere sin nytte, gitt de priser konsumenten står overfor og den inntekt han/hun har. Av dette kan det avledes at de viktigste forklaringsfaktorer bak etterspørselen etter en gitt vare eller tjeneste (gitt visse forutsetninger om bl a konsumentens nyttefunksjon) er prisen på denne, prisene på alle andre varer og tjenester og den inntekt konsumenten har til disposisjon. Konkret betyr det at en konsument etterspørsel etter en gitt teletjeneste ut fra teorien kan forklares av prisen på denne teletjenesten, prisen på alle andre varer og tjenester og konsumentens inntekter.

I praksis vil prisen på de fleste andre varer og tjenester, f eks ost, melk, brød og tennistimer bli utelatt i en modell for beskrivelse av etterspørselen, fordi etterspørselen etter disse er "uavhengige" av etterspørselen etter teletjenester (dvs de påvirker ikke etterspørselen etter teletjenester). Prisen på f eks posttjenester og ulike andre teletjenester er imidlertid viktige fordi disse tjenestene kan være "konkurrerende i etterspørselen", dvs tjenestene er "alternative goder", ved at kundene enten velger den ene eller den andre tjenesten. Viktig er også prisen på terminalutstyret som skal til for å benytte teletjenesten, fordi dette ofte er varer som er "komplementære i etterspørselen", dvs når teletjenesten blir etterspurt, blir også terminalutstyret etterspurt.

I tillegg kommer momenter som slik generell økonomisk grunnteorier ikke sier noe om, dvs som er spesifikk for etterspørselen etter den gitte teletjenesten. Dette kan være forholdet mellom prisene på anskaffelse og bruk av tjenesten, økt konsumentnytte når antall abonnenter tilknyttet nettet vokser (access externality), ekstra nytte av muligheten til å motta meldinger/informasjon fra andre abonnenter (call externality), ekstra nytte ved at abonnenten kan velge om han/hun skal anvende tjenesten (option demand) m m⁸.

5.2 Bedriftenes etterspørsel

I teorien for bedriftenes tilpasning spiller sammenhengen mellom kostnadene ved produksjon av varer og tjenester og de priser som oppnås i markedet en viktig rolle. Hvis det forutsettes at bedriften tilpasser seg ved å maksimere sitt overskudd, gitt de kostnadene den er stilt overfor, kan (under gitte forhold) bedriftenes etterspørsel etter de innsatsfaktorer som må til i produksjonen av en gitt vare/tjeneste, utledes som funksjoner av den prisen som varen oppnår i markedet og prisen på alle innsatsfaktorene. Bedriftens bruk av teletjenester kan anses som slike innsatsfaktorer i produksjonen.

Ved valg av hvilke priser som i praksis skal innlemmes i en god forklaringsmodell for bedriftens etterspørsel etter en gitt teletjeneste, er det også i dette tilfellet viktig å ta med de som vi antar påvirker denne etterspørselen: Prisen på innsatsfaktorer som er "teknisk alternative i produksjonen", f eks posttjenester og ulike andre teletjenester er også her relevante. Likeså vil også f eks prisen på terminalutstyret som skal til for å benytte teletjenesten være viktig, fordi dette er en innsatsfaktor som er "teknisk komplementær i produksjonen". Innsatsfaktorer som er "uavhengige i produksjonen" utelates som regel.

Momenter som slik generell økonomisk grunnteorier ikke sier noe om, kan det også være aktuelt å ta med: For eksempel er det vanlig å anta at bedriftens inntekter/overskudd har en innvirkning på hvor kostnadsbevisst bedriften er: Perioder med store overskudd bidrar til økt etterspørsel, mens perioder med små overskudd eller underskudd bidrar til lavere etterspørsel. I den senere tid er imidlertid dette forholdet blitt komplisert ved at telekommunikasjon i voksende grad blir sett i sammenheng med rasjonalisering av driften. Dette kan i så fall bidra til økende etterspørsel også i perioder hvor overskuddet i bedriftene avtar. Også det generelle aktivitetsnivået i bedriften (f eks målt ved omsetningen) er vanligvis antatt å påvirke etterspørselen: Voksende aktivitetsnivå bidrar til økt behov for teletjenester.

Endelig gjelder også her at forhold som måtte være spesifikke for etterspørselen etter den gitte teletjenesten må ivaretas på samme måte som for konsumentene⁹.

5.3 Segmentprognoser

En viktig lærdom av teoriene for konsumentenes og bedriftenes tilpasning er

at det til dels er ulike forklaringsfaktorer bak etterspørselen etter varer og tjenester fra de to. I tillegg forventer vi at konsumentene og bedriftene i sin etterspørsel legger ulik vekt på de forklaringsfaktorer som måtte være felles. Skal vi lage gode forklaringsmodeller for etterspørsel etter varer og tjenester generelt, og teletjenester spesielt, bør vi derfor segmentere markedet og behandle hvert segment for seg. Intuitivt vil dette gi oss en gevinst i form av økt viten om markedet som igjen vil bidra til sikrere prognoser.

I prinsippet kan det også innenfor hvert av de to segmentene være grupper av kunder hvor etterspørselen kan forklares ut fra faktorer som ikke gjelder for andre grupper eller hvor vektlegging av forklaringsfaktorene er forskjellig fra gjennomsnittet av segmentet. Hensiktsmessighet, mål for de ulike analysene og det faktum at prognoser generelt blir mer usikre jo mindre grupper de gjelder, begrenser hvor langt en i det enkelte tilfellet bør gå i retning av stadig mindre segmenter.

Den enkleste form for segmentering av abonnementsmarkedet er nettopp privat- og bedriftsmarkedet. I forbindelse med omorganiseringen av Televerket har denne segmenteringen fått fornyet aktualitet. De nye markedsdivisjonene har behov for klart å definere hvilke abonnement som tilhører de ulike divisjonene. I den forbindelse ble det utarbeidet nye definisjoner for de såkalte brukskategorier (bkat) for telefonabonnement i TELSIS/AB, og disse ble fordelt til Bedriftsmarkedsdivisjonen (BMD) og Privatmarkedsdivisjonen (PMD) på følgende måte:

- 1 *Boligabonnement* – bkat B (PMD): Abonnement kun eid og i bruk av privatperson
- 2 *Forretningsabonnement* – bkat F (BMD): Abonnement som entydig er eller er lokalisert til bedrifts-virksomhet

⁸ Uttrykkene i parentes er de som anvendes i engelsk/amerikansk litteratur om teorier for etterspørsel etter telefonabonnement og bruk av disse. Se f eks [6] for en sammenfatning av disse momentene.

⁹ Det finnes lite litteratur som beskriver særegenheter ved bedriftenes etterspørsel etter teletjenester for bruk i kvantitative prognosemodeller. Dette skyldes i stor grad at bedriftenes tilpasning er komplekse.

- 3 *Forretnings-/boligabonnement* – bkat FB (BMD): Abonnement som opprettes for selvstendig næringsdrivende som driver næringsvirksomhet fra egen bolig
- 4 *Bolig-/bedriftsabonnement* – bkat BB (PMD): Abonnement som er eid av en bedrift og installert hos en privatperson som arbeider i vedkommende bedrift
- 5 *Hytteabonnement* – bkat H (PMD): Abonnement til hytte eller sommerbolig
- 6 *Televerkets samtaleautomater og fri-luftskiosker* – bkat SM (PMD)
- 7 *Televerkets abonnement* – bkat S (BMD): Samlekategori for abonnement til Televerkets telefoner.

Ulike abonnementspriser påvirker inntektsbudsjettene, og ulike utstyrpriser påvirker investeringsbudsjettene. Dette er i neste omgang med på å påvirke budsjettene detaljeringsgrad, og disse vil langt på vei sette rammer for hvor langt ned det er behov for å segmentere markedet. Per i dag er det naturligvis klare behov for å splitte f.eks. abonnementsprognosene på PMD og BMD pga inntektsbudsjettene. I tillegg er det behov for å splitte BMDs prognoser i regulære analoge abonnement, 2 Mbit/s-abonnement og antall innvalgsnummer, fordi ulike priser påvirker både inntekts- og investeringsbudsjettene.

I framtiden vil det bli økende behov for segmentering. Dette skyldes bl.a. det faktum at det stadig blir teknisk mulig å skreddersy abonnement til ulike brukergrupper og at vi får konkurranse på stadig flere områder. Førstnevnte krever større detaljer i investeringsanalyser/budsjetter, mens sistnevnte på tilsvarende måte gir behov for mer fininddelte inntektsbudsjetter. Dette tvinger fram segmentering både for kunde grupper og for produktgrupper.

Dessverre lar det seg ikke gjøre å ta ut offisielle tidsserier som går særlig langt tilbake for abonnement/tilknyttinger fordelt på brukskategoriene nevnt over. Fra og med årsskiftet 1990/91 ble det imidlertid tilrettelagt muligheter for å splitte R35-tallene i AS, slik at vi får tilgangstallene fordelt på brukskategori. Senere er også tilsvarende rapporter for de ueffektuerte og ikke oppdaterte ordrer kommet til, slik at etterspørselen nå lar seg fordele på brukskategori.

I Region Oslo er det konstruert en grov fordeling av data for årlig netto-etterspør-

sel tilbake til 1977 for intern bruk. Denne konstruerte fordelingen er presentert i figur 6.

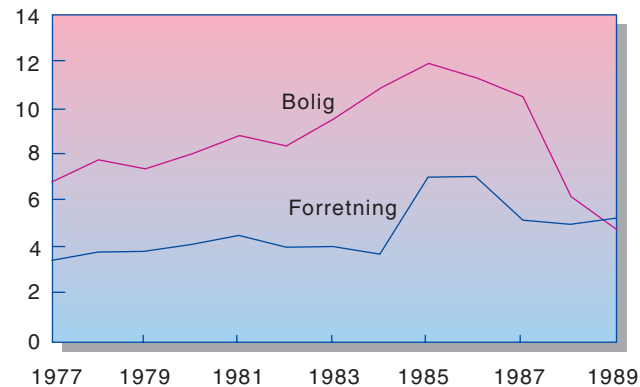
Selv om denne fordelingen er påvirket av de anvendte metoder for fordeling av abonnementene (spesielt antas det at dette er tilfellet i periodene før 1984), illustrerer figuren et viktig faktum: I enkelte perioder utvikler etterspørselen etter abonnement fra forretningskundene seg totalt forskjellig fra tilsvarende etterspørsel fra privatkunder. Vi forventer derfor at segmentanalyser/prognoser vil gi en gevinst i form av bedre viten om markeds-mekanismer og sikrere prognoser.

En tilsvarende beregnet fordeling av HA i brukskategorier for alle teleområder er årlig utført av det tidligere utredningskontoret i Teledirektoratets økonomiavdeling, ØSU, se [2].

Litteratur

- 1 Makridakis, S, Wheelwright, S C, McGee, V E. *Forecasting: Methods and Applications, 2nd edition*. New York, John Wiley, 1983.
- 2 ØSU. *Beregning av etterspørselen etter boligabonnement 1991*. Oslo, Økonomiavdelingen, Teledirektoratet, 1992.
- 3 Bøe, J, Hjelkrem, C. Utarbeiding av prognoser for etterspørsel etter telefonabonnement i Oslo teledistrikt. *Teletronikk*, 87(1), 93–109, 1991.
- 4 Taylor, L D. *Telecommunications demand: a survey and critique*. Cambridge, Mass., Ballinger, 1980.
- 5 Taylor, L D. The demand for telecommunications: a nontechnical exposition. I: Michael A Crew, *Issues in public utility pricing and regulations*. Lexington Books, 1980.
- 6 Taylor, L D. Problems and issues in modelling telecommunications demand. I: Courville, de Fontenay and Dobell (eds.), *Economic analysis of telecommunications: theory and applications*. North-Holland, 1983.
- 7 Serck-Hanssen, J. *Teorier for konsumentenes adferd*. Oslo, Universitetsforlaget, 1975.

Hovedabonnement (x 1000)



Figur 6 Årlig etterspørsel, HA, fordelt på bolig og forretning. Beregnede tall

- 8 Frisch, R. *Innledning til produksjonsteorien, 9. utgave*. Oslo, Universitetsforlaget, 1971.

Glattingsmodeller

AV JOHANNES BØE OG CARLO HJELKREM

001.18:621.39

1 Innledning

En effektiv prognoseteknikk må inneholde et teoretisk fundament. Det må være mulig å beskrive variasjonen i de historiske data i tillegg til en parameterestimering. Glattingsmodeller inneholder slikt fundament og er beskrevet i denne artikkelen.

- Disse teknikkene benyttes ofte innen salg, kvalitetskontroll og markedsplanlegging.
- Metodikken er basert på beskrivelse av fortiden til bruk i framtiden.
- Modellene er enkle å oppdatere og krever liten kapasitet (tid).
- Modellverket er svært enkelt med hensyn på statistisk analyse, men inneholder likevel en stor praktisk verdi.
- Metodene er praktiske når mange tidsreiser skal behandles samtidig.
- Metodene er intuitivt lette å forstå.

Tidsrekkeanalyse er betegnelsen på metoder som prognostiserer framtidige verdier av en variabel utelukkende på grunnlag av dens historiske verdier. På forskjellige måter identifiseres mønstre i historiske data og forlenger dem inn i framtiden. Metodene innedeles etter hvilke typer tidsrekker som de kan anvendes på. De forskjellige typene er:

- Stasjonære (uten trend eller sesong)
- Med trend (uten sesong)
- Med trend og sesongvariasjon (med sesong menes et regelmessig mønster som gjentas fast hvert år).

Vi skal her beskrive metoder som kan brukes på tidsserier som de tre ovennevnte. Disse metodene vil bli forklart ved hjelp av teori og eksempler hentet fra Televerket.

I svært mange tidsrekker vil det være flere faktorer som påvirker selve tidsrekken. Det kan være sesongbetonte svingninger, tilfeldige variasjoner og irregulære (spesielle) hendelser. For å få en oppfatning av hvorledes trenden i selve tidsrekken er, kan det være hensiktsmessig å benytte glattingsmodeller/glattingsmetoder som eliminerer slik faktorer. Slike modeller er basert på å konsentrere informasjon fra flere observasjoner i ett observasjonspunkt på en slik måte at selve forløpet av tidsrekken glattes ut.

1.1 Når er tidsrekke-analyse velegnet?

Ved første øyekast kan det synes som om slike metoder, som kun ser på data i tidsrekken og ikke annen utenforliggende informasjon, som for eksempel i kvalitative metoder eller modeller med forklaringsfaktorer, ikke er egnet som hjelpemiddel i prognosearbeidet. Endrer forutsetningene seg kan tidsrekkemodellene, som utelukkende bygger på historiske data, verken forutsi endringer eller reagere raskt på disse. Likevel finnes det avgrensede områder hvor de både har gitt gode resultater og hevdet seg godt i tester mot andre metoder. Dette viser seg når samme tidsrekke blir testet opp mot andre prognosemetoder, blant annet regresjon, glattingsmodeller, Kalmanfilteret og ARIMA.

Under stabile forhold, der forholdene ikke endrer seg, er det rimelig å anta at de faktorer som tidligere forårsaket bevegelser i variabelen også vil fortsette å gjøre det i framtiden. Tidsrekkeanalyse vil derfor kunne gi gode prognoser under forutsetning av konstante ytre påvirkninger i framtiden som i fortiden.

Tidsrekkeanalyse er spesielt velegnet til å lage kortsiktige prognoser (0,5 – 2 år avhengig av produktets stabilitet). Over kortere perioder er muligheten for endring i ytre påvirkninger liten. De fleste tidsserier har på kort sikt konstante ytre påvirkninger.

Tidsrekkeanalyse er et ypperlig hjelpemiddel når tiden man har til rådighet er kritisk. Den er lett å implementere, bruker liten maskinkapasitet (CPU-tid) og kan lett lages i tradisjonell programvare, ferdigvare eller på regneark. Analysen er også et godt utgangspunkt før man begynner et stort analysearbeid med data. En slik analyse kan også kalles en "første" prognose. Denne "første" prognosen vil kunne vise hva som kan forventes av resultater.

Tidsrekkeanalyse er nesten uvurderlig i forbindelse med vurdering av data. Metoden finner raskt mønstre i historiske data. Slike mønstre kan gi en bedre forståelse av tidligere bevegelser i data og eventuelt finne såkalte "outliers" i data.

Når vi nå skal presentere metodene vil vi starte med de mer enkle tilfellene, de stasjonære, og til slutt ende opp med de mer kompliserte, tidsserier med trend og sesong. Vi vil starte med basisforutsetningene for vanlig glattning og eksponentiell glattning og etter hvert utvide disse

til Holts og Holt-Winters metode (trend og trend med sesong).

1.1.1 Basis forutsetninger

La oss anta at vi har n observasjoner av en tidsrekke y :

$$y_1, y_2, y_3, \dots, y_n$$

La forventningen for tidsrekken ved tidspunkt t være:

$$E y_t = \mu_t$$

La oss forutsette videre at tidsrekken består av to komponenter μ_t og ε_t slik at:

$$y_t = \mu_t + \varepsilon_t$$

ε_t er her et støyledd som genererer tilfeldige variasjoner. Det forutsettes at:

$$E \varepsilon_t = 0 \quad (\text{gjennomsnitt etter uendelig målinger er } 0)$$

$$\text{Var } \varepsilon_t = s^2 \quad (\text{variasjonen er gitt ved kvadratet av } s) \text{ er konstant.}$$

Dette betyr at verdien til y_t genereres av forventningsverdien μ_t og støyleddet ε_t .

Støyleddet ε_t vil vekse mellom positive og negative verdier, og dersom disse legges sammen over et tilstrekkelig antall målinger vil snittet være lik 0.

2 Glidende gjennomsnitt

Oftest refereres glidende gjennomsnitt/glattning til en prosedyre hvor man veier data for å glatte disse ut og fjerne eventuelle uregelmessigheter. Ved å bruke glattning vil man kunne få eventuell informasjon om trender, sesongmønstre eller sykler i dataene. Dette er viktig for å kunne starte på utvikling og valg av en prognosemodell.

Glidende gjennomsnitt vil danne en logisk start for presentasjonen av glattingsmodellene, som egentlig er mer eller mindre sofistikerte utvidelser av den enkle metoden "glidende gjennomsnitt".

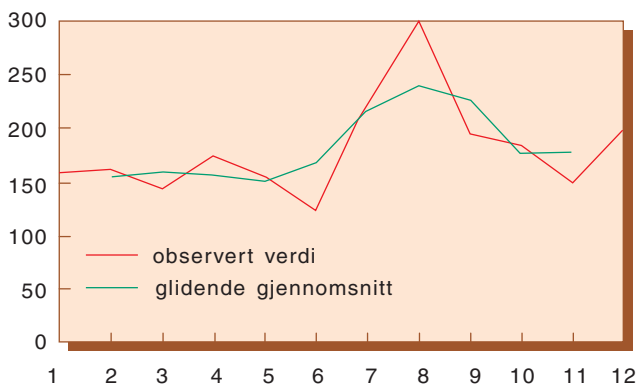
En tidsrekke er stasjonær om den varierer rundt en konstant verdi. Selv om denne konstante verdien vil variere noe over tid vil ikke denne variasjonen kunne tolkes som verken trend eller sesong. Vi kan si:

$$\text{En stasjonær tidsserie har konstant variasjon og inneholder verken trend eller sesong.}$$

I praksis vil ingen tidsserier være stasjonære over en lang periode, men det er mulig over en kort periode. De to metod-

Tabell 1 En tilfeldig valgt rekke hvor trepunkts glidende gjennomsnitt er beskrevet

Periode (t)	Observert verdi	Trepunkts glidende gjennomsnitt	Utrekning
1	159		
2	162	155,00	$(159+162+144)/3$
3	144	160,00	$(162+144+174)/3$
4	174	157,67	$(144+174+155)/3$
5	155	151,00	$(174+155+124)/3$
6	124	168,00	$(155+124+225)/3$
7	225	216,33	$(124+225+300)/3$
8	300	240,00	$(225+300+195)/3$
9	195	226,33	$(300+195+184)/3$
10	184	176,33	$(195+184+150)/3$
11	150	178,00	$(184+150+200)/3$
12	200		



Figur 1a) Tidsserien i tabell 1 glattet med 3 punkter

ene, glidende gjennomsnitt og eksponentiell glatting, som skal gjennomgås, behandler stasjonære tidsserier. Disse er tidligere gjennomgått i annen artikkel, men vi vil likevel gjøre en kort repetisjon her fordi de danner grunnlag for mer sofistikerte glattingsmodeller.

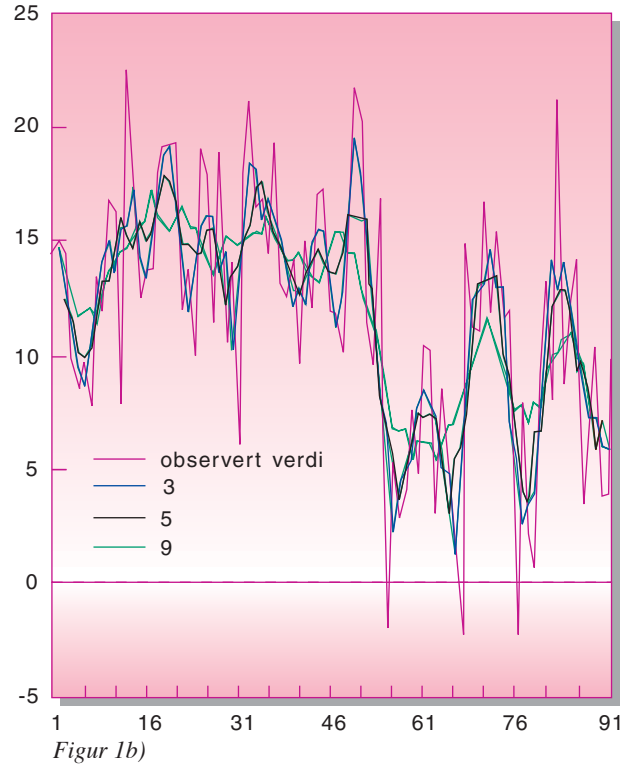
Den glidende gjennomsnittsmo- del tar gjennomsnittet av flere på hverandre følgende observasjoner. Dermed vil effekten av de tilfeldige variasjonene, det vil si støyleddene, være mindre synlig. Dette medfører at trenden i tidsrekken kommer klarere fram.

Den opprinnelige tidsrekken blir erstattet av en glattet rekke som er laget ved å erstatte hver observasjon med gjennomsnittet av den og observasjoner på hver side av den. Hvis hvert gjennomsnitt er beregnet på grunnlag av tre observa- sjoner, beregnes den første glattede

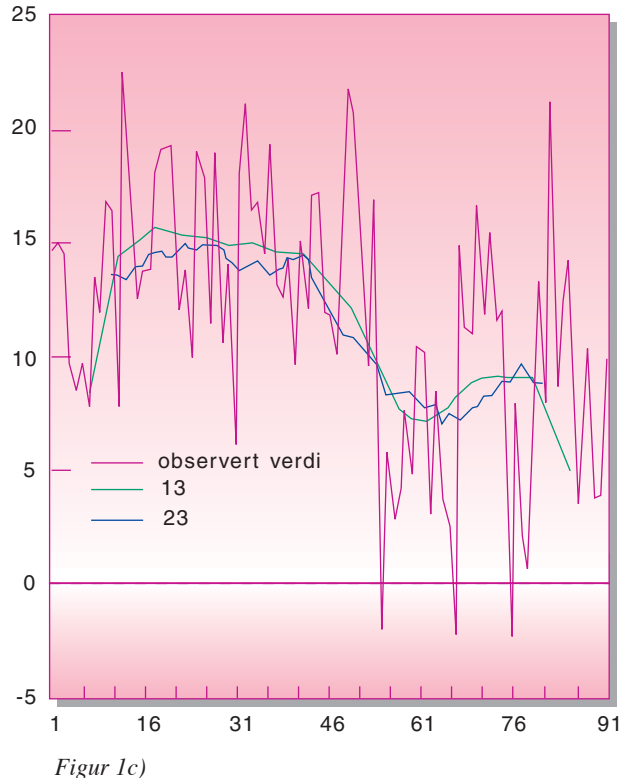
verdien på bakgrunn av verdi nr 1, 2 og 3. Tabell 1 viser en tilfeldig valgt rekke som varierer tilfeldig rundt verdien 180, samt glattinger med 3 punkter. Figur 1a) viser tidsserien i tabell 1 glattet med 3 punkter. Vi vil se av den grafiske framstillingen at de tilfeldige avvikene forsvinner etter hvert som antall punkter øker. Tabell 1 viser i detalj hvordan dette utregnes. Det må advares mot å benytte for stort antall punkter det glattes over, da man kan fjerne viktig informasjon i tids- serien. Det vil alltid være en mulighet at data inneholder f.eks trendskifte som kan bli glattet vekk.

Trepunkts glidende gjennomsnitt for periode t blir følgende:

(x 100)



(x 100)



Tabell 2 De opprinnelige sesongobservasjonene sammen med de glidende gjennomsnittsverdiene

	Observerte verdier	Glidende gjennomsnitt		Observerte verdier	Glidende gjennomsnitt		Observerte verdier	Glidende gjennomsnitt
1984			1987			1990		
1	1466		1	1937	1526	1	1541	945
2	1505		2	1316	1520	2	1167	898
3	1454		3	1260	1487	3	1201	914
4	1002		4	1423	1450	4	-224	889
5	852		5	966	1408	5	792	927
6	972	1223	6	1505	1371	6	218	900
7	780	1289	7	1214	1348	7	63	875
8	1351	1303	8	1711	1420	8	920	896
9	1198	1287	9	1731	1485	9	1329	826
10	1682	1318	10	1194	1461	10	800	908
11	1634	1363	11	1184	1460	11	2126	928
12	785	1432	12	1012	1475	12	871	941
1985			1988			1991		
1	2254	1527	1	1665	1358	1	1233	969
2	1675	1575	2	2180	1264	2	1423	975
3	1259	1637	3	2036	1143	3	356	
4	1376	1597	4	1132	1079	4	761	
5	1385	1576	5	959	1044	5	1035	
6	1807	1594	6	1687	999	6	380	
7	1919	1566	7	-197	948	7	394	
8	1926	1575	8	581	851	8	987	
9	1936	1566	9	287	708	9		
10	1208	1609	10	418	684	10		
11	1383	1582	11	762	636	11		
12	1001	1549	12	483	514	12		
1986			1989					
1	1912	1440	1	1047	512			
2	1786	1429	2	1021	587			
3	1149	1444	3	310	658			
4	1898	1481	4	847	715			
5	1060	1506	5	381	790			
6	1407	1543	6	225	849			
7	615	1545	7	-223	891			
8	1790	1506	8	1490	903			
9	2117	1515	9	1134	977			
10	1648	1476	10	1104	888			
11	1685	1468	11	1664	922			
12	1452	1476	12	1193	921			

(observert verdi periode $t-1$
 + observert verdi periode t
 + observert verdi periode $t+1$)/3.

Beskrivet på en mer matematisk form:

$$S_t = (y_{t-1} + y_t + y_{t+1})/3$$

Et annet eksempel er når vi har sesongdata (det vil si månedsdata eller kvartalsdata, ...). Ved å glatte med antall punkter lik med antall sesongobservasjoner vil hele sesongmønsteret falle bort.

Figur 1b) og 1c) viser en tidsserie som er utvidet i forhold til tabell 2 glattet med henholdsvis 3, 5, 9, 13, 23 punkter. Vi ser

på den grafiske framstillingen at de tilfeldige avvikene forsvinner etter hvert som antall punkter øker.

Tabell 2 viser månedstall for netto etter-spørsmål etter hovedabonnement for telefon fra 1984 til midten av 1991. Det vises der hvordan kurven retter seg ut og blir mer leselig når data glattes. Dette er også vist grafisk. Fordi vi vet at data inneholder sesong kan vi på den måten si at sesongeffekten er fjernet fra data.

Antall perioder som tas med i beregninger av det bevegelige gjennomsnittet er som nevnt vanligvis lik antall sesonger i tidsrekken. I rekker uten sesong bør ikke

antall glattingspunkter bli for stort, da siste glattede punkt vil ligge for langt fra siste observerte verdi. I praksis er da tre til fem punkter mest vanlig.

Når det er snakk om prognosemodell er glidende gjennomsnitt best på tidsrekker som beveger seg rundt en fast verdi og prognosen vil hele tiden være siste beregnede gjennomsnitt. Dette anses som en dårlig prognosemodell. Derimot er glidende gjennomsnitt/ glattede verdier en meget god måte å presentere styringsdata på.

Vi har til nå snakket kun om glidende gjennomsnitt der alle observasjonene i en

snittobservasjon veier likt. Videre kan det være interessant å kunne veie siste observasjon mer enn de tidligere. Vi beveger oss dermed mot en mer intelligent form for tidsrekkemodeller som veier observasjonene forskjellig avhengig av alderen på disse.

3 Eksponentiell glatting

Eksponentiell glatting (og veiet glidende gjennomsnitt), ref innledningsartikkel, er forholdsvis like metoder. De kan være helt ekvivalente dersom parametrene bestemmes på en bestemt måte. Metoden baseres på å veie observasjonene avhengig av avstanden fra siste observasjon. Dette fordi det er rimelig å anta at siste observasjonen har størst betydning og dermed størst vekt.

Den glattede tidsserien er gitt ved likningen

$$\begin{aligned} & \text{ny glattet verdi} \\ &= a(\text{tidligere glattet verdi}) \\ &+ (1-a)(\text{sist observert verdi}) \end{aligned}$$

hvor a er et tall mellom 0 og 1.

Vi må tenke oss å være i tiden $t = 1$ (altså første observasjon), den eneste kjente verdien vi har er y_1 . Så lenge vi bare vet en verdi kan vi bare anslå neste observasjon ut fra den første (y_1), vi kan da si at estimatet y'_2 for y_2 (observerte verdi ved tiden $t = 2$) er det samme som for første verdi $y_1 = y'_2$.

Når vi senere observerer verdi y_2 , blir spørsmålet :

Hvor mye vekt skal siste observerte verdi ha i forhold til de tidligere observerte verdier?

Generelt kan vi si:

y'_t er altså et estimat for observasjon i tiden t , y'_t er beregnet på bakgrunn av tidligere observasjoner.

De tidligere observerte verdier er lagt inn i estimatet for y_2 . Nå er y_2 observert, og vi fortsetter med å estimere y_3 på bakgrunn av y_1 og y_2 , observerer så y_3 og veker gammel informasjon og ny informasjon og estimerer neste.

Uttrykt matematisk blir dette:

$$y'_t = ay'_{t-1} + (1-a)y_t$$

Dette kalles en rekursiv funksjon og alle rekursive funksjoner må "sparkes" i gang. Hva er estimatet for y_1 ved tiden $t = 1$ når y'_0 ikke finnes? Vi har da ingen informasjon og vi trenger en "startverdi".

Denne startverdien kan enten gjettes på bakgrunn av informasjon om tids-serien, det kan brukes backcasting (prognose på omvendt tidsserie) eller den kan settes fritt. Dette kan avhenge av antall data som er til rådighet. Et godt alternativ til backcasting er som nevnt over å bruke den første verdien som startverdi.

Parameteren a er mellom 0 og 1, denne bestemmes ut fra hvilken a som passer best avhengig av hvor stor følsomhet denne har på hver enkel tidsserie. En høy verdi på a vil gi en modell som føyer seg meget godt inn til tidsserien uten at dette nødvendigvis vil gi gode prognoser.

Et eksempel her kan være månedlige netto etterspørselstall for telefon, se tabell 2, og glatte disse hvor glattingsparameter a varierer ($a = 0,9$, $a = 0,5$). Figur 3 viser hvordan den glattede kurven endrer seg med variasjon i glattingsparameteren a .

Vi ser av de observerte data at disse varierer voldsomt mellom verdiene 500 til 2500. Videre kan vi se at ved observasjon nr 55 endres nivået på kurven. Den glattede kurven med glattingsparameter $a = 0,5$ reagerer raskere enn kurven med glattingsparameter $a = 0,975$. Dette kan forklares ut fra modell-likningen. Med en stor a legges det mye vekt på tidligere observasjoner og liten vekt på den siste, og man kan si at stor a medfører treghet i glattingsmodellen. Ved liten a reagerer den glattede kurven raskt på nye endringer.

Vi skal nå vise hvorfor metoden kalles eksponentiell glatting. Med litt algebraisk manipulering vises dette raskt. Likningen for eksponentiell glatting er følgende:

$$y'_t = ay'_{t-1} + (1-a)y_t \quad (1)$$

Likningen for forrige periode var:

$$y'_{t-1} = ay'_{t-2} + (1-a)y_{t-1} \quad (2)$$

Settes likning (2) inn i likning (1) får vi:

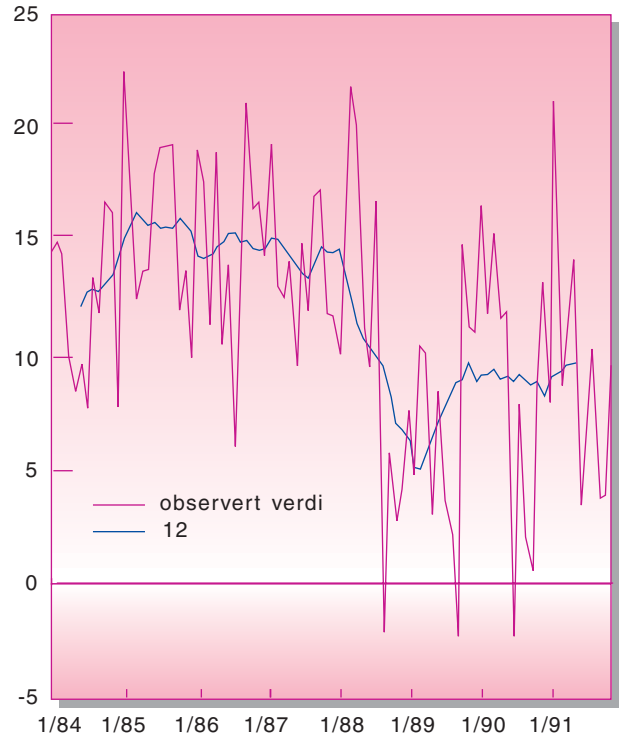
$$y'_t = aay'_{t-2} + (1-a)ay_{t-1} + (1-a)y_t$$

eller

$$y'_t = (1-a)y_t + (1-a)ay_{t-1} + aay'_{t-2}$$

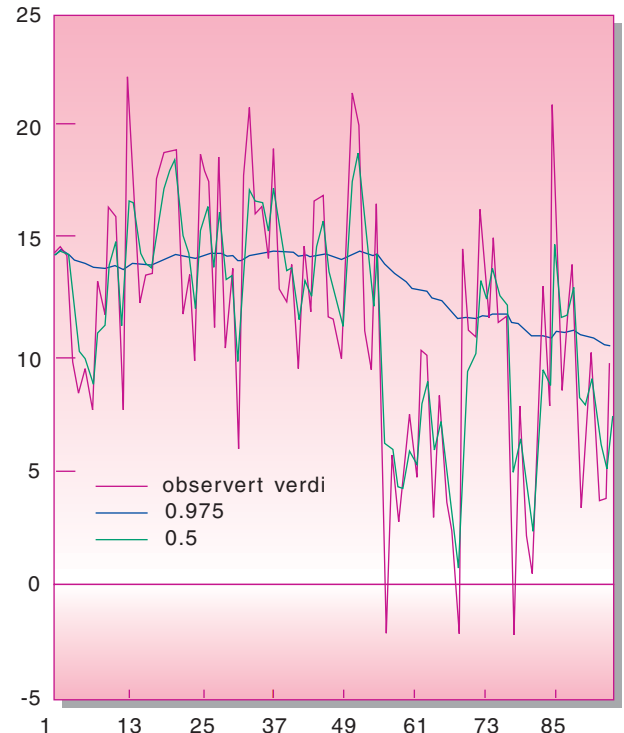
Fortsetter vi på denne måten får vi til slutt:

(x 100)



Figur 2 Grafisk framstilling av tabell 2 i form av rådata og 12 måneders glidende gjennomsnitt

(x 100)



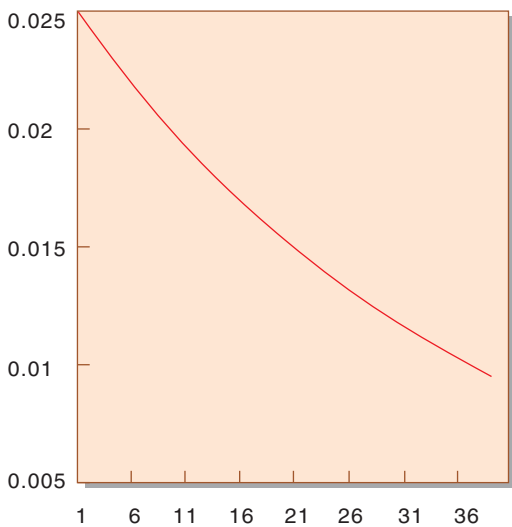
Figur 3 Månedlig netto etterspørsel etter telefon og glattede kurver

$$y'_t = (1-a)y_t + (1-a)ay_{t-1} + (1-a)aa_{t-2} + \dots$$

Vektene som tildeles tidligere blir følgende:

$$(1-a), (1-a)a, (1-a)aa, \dots$$

Figur 4 viser hvordan vektene avtar eksponentielt med avstanden fra siste observasjon. Dette kan anses som en rimelig antakelse ettersom det er mest



Figur 4

naturlig at de siste observasjonene har største betydning.

3.1 Prognosemodell, eksponentiell glatting

Prognosemodellen basert på eksponentiell glatting kan uttrykkes enkelt på følgende matematisk form:

$$F_{t+m} = y'_t$$

Ut fra denne formelen ser vi at prognosen er en fast verdi som vil endre seg når nye observasjoner kommer til. y'_t finnes ved hjelp av rekursjonsformelen som er beskrevet tidligere.

3.1.2 Eksempel

Vi tar utgangspunkt i to forskjellige tidsrekk hvor én er med trend og én uten trend (stasjonær). På begge tidsrekkene bruker vi glattingskonstant $a = 0,75$.

Figur 6 viser historikk, tilpasning og prognose for tidsrekke med trend (ikke stasjonær).

Intuitivt vil vi forstå at når vi har en tidsrekke som er voksende vil prognosen få en konstant verdi. Ser vi ut fra de historiske data at tidsrekken er voksende er det naturlig å anta at kurvens observasjoner

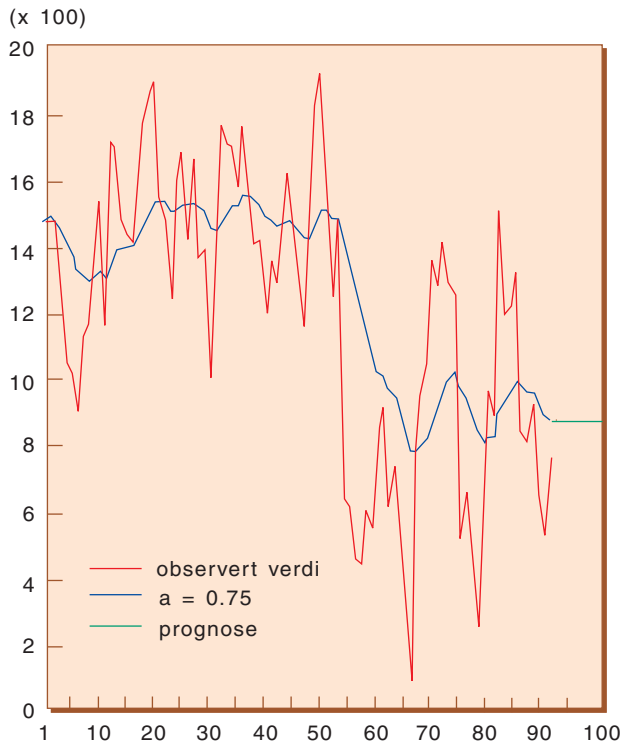
vil vokse også i framtiden. Konklusjonen blir altså at en prognosemodell basert på eksponentiell glatting ikke vil være aktuell å bruke på andre tidsrekk enn på stasjonære tidsrekk. Figur 6 viser dette.

Figur 5 viser en tidsrekke som er stasjonær, dog med et nivåskift. Her passer modell og tidsrekke sammen, modellen estimerer kun ett nivå og tidsrekken inneholder kun denne komponenten.

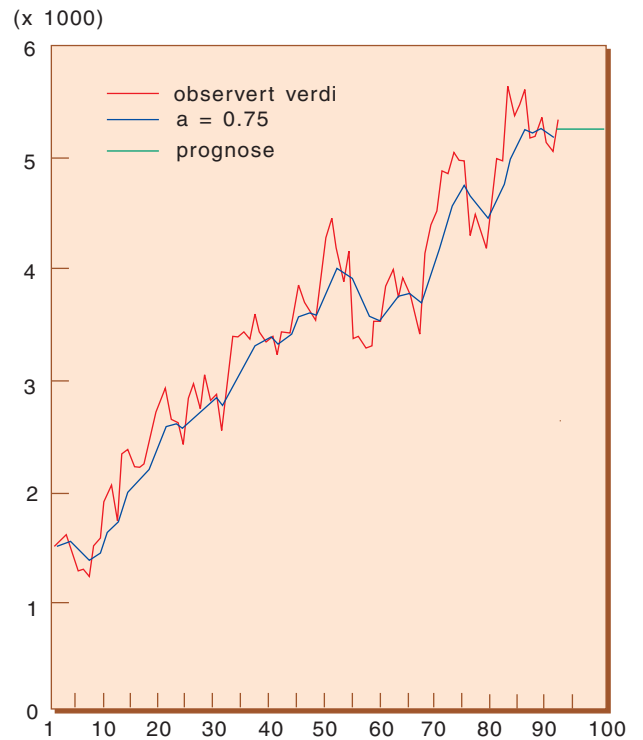
Disse eksemplene viser hvor viktig det er å vite noe om tidsrekken før en lager en prognose.

4 Tidsrekker med trend, Holts metode

Bruk av bevegelig gjennomsnitt eller eksponentiell glatting kan, som nevnt i kapittel 3, avdekke en eventuell trend. Eller trenden kan være åpenbar fordi vi allerede har kjennskap til hvordan tidsrekken har oppført seg tidligere. For å kunne håndtere en tidsrekke med trend finnes flere andre modeller utledet fra eksponentiell glatting. Vi velger å se på en metode som kalles Holts metode som er en form for eksponentiell glatting, men har i tillegg tatt med en trendfaktor.



Figur 5 Historikk og prognose for tidsrekke uten trend (stasjonær)



Figur 6 Historikk, tilpasning og prognose for tidsrekke med trend (ikke stasjonær)

Prinsippet for eksponentiell glatting gikk ut på å estimere μ'_t som den underliggende forventede verdi på nivået ved tidspunkt t . La som i eksponentiell glatting tidsseriens forventningsverdi ved tidspunkt t være μ_t og rekursjonslikningen blir da:

$$\mu'_t = (1 - a) y_t + a\mu'_{t-1}$$

Her gis siste observasjon y_t vekt $(1-a)$ og det tidligere beregnete nivået μ'_{t-1} gis vekten a .

Hvis tidsrekken inneholder en trend, vil den glattede verdien y'_t (som prognose for framtidige perioder) generelt være for lav eller for høy avhengig av om trenden er positiv eller negativ, se eksempel i avsnitt 3.1.2. Dette fordi funksjonen kun tar hensyn til verdiene som er observert tidligere og ikke til endringen mellom verdiene. Derfor estimeres ikke en trend for framtidige verdier. Hvis det er en trend, bør den komme til syne i de glattede verdiene.

Figur 7 viser data med trend som er glattet ved hjelp av eksponentiell glatting. Det vises der tydelig at kurven inneholder en trend.

I Holts metode føres det inn en ekstra parameter som beskriver trenden i tidsserien. Trenden beskrives ved hjelp av parameteren β_t . β_t vil være stigningskoeffisienten til tidsserien ved tidspunkt t . Dermed vil vi kunne beskrive en tidsserie med trend på følgende måte:

$$y_t = y_{t-1} + \beta_{t-1} \quad y_{t-1} = y_{t-2} + \beta_{t-2}$$

Vårt mål blir nå å finne en hensiktsmessig måte å lage et godt estimat for y_t , μ'_t . Dette må gjøres slik at vi også tar hensyn til stigningskoeffisienten β_t . Vi tar da utgangspunkt i y_t som er de observerte verdiene. Under eksponentiell glatting kom vi fram til funksjonen

$$\mu'_t = (1 - a) y_t + a\mu'_{t-1}$$

og det er nå nødvendig å legge til en faktor som beskriver trenden i tidsserien. Uten å forklare for meget vil vi bare sette opp funksjonsformelen og la denne tale for seg. Intuitivt vil vi forstå at tidligere estimert verdi for nivået må økes med stigningen som tidsserien inneholder og vi får dermed funksjonen:

$$\mu'_t = (1 - a) y_t + a(\mu'_{t-1} + \beta'_{t-1})$$

Spørsmålet er nå følgende:

Hvordan beregne et estimat for trendfaktoren til formelen?

Den observerte trendfaktoren β_t kan bestemmes ved å ta differansen mellom y_t og y_{t-1} .

$$\beta_t = y_t - y_{t-1}$$

Brukes dette estimatet for trenden i tidspunkt t vil vi få et dårlig estimat hvis det er store variasjonen i trenden. Vi velger heller å bruke endringen i de estimerte nivåene i stedet. Altså bytter vi ut y_t med μ'_t og får følgende uttrykk:

$$\beta_t = \mu'_t - \mu'_{t-1}$$

Det må her gjøre klart at β_t er å betrakte som en verdi på siste observerte trend, og ikke den estimerte trendfaktoren β'_t .

Vi har nå "observert verdi på trenden" og kan sette sammen en funksjon basert på eksponentiell glatting.

$$\beta'_t = (1-b)\beta_t + b\beta'_{t-1}$$

$$\beta'_t = (1-b)(\mu'_t - \mu'_{t-1}) + b\beta'_{t-1}$$

b er her en parameter, slik som a , som skal ligge mellom 0 og 1 og som angir hvor mye siste observasjon for trenden skal veie i forhold til de tidligere.

4.1 Oppsummering

Når en tidsrekke inneholder en trend beregnes glattingen basert på Holts metode med følgende to formler:

$$\mu'_t = (1 - a) y_t + a(\mu'_{t-1} + \beta'_{t-1})$$

$$\beta'_t = (1-b)(\mu'_t - \mu'_{t-1}) + b\beta'_{t-1}$$

Tabell 3 viser hvordan dette gjøres på en forenklet tidsserie som beskriver årlig vekst i hovedabonnement og hvordan prognosen fire år fram utvikler seg. $a = 0,5$, $b = 0,6$.

Tabell 3 viser hvordan Holts metode fungerer.

4.2 Prognosemodell, Holts metode

Prognosemodellen basert på Holts metode kan uttrykkes enkelt på følgende matematisk form:

$$F_{t+m} = \mu'_t + m\beta'_t$$

Ut fra denne formelen ser vi at prognosen er en fast stigning som vil endre seg når nye observasjoner kommer til. μ'_t og β'_t finnes ved hjelp av rekursjonsformelen som er beskrevet tidligere.

Intuitivt vil vi forstå at har vi en tidsserie som er voksende vil vi i prognosene få en økende verdi med konstant trend. Har vi for eksempel en tidsserie som inneholder

sesong vil det være naturlig å anta at kurven, observasjonene i framtiden, vil inneholde sesong også i framtiden. Konklusjonen blir altså at en prognosemodell basert på Holts metode ikke vil være aktuell å bruke på andre tidsserier enn de som inneholder en lineær trend og ikke sesongkomponenter.

4.2.1 Eksempel

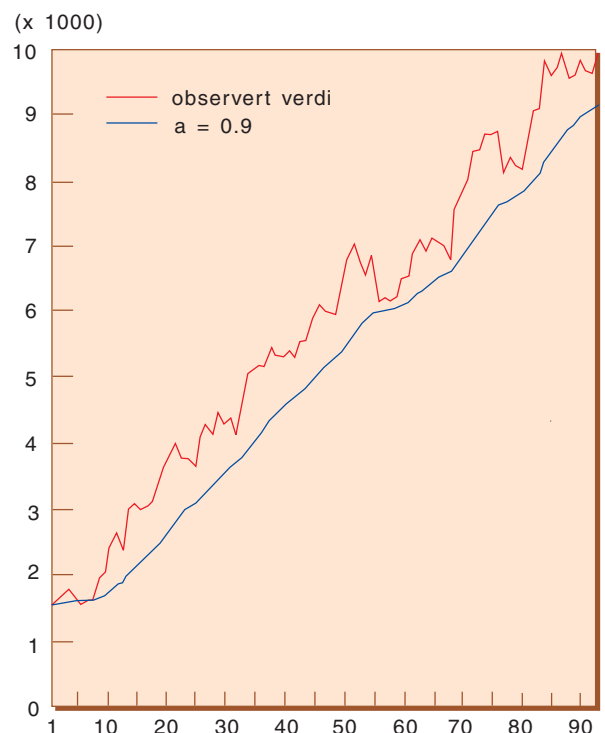
For å kunne vise Holts metode som prognosemetode rekapitulerer vi forrige tal eksempell hvor vi i tillegg får inn beregning av prognosene framover. Historikken er fra 1982 til 1990 og prognosen fra 1991 til 1992.

Tabell 4 viser hvordan Holts metode fungerer når en prognose skal konstrueres. Denne tabellen kan ses på som en fortsettelse av tabell 3.

Figur 8 viser observerte verdier og glattede verdier, samt prognose.

La oss ta tre forskjellige tidsserier hvor én er med trend og sesong, én er med trend og uten sesong (lineær trend), og én uten trend og sesong (stasjonær). På alle tidsseriene bruker vi Holts metode med glattingskonstanter $a = 0,3$ $b = 0,5$.

I figur 9 ser vi at prognosen ville blitt en negativ kurve fordi data inneholder en



Figur 7 Data med trend som er glattet ved hjelp av eksponentiell glatting

Tabell 3 Holts metode

År	t	Antall abonnem-ent	Årlig endring	a=0,5 μ_t	Beregning	b=0,6 β_t	Beregning
0				78000,0	Valgt startverdi	4000,0	Valgt startverdi
1982	1	82324		82162,0	$(1-0,5)*82324 + 0,5*(78000,0+4000)$	4064,8	$(1-0,6)*(82162,0-78000,0) + 0,6*4000,0$
1983	2	86692	4368	86459,4	$(1-0,5)*86692 + 0,5*(82162,0+4064,8)$	4157,8	$(1-0,6)*(86459,4-82162,0) + 0,6*4064,8$
1984	3	90744	4052	90680,6	$(1-0,5)*90744 + 0,5*(86459,4+4157,2)$	4183,2	$(1-0,6)*(90680,6-86459,4) + 0,6*4157,8$
1985	4	98744	8000	96803,9	$(1-0,5)*98744 + 0,5*(90680,6+4183,2)$	4959,2	$(1-0,6)*(96803,9-90680,6) + 0,6*4183,2$
1986	5	106494	7750	104128,6	$(1-0,5)*106494 + 0,5*(96803,9+4959,2)$	5905,4	$(1-0,6)*(104128,6-96803,9) + 0,6*4959,2$
1987	6	112850	6356	111442,0	$(1-0,5)*112850 + 0,5*(104128,6+5905,4)$	6468,6	$(1-0,6)*(111442,0-104128,6) + 0,6*5905,4$
1988	7	119421	6571	118665,8	$(1-0,5)*119421 + 0,5*(111442,0+6468,6)$	6770,7	$(1-0,6)*(118665,8-111442,0) + 0,6*6468,6$
1989	8	125079	5658	125257,7	$(1-0,5)*125079 + 0,5*(118665,8+6770,7)$	6699,2	$(1-0,6)*(125257,7-118665,8) + 0,6*6770,7$
1990	9	131347	6268	131652,0	$(1-0,5)*131347 + 0,5*(125257,7+6699,2)$	6577,2	$(1-0,6)*(131652,0-125257,7) + 0,6*6699,2$

Tabell 4 Hvordan Holts metode fungerer når en prognose skal konstrueres

År	t	Antall abonnem-ent	Årlig endring	a=0,5 μ_t	Beregning	b=0,6 β_t	Beregning
1990	9	131347	6268	131652,0	$(1-0,5)*131347 + 0,5*(125257,7+6699,2)$	6577,2	$(1-0,6)*(131652,0-125257,7) + 0,6*6699,2$
1991	10			138229	$131652,0+6577,2$		
1992	11			144806	$138229+6577,2$		

endring i nivået som modellen tolker som en trend.

I figur 10 virker prognosen fornuftig sett i relasjon til de data vi arbeider på. Det vises at data inneholder et gitt nivå og en gitt trend, og det rette valget vil derfor være en glattingsmodell med nivå og trend.

I figur 11 ser vi at dataene inneholder både nivå, trend og sesong. Når vi bruker en modell som forutsetter bare nivå og trend tolkes sesongsvingningene som trendendringer, og resultatet blir ubrukelig.

5 Tidsrekker med både trend og sesong

En sesongvariasjon er definert som et regelmessig mønster som gjentar seg år for år. Dette mønsteret er avhengig av at målingene foretas på samme tidspunkt år for år. Det finnes også andre metoder som behandler dette, men vi skal her ta for oss Holt-Winters metode som er en utvidelse av Holts metode.

Tidligere ble det ved hjelp av en konstant a, tatt hensyn til at parametrene i modellen endrer seg over tiden. Alle parametrene ble imidlertid påført den samme endringen over tid. I en tidsrekke med til-

feldige svingninger er det ikke sikkert at trenden og sesongsvingningene endrer sitt mønster like raskt.

Vanligvis vil sesongkomponentene være mer stabile enn trenden. Winter utviklet to modeller som tar hensyn til dette. Modellene har to parametre i tillegg til sesongparametrene som oppdateres rekursivt. Parametrene er:

- nivåparameter
- trendparameter β
- sesongparametere S_k

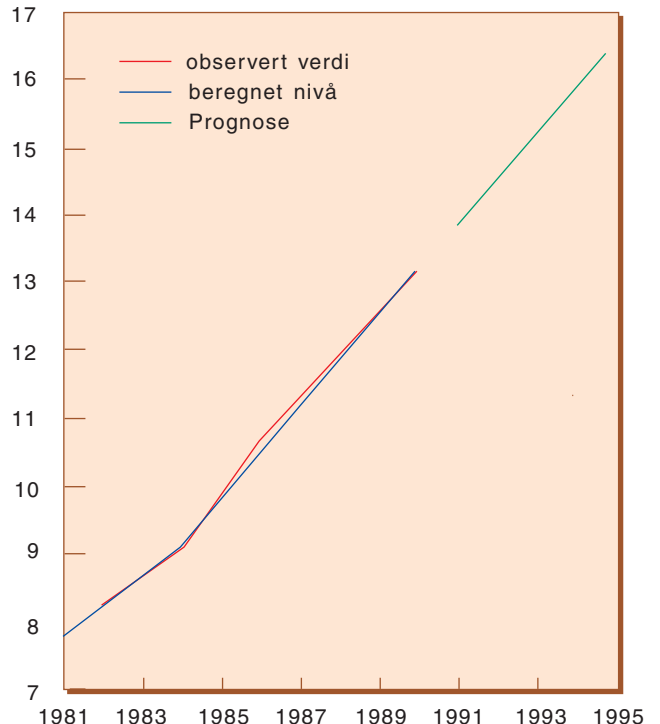
Metoden er basert på lineær trend, men den kan også brukes i andre tilfeller.

Det er utviklet både en additiv og en multiplikativ modell

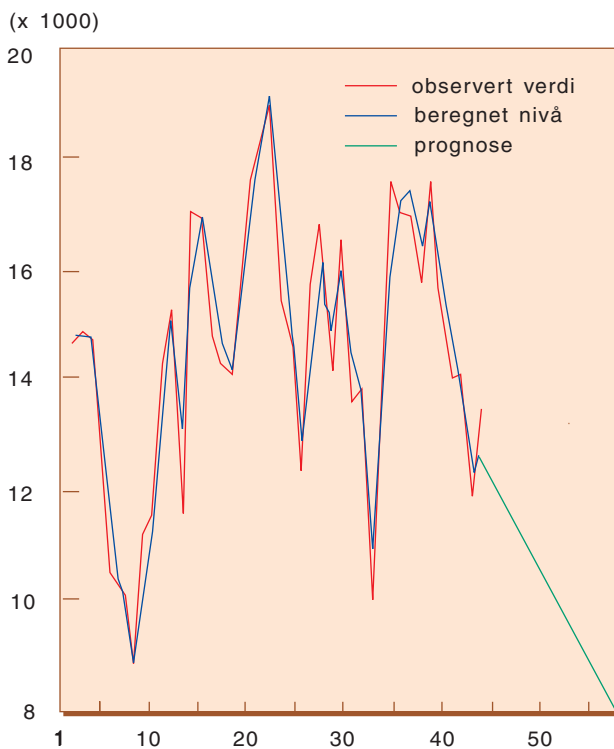
Additiv modell:

$$y_t = Trend_t + Sesong_t + \epsilon_t$$

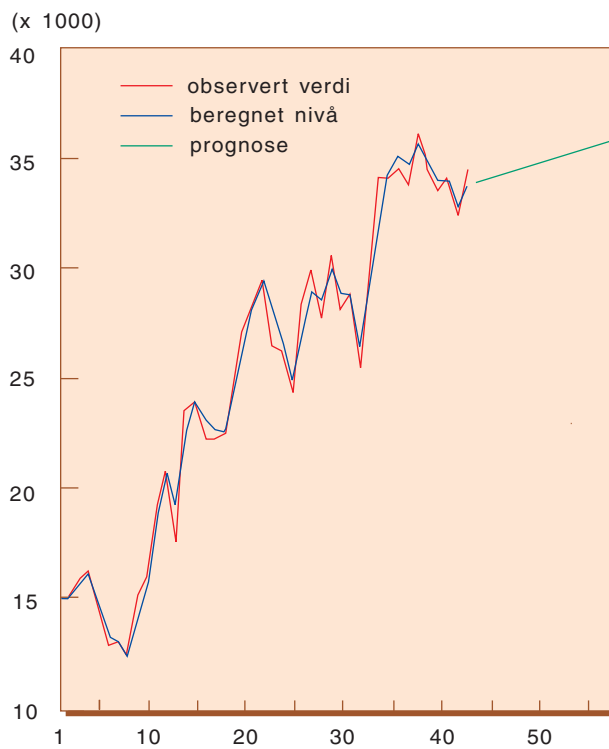
(x 10 000)



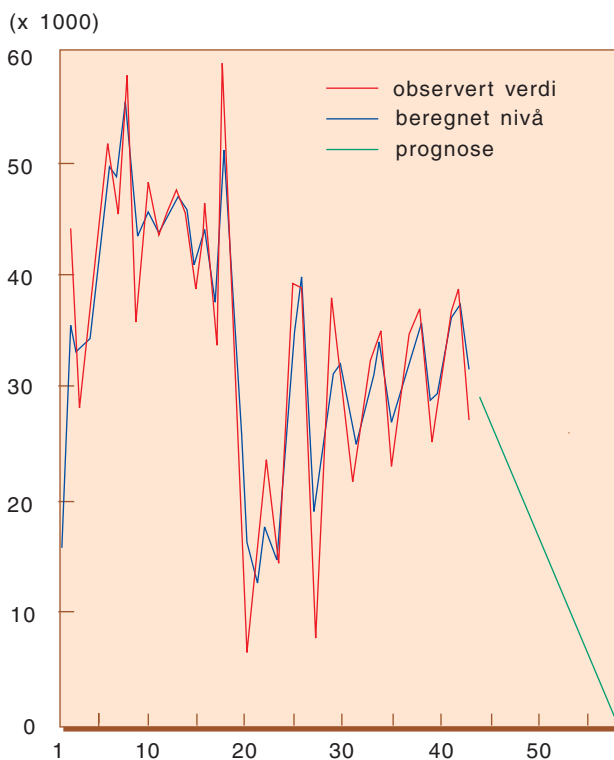
Figur 8 Observerte verdier og glattede verdier, samt prognose



Figur 9 Historikk, tilpasning og prognose for tidsserie uten trend og uten sesong (stasjonær)



Figur 10 Historikk, tilpasning og prognose for tidsserie med trend og uten sesong (ikke stasjonær)



Figur 11 Historikk, tilpasning og prognose for tidsserie med trend og sesong

Multiplikativ modell:

$$Y_t = Trend_t * Sesong_t + \epsilon_t$$

$$\mu_t = \mu_{t-1} + \beta_{t-1}$$

I tillegg forutsettes at sum av sesongutslagene er lik 0.

$$\sum_{k=1}^m S_{t+k} = 0$$

Holt-Winters metode inneholder tre glattingsparametere a, b, c , i motsetning til Holts metode som er beskrevet i avsnitt 2. Metoden er basert på Holts metode og kan betraktes som en utvidelse av denne.

Videre vil vi begrense oss til å beskrive den additive metoden.

Additiv modell

Modellen er som tidligere nevnt:

$$y_t = T_t + S_t + \epsilon_t$$

der

$$T_t = \mu_t$$

og

m er antall sesongutslag pr år.

5.1 Beregning av nivået

La oss ta utgangspunkt i Holts metode og utvide denne slik at sesongutslagene kan bli tatt hensyn til i beregningene. Anta at vi befinner oss i tidspunkt t . Holts metode for beregning av nivået er da:

$$\mu'_t = (1 - a) y_t + a(\mu'_{t-1} + \beta'_{t-1})$$

For å innlemme sesongutslagene kalles estimatet for sesongutslaget i tiden $t+1$ for S'_{t+1-m} .

Det nye nivået μ'_t kan estimeres ved differansen mellom observasjonen og sesongutslaget S'_{t-m} (observasjonen er da justert for sesongutslaget). Videre kan det nye nivået også estimeres ved det gamle nivået μ'_{t-1} pluss stigningen (trenden) β'_{t-1} , som i Holts metode. Vi får da en veid sum av de to anslagene:

$$\mu'_t = (1 - a) (y_t - S'_{t-m}) + a(\mu'_{t-1} + \beta'_{t-1})$$

Tabell 5 Hvordan Holt-Winters additive metode fungerer når en prognose skal konstrueres

År	Se- song	t	Antall ab. y_t	$a=0,4$ μ_t	$b=0,5$ β_t	$c=0,6$ S_t			
	-3	-3				500			
	-2	-2				-500			
	-1	-1				0			
	0	0	3000	Startverdi	0	250			
1984	1	1	4425	3555	(1-0,4)*(4425-500) + 0,4*(3000+0)	278	(1-0,5)*(3555-3000) + 0,5*0	648	(1-0,6)*(4425-3555) + 0,6*500
	2	2	2826	3529	(1-0,4)*(2826+500) + 0,4*(3555+278)	126	(1-0,5)*(3529-3555) + 0,5*278	-581	(1-0,6)*(2826-3529) - 0,6*500
	3	3	3329	3459	(1-0,4)*(3329+0) + 0,4*(3529+126)	28	(1-0,5)*(3459-3529) + 0,5*126	-52	(1-0,6)*(3329-3459) + 0,6*0
	4	4	4101	3705	(1-0,4)*(4101+250) + 0,4*(3459+28)	137	(1-0,5)*(3705-3459) + 0,5*28	308	(1-0,6)*(4101-3705) + 0,6*250
1985	1	5	5188	4261	(1-0,4)*(5188+648) + 0,4*(3705+137)	346	(1-0,5)*(4261-3705) + 0,5*137	760	(1-0,6)*(5188-4261) + 0,6*648
	2	6	4568	4932	(1-0,4)*(4568-581) + 0,4*(4261+346)	509	(1-0,5)*(4932-4261) + 0,5*346	-494	(1-0,6)*(4568-4932) - 0,6*581
	3	7	5781	5676	(1-0,4)*(5781-52) + 0,4*(4932+509)	626	(1-0,5)*(5676-4932) + 0,5*509	11	(1-0,6)*(5781-5676) - 0,6*52
	4	8	3592	4491	(1-0,4)*(3529+308) + 0,4*(5676+626)	-279	(1-0,5)*(4491-5676) + 0,5*626	-175	(1-0,6)*(3592-4491) + 0,6*308
1986	1	9	4847	4137	(1-0,4)*(4847+760) + 0,4*(4491-279)	-317	(1-0,5)*(4137-4491) - 0,5*279	740	(1-0,6)*(4847-4137) + 0,6*760
	2	10	4365	4444	(1-0,4)*(4365-494) + 0,4*(4137-317)	-5	(1-0,5)*(4444-4137) - 0,5*317	-328	(1-0,6)*(4365-4444) - 0,6*494
	3	11	4522	4482	(1-0,4)*(4522+11) + 0,4*(4444-5)	17	(1-0,5)*(4482-4444) - 0,5*5	22	(1-0,6)*(4522-4482) + 0,6*11
	4	12	4785	4776	(1-0,4)*(4785+175) + 0,4*(4482+17)	155	(1-0,5)*(4776-4482) + 0,5*17	-101	(1-0,6)*(4785-4776) - 0,6*175
1987	1	13	4513	4236	(1-0,4)*(4513+740) + 0,4*(4776+155)	-192	(1-0,5)*(4239-4776) + 0,5*155	555	(1-0,6)*(4513-4236) + 0,6*740
	2	14	3894	4151	(1-0,4)*(3894-328) + 0,4*(4236-192)	-139	(1-0,5)*(4151-4236) - 0,5*192	-300	(1-0,6)*(3894-4151) - 0,6*328
	3	15	4656	4385	(1-0,4)*(4656+22) + 0,4*(4151-139)	48	(1-0,5)*(4385-4151) - 0,5*139	122	(1-0,6)*(4656-4385) + 0,6*22
	4	16	3390	3868	(1-0,4)*(3390-101) + 0,4*(4385+48)	-235	(1-0,5)*(3868-4656) + 0,5*48	-252	(1-0,6)*(3390-3868) - 0,6*101
1988	1	17	5881	4649	(1-0,4)*(5881+555) + 0,4*(4649-235)	273	(1-0,5)*(4649-3868) - 0,5*235	825	(1-0,6)*(5881-4649) + 0,6*555
	2	18	3778	4416	(1-0,4)*(3778-300) + 0,4*(4236+273)	20	(1-0,5)*(4416-4649) + 0,5*273	-435	(1-0,6)*(3778-4416) - 0,6*300
	3	19	671	2104	(1-0,4)*(671+122) + 0,4*(4416+20)	-1146	(1-0,5)*(2104-4416) + 0,5*20	-500	(1-0,6)*(671-2104) + 0,6*122
	4	20	1663	1532	(1-0,4)*(1663+252) + 0,4*(2104-1146)	-859	(1-0,5)*(1532-2104) - 0,5*1146	-99	(1-0,6)*(1663-1532) - 0,6*252
1989	1	21	2378	1201	(1-0,4)*(2378+825) + 0,4*(1532-859)	-595	(1-0,5)*(1201-1532) - 0,5*859	966	(1-0,6)*(2378-1201) + 0,6*825
	2	22	1453	1375	(1-0,4)*(1453-435) + 0,4*(1201+595)	-210	(1-0,5)*(1375-1201) - 0,5*595	-230	(1-0,6)*(1453-1375) - 0,6*435
	3	23	2401	2206	(1-0,4)*(2401+500) + 0,4*(1375-210)	311	(1-0,5)*(2206-1375) - 0,5*210	-222	(1-0,6)*(2401-2206) - 0,6*500
	4	24	3961	3443	(1-0,4)*(3961-99) + 0,4*(2206+311)	773	(1-0,5)*(3443-2206) + 0,5*311	148	(1-0,6)*(3961-3443) - 0,6*99
1990	1	25	3909	3452	(1-0,4)*(3909+966) + 0,4*(3443+773)	391	(1-0,5)*(3452-3443) + 0,5*773	762	(1-0,6)*(3909-3452) + 0,6*966
	2	26	786	2147	(1-0,4)*(786-230) + 0,4*(3452+391)	-457	(1-0,5)*(2147-3452) + 0,5*391	-682	(1-0,6)*(786-2147) - 0,6*230
	3	27	2312	2196	(1-0,4)*(2312-222) + 0,4*(2147-457)	-204	(1-0,5)*(2196-2147) - 0,5*457	-87	(1-0,6)*(2312-2196) - 0,6*222
	4	28	3797	2986	(1-0,4)*(3797+148) + 0,4*(2196-204)	293	(1-0,5)*(2986-2196) - 0,5*204	413	(1-0,6)*(3797-2986) + 0,6*148
1991	1	29	3012	2662	(1-0,4)*(3012+762) + 0,4*(2986+293)	-16	(1-0,5)*(2662-2986) + 0,5*293	598	(1-0,6)*(3012-2662) + 0,6*762
	2	30	2176	2773	(1-0,4)*(2176-682) + 0,4*(2662-16)	48	(1-0,5)*(2773-2662) - 0,5*16	-648	(1-0,6)*(2176-2773) - 0,6*682

Her er a glattingsparameteren som må bestemmes på best mulig måte på samme måte som for eksponentiell glatting og Holts metode.

5.2 Beregning av trenden

La oss igjen ta utgangspunkt i Holts metode og utvide denne slik at sesongutslagene kan bli tatt hensyn til i beregningen av trenden. Anta at vi befinner oss i tidspunkt t . Holts metode for beregning av trenden er da:

$$\beta'_t = (1-b)(\mu'_t - \mu'_{t-1}) + b\beta'_{t-1}$$

Den nye stigningskoeffisienten β'_t kan estimeres ved differansen mellom de estimerte nivåene μ'_t og μ'_{t-1} og disse estimatene er allerede justert for sesongutslagene. Derfor er det ikke nødvendig med sesongkorleksjon her. Derfor kan den nye stigningen også estimeres på

samme måte som i Holts metode. Vi får da en veid sum av de to anslagene som er lik Holts metode:

$$\beta'_t = (1-b)(\mu'_t - \mu'_{t-1}) + b\beta'_{t-1}$$

Her er b glattingsparameteren som må bestemmes på best mulig måte på samme måte som for eksponentiell glatting og Holts metode.

5.3 Beregning av sesongutslagene

For å beregne trenden er det nødvendig å vite hvordan sesongutslaget kan uttrykkes. Sesongutslaget, S_t , kan estimeres som differansen mellom observasjonen y_t og nivået μ'_t . Videre kan sesongutslaget også estimeres ved det gamle estimatet for sesongutslaget, S'_{t-m} (der m er antall sesongutslag). Vi får da en veid sum av de to anslagene:

$$S'_t = (1-c)(y_t - \mu'_t) + cS'_{t-m}$$

Her er c glattingsparameteren som må bestemmes på best mulig måte, tilsvarende som for eksponentiell glatting og Holts metode. m er antall sesongutslag pr år.

5.4 Oppsummering

Det fullstendige formelverk for Holt-Winters metode er:

$$\mu'_t = (1-a)(y_t - S'_{t-m}) + a(\mu'_{t-1} + \beta'_{t-1})$$

$$\beta'_t = (1-b)(\mu'_t - \mu'_{t-1}) + b\beta'_{t-1}$$

$$S'_t = (1-c)(y_t - \mu'_t) + cS'_{t-m}$$

For å kunne vise Holt-Winters metode som prognosemetode rekapitulerer vi forrige talleksempel hvor vi i tillegg får inn beregning av prognosene framover.

Historikken er fra 1982 til 1990 og prognosen fra 1991 til 1992.

Tabell 5 viser hvordan Holt-Winters additive metode fungerer når en prognose skal konstrueres.

5.5 Prognosemodell, Holt-Winters metode

Prognosemodellen basert på Holt-Winters metode kan uttrykkes enkelt på følgende matematisk form:

$$F_{t+k} = \mu'_t + k\beta'_t + S'_{t+k-Im}$$

Velg det laveste positive heltall for l slik at likningen under er oppfylt:

$$k - Im \leq 0$$

Programrutinen rundt Holt-Winters prognosemodell er følgende uttrykk. NB! k er her år og m er sesong:

FOR $k:=1$ TO 2 DO

FOR $m:=1$ TO 4 DO

$$F_{t+k} * m = \mu'_t + k * m\beta'_t + S'_{t+k-m}$$

END FOR;

END FOR;

Denne rutinen lager prognoser pr kvartal to år fram i tiden.

Ut fra denne formelen ser vi at prognosen er en fast stigning som vil endre seg når nye observasjoner kommer til. μ'_t , β'_t og alle sesongutslagene, S'_t , finnes ved hjelp av rekursjonsformelen som er beskrevet tidligere.

5.5.1 Eksempel

For å kunne vise Holt-Winters metode som prognosemetode rekapitulerer vi forrige talleksempel hvor vi i tillegg får inn beregning av prognosene framover.

Historikken er fra 1982 til 1990 og prognosen fra 1991 til 1992.

Tabell 6 viser hvordan Holt-Winters additive metode fungerer når en prognose skal konstrueres. Denne tabellen kan ses på som en fortsettelse av tabell 5.

Figur 12 viser historikk, tilpasning og prognose for tidsserien over.

La oss se på de samme eksemplene som sist og i tillegg få prognosen fram ved hjelp av formlene som er presentert over. Vi tar utgangspunkt i tre forskjellige tidsserier hvor én er med trend og sesong, én er med trend og uten sesong (lineær trend), og én uten trend og sesong (stasjonær). På alle tidsseriene bruker vi glattingskonstant $a = 0,3$, $b = 0,5$, $c = 0,3$.

I figur 13 ser vi at prognosen får en negativ utvikling fordi data kun inneholder ett nivå og modellen prøver da å beregne trend og sesong ut fra variasjonen i nivået.

I figur 14 ser vi at prognosen blir negativ selv om dataene viser en svak positiv utvikling. Dette skyldes at modellen prøver å estimere sesongutslag i tillegg til trend.

I figur 15 ser vi at dataene inneholder både nivå, trend og sesong. Når vi bruker en modell som forutsetter nivå, trend og sesong estimerer modellen de komponentene dataene inneholder.

6 Hvordan bør glattingsparametrene bestemmes

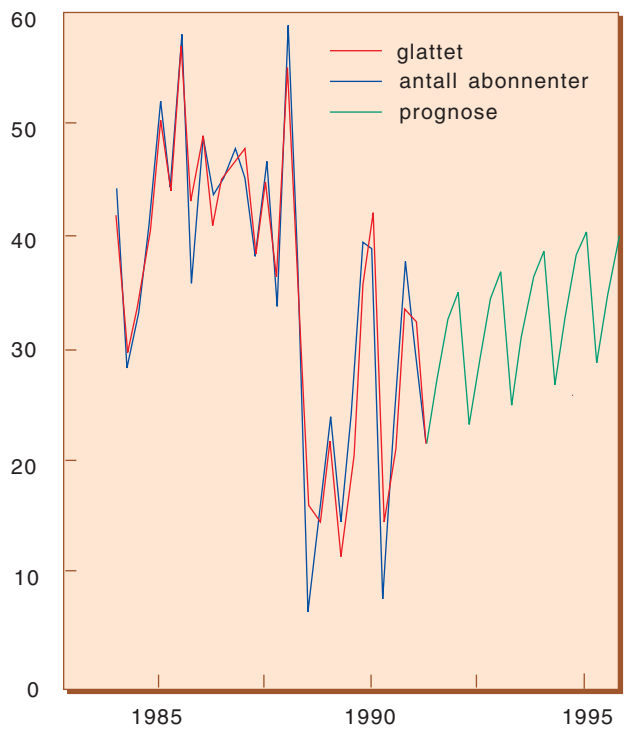
Til sist bør det nevnes at selv om metoden er gjennomgått mangler det noe for å kunne få et helhetlig inntrykk

av hva som må gjøres. Parametrene a , b , c må bestemmes på best mulig måte. Dette gjøres ved å minimalisere kvadratavviket for prognosen en tidsenhet fram i forhold til den observerte verdien. Dette gjøres ved å justere glattingsparametrene slik at summen av kvadratet av ettskrittprognosefeilen, Q , blir minst mulig.

Dette kan uttrykkes på følgende form:

$$\min_{a, b, c} Q = \sum_{t=1}^n (y_t - F_{(t-1)+1})^2$$

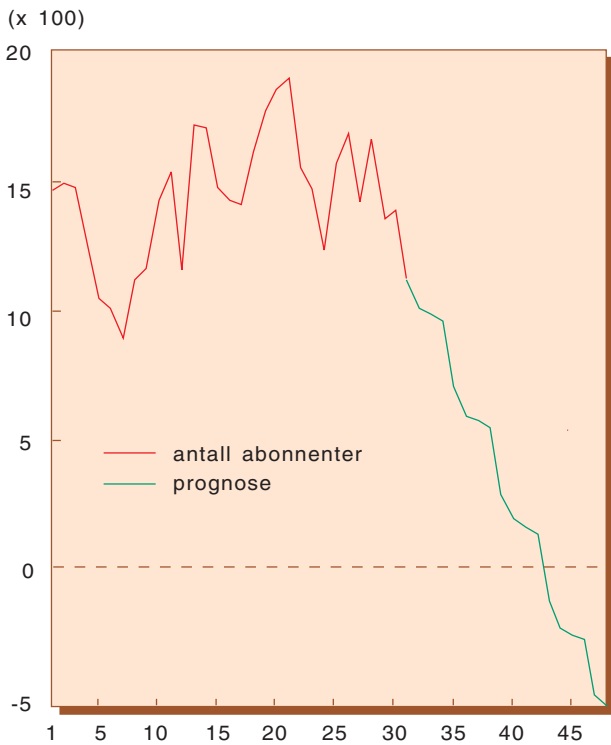
(x 100)



Figur 12

Tabell 6 Hvordan Holt-Winters additive metode fungerer når en prognose skal konstrueres. Denne tabellen kan sees på som en fortsettelse på tabell 5

År	Sesong	t	Antall ab. y_t	$a=0,4 \mu'_t$	$b=0,5 \beta'_t$	$c=0,6 S'_t$
1991	1	29	3012	2662	(1-0,4)*(3012+762) + 0,4*(2986+293)	598
	2	30	2176	2773	(1-0,4)*(2176-682) + 0,4*(2662-16)	-648
	3	31	2734		2773+(1*48)-87	-87
	4	32	3282		2773+(2*48)+413	413
1992	1	33	3514		2773+(3*48)+598	598
	2	34	2317		2773+(4*48)-648	-648
	3	35	2925		2773+(5*48)-87	-87
	4	36	3474		2773+(6*48)+413	413



Figur 13 Historikk, tilpasning og prognose for tidsserie uten trend og uten sesong (stasjonær)

7 Hva skal velges: tidsrekkeanalyse eller regresjon

Til tross for at tidsrekkemodeller basert på glatting er fleksible og har vist seg å være effektive blir de ofte undervurdert. Dette fordi når en variabel framskrives utelukkende på bakgrunn av historiske verdier, kan ikke metoden svare på endringer i en bransje eller endrede forutsetninger om framtiden. Disse modellene her virker ut fra at forhold i framtiden blir som de var i historien.

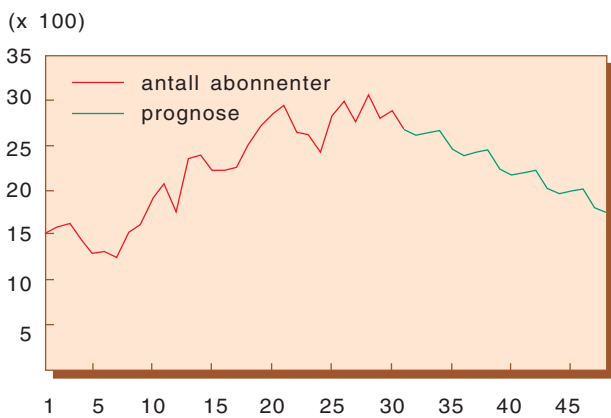
Likevel har det vist seg at disse modellene gir godt resultat innen et kortsiktig perspektiv, fordi endringer som påvirker en bransje sjelden kommer raskt, og som oftest er forutsetningene stabile på kort sikt. Andre fordeler er at modellene trenger lite data før de gir relativt gode resultater, i tillegg til at behandlingstid er relativt lav er det mulig å benytte metodene på mange delprognoser samtidig.

Regresjon eller lineære forklaringsmodeller gir best resultat der det finnes klart definerte sammenhenger mellom forklaringsvariabel og det som skal prognostiseres. Er det derimot en variabel som det ikke finnes data for eller at sammenhengene er uklare, vil regresjon trolig gi et dårlig resultat. I slike situasjoner er det bedre å bruke tidsrekkemodeller, som prognostiserer ut fra tidligere verdier på variabelen og ikke lar prognosen bli påvirket av en mengde uklare faktorer.

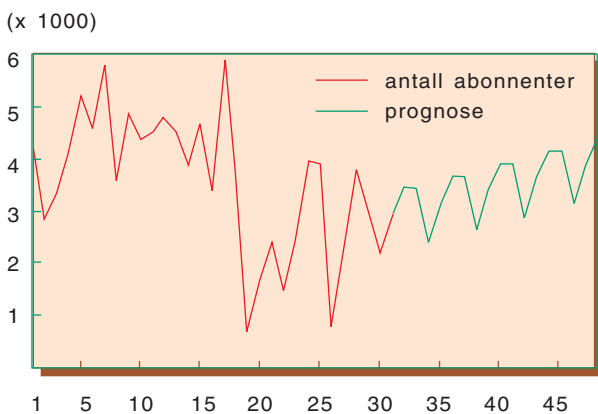
Mangelen ved tidsrekkemodellene er at all informasjon som er nødvendig for å lage prognoser for framtiden på en eller annen måte må være innbakt i de historiske observasjonene. Enten de historiske dataene ses på som et "mønster" eller som en forklaringsvariabel, er det lite trolig at dette er nok informasjon til å lage gode prognoser utover en kort tidshorisont. Hvis da ikke tidsserien er usedvanlig stabil og lite påvirket av ytre faktorer.

Rent praktisk er tidsrekkemodeller best når det er snakk om prognoser med kort tidshorisont, for eksempel ut året og neste år. For de mer langsiktige modeller, som for

eksempel etterspørsel etter hovedabonnement fram til år 2000, er en modell med forklaringsvariable å foretrekke.



Figur 14 Historikk, tilpasning og prognose for tidsserie med trend og uten sesong (ikke stasjonær)



Figur 15 Historikk, tilpasning og prognose for tidsserie med trend og sesong

Regresjonsmodeller

AV KJELL STORDAHL, CARLO HJELKREM OG JOHANNES BØE

1 Generelt

Regresjonsmodeller er en av de mest benyttede modelltyper for prognostisering. Styrken til regresjonsmodellene er at det er lett og samtidig modellmessig enkelt å bygge modeller med flere forklaringsvariable. Økonomene bruker i stor grad regresjonsmodeller der det inkluderes ulike økonomiske variable for å lage prognoser.

Regresjonsmodeller kan brukes til å lage både kort-, medium- og langtidspregninger. Generelt kan det sies at regresjonsmodeller kan egne seg svært godt til å lage langtidspregninger fordi det er lettere å modellere inn årsakssammenhenger i regresjonsmodellene i forhold til andre prognosemodeller. På den annen side er det viktig å være klar over at selv om en har laget en svært god modell ved å benytte flere forklaringsvariable, er det ikke sikkert at mulighetene er til stede for å lage gode prognoser. Det skyldes at de prognosene som skal lages betinger at det også må lages prognoser for utviklingen av de ulike forklaringsvariable som inngår i modellen. Dersom disse prognosene er svært usikre hjelper det ikke å selve prognosemodellen har en svært god tilpasning. Det er kombinasjonen av kvaliteten på prognosemodell og på prognosene for forklaringsvariable i modellen som til sammen står for kvaliteten på prognosene.

Av klassiske verker innen regresjonsanalyse henvises det til [1], [2], [3], [4] og [5].

I økonometri er en opptatt av å måle og empirisk teste økonomiske sammenhenger. Viktige momenter er teori, observasjon og statistiske sammenhenger. Fra økonomisk teori kan det, under gitte forutsetninger, utledes utsagn om sammenhenger mellom økonomiske variable, ofte uttrykt som likninger. Imidlertid er koeffisientene i slike likninger sjelden spesifisert numerisk. Mens økonomisk teori ofte kan angi fortegnet eller retningen til responsen i en økonomisk variabel når relevante forklaringsvariable endres, er teorien ikke i stand til å angi størrelsen på denne responsen. Det er her økonometrien hjelper oss ved å konfrontere teorien med faktiske observasjoner. Den statistiske modelleringen foregår som regel ved hjelp av regresjon. En naturlig forlengelse av de modeller som bygges vha økonometriske metoder er å la disse foreta projiseringer inn i fremtiden, dvs prognoser.

Av klassiske verker innen økonometri henvises det f eks til [6], [7], [8] og [9].

I denne artikkelen går vi spesielt inn på å lage prognoser for tellerskritt basert på den utvikling vi har sett de siste årene – det vil si fra og med 1989. Tilrettelegging av tellerskrittdata med de justeringer av data som må gjøres, er omtalt i detalj i [10]. Kort skissert er problemstillingen at tellerskrittene lengde endres fra tid til annen. Dette er en av to måter å endre tellerskrittprisen på, – den andre er å endre selve kroneprisen. Endringer i tellerskrittprisen gir imidlertid tydelige hopp i data-rekken for tellerskritt, fordi tellerskrittene i ulike takstklasser som regel er forlenget (dvs at prisen er redusert) uten at dette er kompensert med tilsvarende økning i antall tellerskritt. Slike hopp kan i prinsippet modelleres inn i en prognosemodell ved hjelp av teknikker for beskrivelse av såkalte nivåskift. Imidlertid vil stadige nivåskift gi en så stor modell med så mange parametere at dette reduserer sikkerheten betraktelig. Vi har derfor valgt å regne om alle historiske tellerskrittdata til "volum tellerskritt", med tellerskrittprisen i 1993 som enhet. Vi får da en tidsserie som er

Tabell 1.1 Tellerskritt og volum tellerskritt pr kvartal. Region Oslo. Konstruerte tall

	Avleste ts	Volum ts
1.kv.89	61965	47586
2.kv.89	62438	47949
3.kv.89	57702	44480
4.kv.89	64169	49808
1.kv.90	61875	52378
2.kv.90	57287	48494
3.kv.90	52156	47446
4.kv.90	58736	53432
1.kv.91	54375	54150
2.kv.91	56051	55819
3.kv.91	52367	52151
4.kv.91	57984	57744
1.kv.92	60869	61051
2.kv.92	57151	57321
3.kv.92	57064	57196
4.kv.92	62131	62275
1.kv.93	66964	65005
2.kv.93	64370	62487

sammenliknbar over tid og uavhengig av historiske endringer i tellerskrittprisen. Omregningen fra tellerskritt til volum tellerskritt tar hensyn til ulike takstklassers vekt, fordeling mellom full og redusert takst og samtaletidsfordelingen for de ulike takstklasser. Se mer om dette i [10].

De modellene som utvikles og presenteres her, brukes ikke offisielt i Televerket. Dataene er noe fordreid av konfidensielle og forretningsmessige hensyn. For øvrig er modellene laget som om data var virkelige.

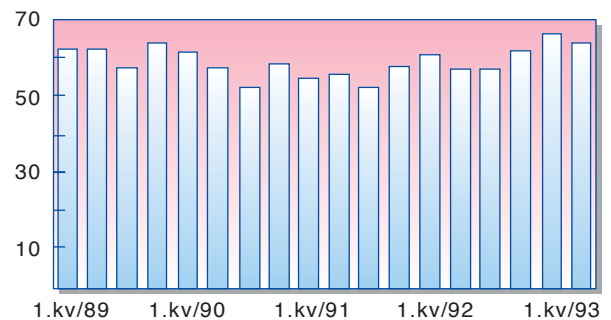
2 Aktuelle prognosemodeller

Før vi ser etter forklaringsfaktorer og på aktuelle modeller, er det naturlig å studere selve tidsrekken som da angir etterspørselen etter tellerskritt fra primo 1989 til medio 1993. Etterspørselsrekken som er vist i figur 2.1 er angitt i kvartalsvise observasjoner.

Det kan ses at denne tidsrekken skifter nivå ved enkelte årsskifter. Dette skyldes at lengden på tellerskrittet på ulike

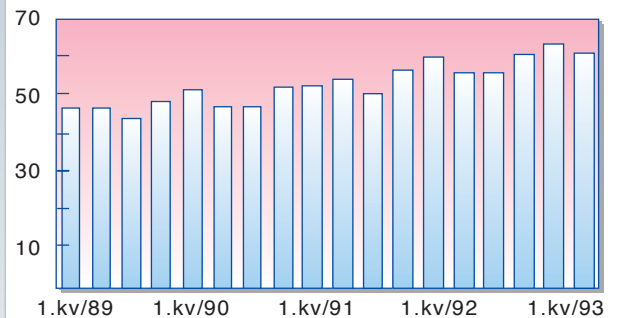
001.18:621.39

Tellerskritt (x1000)



Figur 2.1 Kvartalsvis tellerskrittutvikling fra 1/89 til 2/93

Volum tellerskritt (x1000)

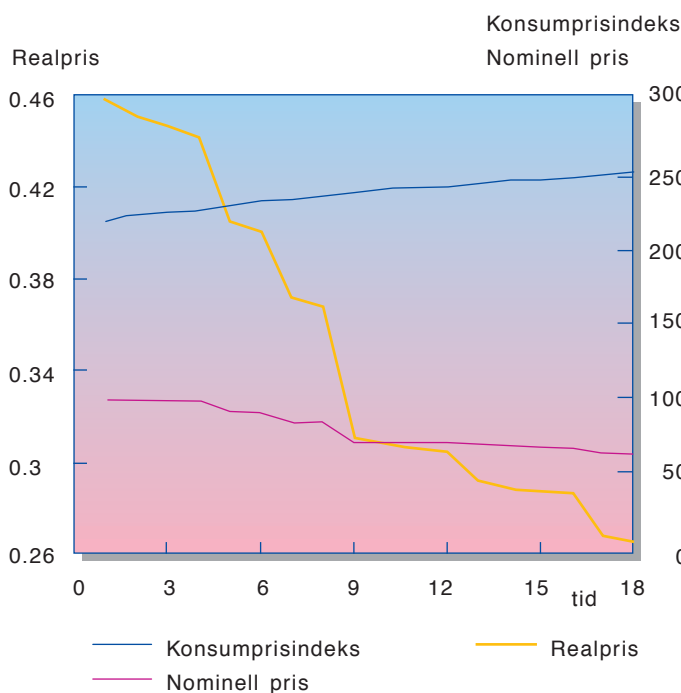


Figur 2.2 Kvartalsvis utvikling av volum tellerskritt fra 1/89 til 2/93

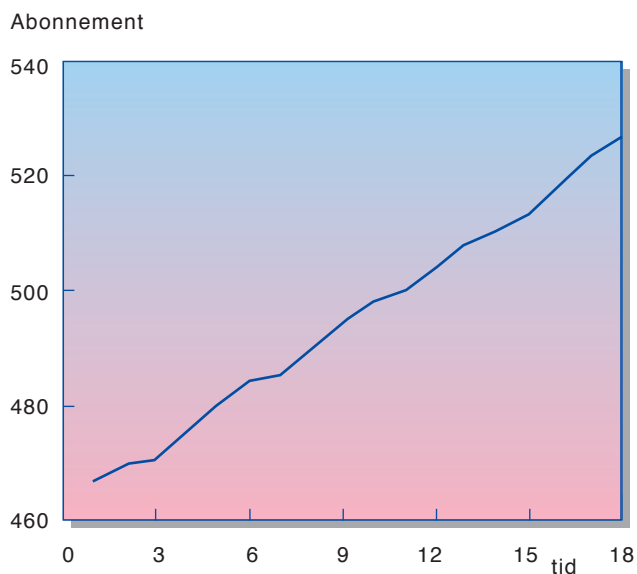
nasjonale eller internasjonale takstklasser endres ved årsskiftet. Dersom ikke alle abonnenter umiddelbart endrer på sin bruk proporsjonalt når tellerskritt lengden økes, vil tellerskrittantallet reduseres. Hopp eller nivåskift i tidsrekken gjør den også vanskelig å prognostisere. Det er derfor mer naturlig å se på trafikkvolumet som er upåvirket av endring av teller-

skrittlengden. Dette gjøres her ved å ta utgangspunkt i tellerskritt lengden siste år og transformere tellerskritt volumet i tidligere kvartaler ut fra forholdet mellom de respektive tellerskritt lengder. Figur 2.2 viser den justerte tellerskritt utviklingen, volum tellerskritt.

Det er denne tidsrekken som er grunnlaget for prognosemodellene vi skal vise. Tidsrekken har også det rette nivået som er korrekt tellerskritt lengde for siste år og som da vil være utgangspunkt for prognostiseringen.



Figur 2.3 Kvartalsvis utvikling av pris på volum tellerskritt: Realpris (venstre akse), konsumprisindeks (høyre akse) og nominell pris (høyre akse). Indekser fra 1/88 til 2/93



Figur 2.4 Kvartalsvis utvikling av antall innkoblede abonnement fra 1989 til medio 1993

vår pris-variabel hvis konsumprisindeksen endrer seg proporsjonalt med de nominelle prisendringene.

Vi forventer at en økning i prisene medfører en reduksjon i volum tellerskritt.

I slike analyser er det også vanlig å modellere endringen i prisene, ut fra en teori om at det ikke er prisnivået som påvirker etterspørselen, men *endringer* i prisnivået. Det ligger imidlertid utenfor rammen for denne artikkelen å prøve alle tenkelige teori-varianter. Vi følger derfor ikke denne tanken videre.

Spørsmålet blir så hvor raskt abonnentene reagerer på prisendringer. Når det gjelder priselastisiteter, opereres det med kort- og langtids priselastisiteter. Dette skyldes erfaring som viser at ikke alle abonnenter reagerer umiddelbart på prisendringer. Det kan ta opptil flere kvartaler før en prisendring får tilnærmet full effekt. Vi har derfor introdusert fem ulike prisforklaringsvariable. Den første er som beskrevet ovenfor: Realprisen forklarer etterspørselen etter volum tellerskritt i det kvartalet prisen gjelder. Den andre prisvariabelen har et "lag" på en periode, dvs den er forsinket med en periode. Det betyr at volum tellerskritt i f.eks 1. kvartal 1989 forklares med realprisen i 4. kvartal 1988. Den tredje prisvariabelen har et lag på to perioder, dvs at etterspørselen etter volum tellerskritt 1. kvartal 1989 forklares med realprisen i 3. kvartal 1988. Tilsvarende gjelder de to siste prisvariabelene som har lag på hhv tre og fire perioder. Med vårt underlag har det ikke vært mulig å gå lengre tilbake enn fire perioder.

2.1 Forklaringsvariable

2.1.1 Pris

Generell økonomisk teori for konsumentenes og bedriftenes tilpasning sier at etterspørselen etter en vare/tjeneste kan forklares ved forholdet mellom prisen på varen/tjenesten og prisen på alle andre varer/tjenester.

I modellering av etterspørselsrelasjoner vil som regel prisen på "alle andre" varer og tjenester representeres ved hjelp av en prisindeks, ofte gitt ved konsumprisindeksen. Vi har i vår analyse valgt å bruke gjennomsnittet for Statistisk Sentralbyrås konsumprisindeks i hvert kvartal. 1979 = 100.

I innledningen sa vi at prisen på tellerskritt kan endres på to måter; ved endringer i tellerskritt lengden eller ved endring i kroneprisen. Vi har derfor konstruert en prisindeks for volum tellerskritt som tar hensyn til begge disse faktorene. 1989 = 100. Indeksen er basert på gjennomsnittsprisene for kvartalet.

Ved å dividere de to prisindeksene på hverandre får vi frem det relative prisforholdet. Dette er da et uttrykk for realprisene i de enkelte periodene. Vi vil undersøke om etterspørselen etter volum tellerskritt avhenger av realprisnivået for tellerskrittene. Nominelle endringer i prisen på tellerskritt gir dermed ikke utslag i form av endringer i

2.1.2 Antall abonnement

Det er helt innlysende at det må være en sammenheng mellom antall abonnement og antall forbrukte tellerskritt. Dersom det ikke er noen abonnenter, vil det heller ikke bli produsert noen tellerskritt. Antall tellerskritt øker også med antall abonnement. Det er grunn til å tro at det er en sterk avhengighet mellom disse variablene. Figur 2.4 viser hvorledes den gjennomsnittlige abonnementsutviklingen for hver periode har vært i løpet av de siste årene.

Vi forventer at en økning i antall abonnement vi gi en økning i volum tellerskritt.

2.1.3 Tiden

Det er en generell oppfatning av markedet for tellerskritt at nye abonnement i gjennomsnitt ikke genererer like mange tellerskritt som de eksisterende: Når real-

prisene faller, vil nye abonnenter som ikke har like stor kjøpekraft etterspørre abonnement og eksisterende (bedrifts-)abonnenter øke det antall de allerede har, alt annet konstant (cet. par).

Når vi benytter antall abonnement som forklaringsfaktor, får vi estimert det gjennomsnittlige antall tellerskritt som abonnementsene har generert over dataperioden. Hvis vi antar at de nye abonnentene vil bidra til å redusere dette forholdet, må vi trekke inn noe som kan forklare dette i modellen. En måte å gjøre dette på, er å benytte tiden t som forklaringsvariabel. Den vil da representere en negativ trend i forhold til veksten som abonnementsmassen bidrar med.

Fordelen med å ha med tiden som forklaringsvariabel er at den er lett å prognostisere. Vi vil aldri ha noen usikkerhet beheftet til fremskrivning av tiden.

Variabelen er konstruert ved hjelp av verdier fra 1 til 18, voksende med en i hver periode.

Vi forventer at volum tellerskritt reduseres ettersom tiden går.

2.1.4 Virkedager

Vi vet at antall tellerskritt per kvartal er avhengig av antall virkedager det er i kvartalet. Telefonbruken er forskjellig på hverdager i forhold til lørdager og helligdager. Derfor tillegges forskjellig type dager ulik vekt – det vil si andre dager enn hverdager vil representere fraksjoner av hverdager der størrelsen på fraksjonen er avhengig av type dag. Vi har målt antall tellerskritt som etterspørres på de ulike hverdager, lørdager, søn- og helligdager og innklemte virkedager (f.eks. romjul og dager mellom Palmesøndag og Skjærtorsdag). Dette har gitt opphav til vekter gitt i tabell 2.1.

Dette er for øvrig også omtalt i [10].

Dersom antall virkedager var likt hvert kvartal, ville den ikke ha effekt som forklaringsvariabel. Her er det imidlertid endringer mellom kvartaler bl. a. på grunn av påske og pinse og innen kvartalene på

Tabell 2.1 Vekting for beregning av virkedager

	Vekt
Lørdag og søndag	0,3
Helligdager og innklemte virkedager	0,2 – 0,7
Resten	1

grunn av hvilke dager helligdagene faller på i julen og de offentlige høytidsdager faller på i mai, spesielle flytende helligdager og fordi det kan være ulike antall helger i de forskjellige kvartaler. På månedsbasis kan det være enda større individuelle variasjoner. Figur 2.5 viser tidsrekken.

Vi forventer at volum tellerskritt øker når antall virkedager øker.

2.1.5 Sesongvariable

Det ses av figur 2.2 at det er klare sesongmessige svingninger, noe som indikerer at det må bygges inn sesongvariable. Variabelen *Virkedager* vil forklare noe av disse sesongsvingningene, mens f.eks. 3. kvartal med fellesferie ikke fanges opp i denne variabelen. I tillegg introduserer vi derfor noe som i regresjonsanalysen kalles dummyvariable. Disse variablene har den egenskap at de er 1 på bestemte tidspunkter og 0 ellers. Vi må da ha dummyvariable som reflekterer endringene i de fire sesongene i løpet av året.

Det er imidlertid kun nødvendig å inkludere tre sesongvariable i modellen ved å bruke den ene sesongen som referansesesong. I vårt tilfelle brukes sesong 1 – det vil si 1. kvartal – som referanse, hvilket betyr at sesongutslagene 2, 3 og 4 er utslag i forhold til nivået på tidsrekken i 1. kvartal. På denne måten sparer vi en variabel, samtidig som vi unngår å få perfekt multikolinearitet under estimeringen, se avsnittet om korrelasjonsmatrisen for parametrene i kapittel 5.

Tidsrekken i figur 2.2 har 18 observasjoner. Siden variabelen *Tid* er definert med verdier fra 1 til 18, kan vi definere sesongvariablene som følger:

$$S_{2t} = 1 \text{ for } t = 2, 6, 10, 14, 18 \\ 0 \text{ ellers}$$

$$S_{3t} = 1 \text{ for } t = 3, 7, 11, 15 \\ 0 \text{ ellers}$$

$$S_{4t} = 1 \text{ for } t = 4, 8, 12, 16 \\ 0 \text{ ellers}$$

I selve modellen vil hver sesongvariable multipliseres med en parameter. Størrelsen på denne parameteren vil være det estimerte sesongutslaget. Vi ser av definisjonen på sesongvariablene at de vil være forskjellig fra 0 til forskjellige tider.

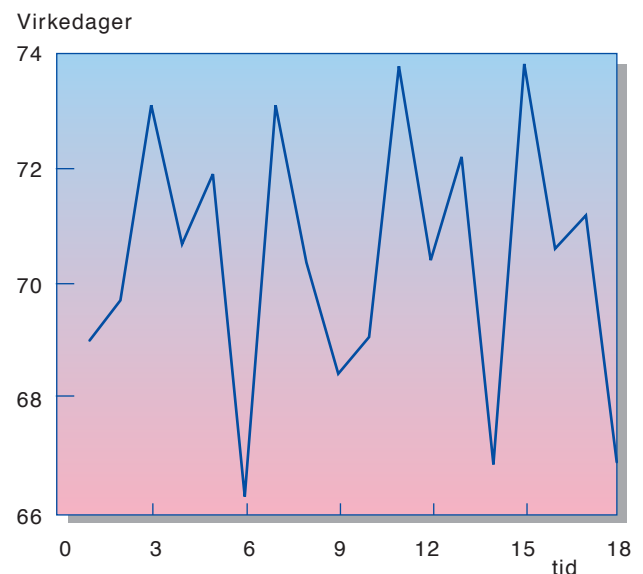
Etter sesongmønsteret i volum tellerskritt å dømme, forventer vi at sesong tre vil ligge lavere enn sesong en. De andre sesongene er litt ujevne – vi har ingen klar forventning om disse.

2.2 Ulike modelltyper

Totalt har vi nå introdusert 11 ulike forklaringsvariable, noe som med et eventuelt konstantledd utgjør 12 parametere. Totalt har vi 18 observasjoner. Dette indikerer i seg selv at antall variable er for stort. Hvis forutsetningene for en god regresjonsmodell skal være oppfylt (se kapittel 3 og 5) vil neppe mer enn en av prisvariablene kunne aksepteres på grunn av stor avhengighet. I så fall vil dette redusere antall parametere med 4.

2.2.1 Tradisjonell lineær modell

For å kunne estimere parametrene i en regresjonsmodell, brukes maksimum likelihood-prinsippet eller estimering basert på minste kvadratets metode. Disse to metodene fører til samme resultat dersom vi har en lineær regresjonsmodell. Med lineær form menes lineær i parametrene. Dersom imidlertid parametrene forekommer i komplekse former i modellen, må vi benytte andre måter å estimere parametrene på. Den metode eller klasse av metoder som da benyttes, kalles ikkelineær estimering. Her er det ikke mulig å beregne et analytisk uttrykk for de estimerte parametere. Det må isteden settes i gang en rekursiv estimer-



Figur 2.5 Antall virkedager per kvartal fra 1989 til 1993

ingsprosedyre som skal iterere seg frem til estimerte verdier.

Tar vi med alle forklaringsvariablene vil utgangspunktet for modellen være:

$$\begin{aligned} TS_t = & \beta_0 + \beta_{10}P_t + \beta_{11}P_{t-1} + \beta_{12}P_{t-2} \\ & + \beta_{13}P_{t-3} + \beta_{14}P_{t-4} + \beta_2AB_t \\ & + \beta_3T_t + \beta_4Vdag_t + \beta_{52}S_2t \\ & + \beta_{53}S_3t + \beta_{54}S_4t \end{aligned} \quad (2.1)$$

hvor

TS_t = etterspurt volum tellerskritt i periode t

P_{t-i} = realpris i periode t , $i = 0, 1, 2, 3, 4$.

AB_t = innkoblede abonnement i periode t

T_t = tiden i periode t

$Vdag_t$ = antall virkedager i periode t

S_j = dummyvariabel for sesong j , periode t

β_0, β_{1i} ($i = 1, 2, 3, 4$), $\beta_2, \beta_3, \beta_4, \beta_{5j}$ ($j = 2, 3, 4$) er parametere som skal estimeres.

Etter at regresjonsanalysen er foretatt, vil det vise seg om noen av de aktuelle variablene ikke får signifikant innvirkning. Da skal de i utgangspunktet utelates fra modellen.

Ser vi f.eks. på β_{10} angir den hvor mye volum tellerskritt endrer seg med når realprisen i den perioden prisen gjelder endrer seg med en enhet, cet. par., dvs den er den partielt deriverte av TS_t mhp P_t . For å få frem korttids priselastisiteten for denne perioden kan vi da multiplisere med P_t og dividere med TS_t . Vi har altså at korttids priselastisiteten er gitt ved

$$\beta_{10} * P_t / TS_t \quad (2.2)$$

Mens β_{10} er en konstant (punkt-estimat) ser vi at priselastisiteten er avhengig av forholdet mellom P_t og TS_t og vil derfor variere over tid.

Hvis effekten av en prisendring har slått tilnærmet fullt ut i etterspørselen etter volum tellerskritt etter 4 perioder, har vi at langtids priselastisiteten er gitt ved

$$\sum_{i=0}^4 \beta_{1i} * P_{t-i} / TS_t \quad (2.3)$$

På samme måte angir β_2, β_3 og β_4 hvor mye volum tellerskritt endrer seg når hhv abonnement, tid og virkedager endrer seg

med en enhet, og det kan her også beregnes elastisiteter. β_{5j} ($j = 2, 3, 4$) sier hvor mye sesongutslagene utgjør i forhold til 1. kvartal.

2.2.2 Loglineær modell

Denne type modeller er i utgangspunktet ikkelineære. Det er imidlertid ofte mulig ved bruk av en logaritmisk transformasjon å overføre den opprinnelige modellen til en lineær form.

Ofte vil begrunnelsen for valg av en slik modellform ligge i observasjon og teori for en variabls utvikling: En eksponentiell utvikling kan f.eks beskrives ved en eksponential-funksjon eller en funksjon der variablene inngår multiplikativt. En mye brukt variant er hentet fra økonomisk teori om bedriftenes tilpasning. Der benyttes en såkalt Cobb-Douglas produktfunksjon, hvor variablene (og konstanten) inngår multiplikativt med parametrene som eksponenter. Dette er overførbart til vårt eksempel og gir følgende modell når vi innlemmer alle de relevante forklaringsvariablene:

$$\begin{aligned} TS_t = & \beta_0 * P_t^{\beta_{10}} * P_{t-1}^{\beta_{11}} * P_{t-2}^{\beta_{12}} \\ & * P_{t-3}^{\beta_{13}} * P_{t-4}^{\beta_{14}} * AB_t^{\beta_2} \\ & * T_t^{\beta_3} * Vdag_t^{\beta_4} \\ & * e^{(\beta_{52}S_2t + \beta_{53}S_3t + \beta_{54}S_4t)} \end{aligned} \quad (2.4)$$

hvor e er grunntallet i eksponentialfunksjonen.

Også her vil det bli reduksjon i modellen etter at regresjonsanalysen er foretatt og kun de signifikante variable er inkludert.

I praksis gir ofte en slik utforming av etterspørselsfunksjonen en bedre tilpasning til data enn den lineære. Spesielt gjelder dette i tilfeller hvor vekst og sesongutslag er økende. Ofte kan imidlertid dette være vanskelig å avgjøre i utgangspunktet, og derfor prøves begge funksjonsformer for å se hvilken som gir best resultater.

En transformasjon over til lineær form i parametrene gjøres ved å ta logaritmen til (2.4). Det gir følgende:

$$\begin{aligned} \ln TS_t = & \ln \beta_0 + \beta_{10} \ln P_t + \beta_{11} \ln P_{t-1} \\ & + \beta_{12} \ln P_{t-2} + \beta_{13} \ln P_{t-3} \\ & + \beta_{14} \ln P_{t-4} + \beta_2 \ln AB_t \\ & + \beta_3 \ln T_t + \beta_4 \ln Vdag_t \\ & + \beta_{52} S_2t + \beta_{53} S_3t \\ & + \beta_{54} S_4t \end{aligned} \quad (2.5)$$

Elastisiterer vi TS_t mhp P_t direkte i (2.4) gir dette β_{10} , som er korttids priselastisiteten. I denne utformingen av etterspørselsfunksjonen er altså korttids priselastisiteten konstant, og gitt direkte ved denne parameteren. Langtids priselastisiteten er da gitt som

$$\sum_{i=0}^4 \beta_{1i} \quad (2.6)$$

2.2.3 Ikkelineær regresjonsmodell

I denne artikkelen er det ikke nødvendig å se på ikkelineære modeller fordi vi får meget god tilpasning med de prognosemodeller som vi har valgt ut. Dersom en kan unngå ikkelineære modeller, så er det en fordel. Det er her en risiko for at estimeringsprosessen kan havne i lokalt og ikke globalt minimumspunkt – noe som vil føre til faktisk feil resultater.

2.2.4 Veiet regresjonsmodell

I de tilfeller hvor variansen til støyledet øker eksempelvis med tiden, kan det være aktuelt å ta en logaritmisk transformasjon for å dempe denne økningen. Eventuelt kan det være aktuelt å veie de enkelte observasjoner etter et gitt mønster. En slik metode kalles veiet regresjon. I det data-underlaget som er her, har det vist seg unødvendig å foreta veiet regresjonsanalyse.

Før vi nå foretar selve modellbyggingen med basis i de forklaringsvariablene som er trukket frem, går vi igjennom de forutsetninger som skal være til stede i regresjonsanalysen og det verktøy som kan brukes til utvikling av prognosemodellen.

3 Stokastisk utforming – modellforutsetninger i regresjon

Prosessen med utarbeiding av en prognosemodell er godt beskrevet i [12]. De fleste av de prinsipper for modellering som gjennomgås der, benyttes i regresjonsanalysen.

Det fundamentale i modellbyggingsprosessen er residualene. Arbeidet går kort ut på å analysere residualene og ut fra dette eventuelt komme frem med forslag til forbedring av modellen. Residualet ved tidspunkt t betegnes med e_t . Det er definert som differansen mellom observasjonen ved tidspunkt og den tilpassede verdi som er den verdi som modellen anslår ved tidspunkt t .

I regresjonsmodellene innfører vi en variabel ε_t som også kalles støyledet. Denne

henges additivt på den lineære modellen (2.1) og multiplikativt på modellen (2.4). Denne variabelen angir det vi ikke greier å forklare i modellen. Det kan være variasjon i variable som ikke er spesifisert i modellen, avvik mellom faktisk funksjonsform og den vi har valgt, eventuelt målefeil i variablene samt tilfeldige avvik. Målet er selvsagt at støyleddene skal være små i forhold til observasjonene. Det er viktig under modellbyggingen å velge ut variable og finne frem til en funksjonsform av variable og parametere som inngår i modellen som gjør at det vi ikke greier å forklare – støyen – er minst mulig.

Residualene er realiseringen av støyen. Det er med andre ord de kvantitative verdiene på støyen som fremkommer når modellen er gitt og de tilpassede verdier beregnet.

Et krav til prognosemodeller er selvsagt at støyen er minst mulig. I regresjonsmodeller ivaretas dette i estimeringen ved bruk av minste kvadraters metode. For øvrig må det ivaretas i selve modellarbeidet ved riktig valg av forklaringsvariable og funksjonsform.

De krav som stilles til støyleddet, er:

- i) $E \varepsilon_t = 0$ for alle t (3.1)
- ii) $\text{Var } \varepsilon_t = \sigma^2$ for alle t (3.2)
- iii) $\text{Cov}(\varepsilon_t, \varepsilon_{t+k}) = 0$ for alle $t \neq k$ (3.3)
- iv) ε_t skal være Normalfordelt (3.4)

Med innføring av disse forutsetningene for støyleddene sier vi at modellene er gitt en stokastisk formulering.

Med hensyn til krav nr i) om at støyen skal ha forventningsverdi 0, så ivaretas dette automatisk i regresjonsanalysen. Minste kvadraters metode estimerer parametrene i modellen slik at gjennomsnittet av alle residualene i en regresjonsmodell er lik 0.

Som sagt må residualene analyseres for å undersøke om de tilfredsstillende disse kravene. Dersom ett eller flere krav ikke er tilfredsstillende, betyr det at estimeringen av parametrene i modellen blir mer usikker. Dette kan eksempelvis føre til forventningsskjevhet i parametrene.

4 Minste kvadraters metode og estimering i multipl regressjon

En lineær multipl regressjonsmodell kan skrives på form:

$$y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \dots + \beta_{p-1} X_{p-1,t} + \varepsilon_t \quad (4.1)$$

I denne likningen har vi p parametere, β , og $p-1$ forklaringsvariable X . Den første parameteren angir konstantleddet i modellen. De øvrige parametrene er relatert til de $p-1$ forklaringsvariable som er trukket inn i modellen. Anta nå at vi har n observasjoner. Da vil vi ha n slike likninger for $t = 1, t = 2, \dots, t = n$ som er angitt i (4.1) Disse kunne vi plassere rett under hverandre. Dette likningssystemet med n likninger kan da skrives på vektorform ved:

$$y = X\beta + \varepsilon \quad (4.2)$$

Her er y en n -dimensjonal vektor som består av tidsrekken (observasjonene), mens X er en $(n \times p)$ dimensjonal matrise som består av de respektive verdier på forklaringsvariablene i hvert av de n tidspunktene. β er en p -dimensjonal vektor som består av parametrene, og ε er n -dimensjonal vektor med støyleddene.

En konsekvens av kravet (3.1) om at forventningen til støyleddet skal være lik 0 ($E \varepsilon = 0$) er:

$$E y = X\beta \quad (4.3)$$

En konsekvens av kravene (3.2) og (3.3) er

$$\text{Cov } y = \text{Cov } \varepsilon = \sigma^2 I \quad (4.4)$$

der I er en $(n \times n)$ dimensjonal diagonalmatrise med bare 1-ere på diagonalen og 0 ellers i matrisen.

I multipl regressjon er da regresjonsmodellen gitt ved matriselikningen (4.2), mens kravene til modellen er gitt ved likning (4.3) og (4.4). I tillegg kreves det at ε er Normalfordelt.

Minste kvadraters estimatorene finnes ved å minimere uttrykket:

$$Q = \sum_{t=1}^n (y_t - E y_t)^2 \quad (4.5)$$

Her er y_t observasjonene og $E y_t$ er forventet observasjonsverdi gitt ved likning (4.3) eller (4.1) når $\varepsilon_t = 0$. Minste kvadraters estimatorene vil være de verdier på β som minimerer likning (4.5). I likning (4.5) inngår det p parametere. For å finne de optimale verdier på β må likning (4.5) deriveres med hensyn på hver enkelt β og settes lik 0. Vi vil da ha et lineært likningssystem med p likninger og p ukjente parametere. Løsningen av dette likningssystemet gir minste kvadraters estimatorene. På matriseform kan likning (4.5) uttrykkes på følgende måte:

$$Q = (y - X\beta)'(y - X\beta) \quad (4.6)$$

Derivasjon av dette uttrykket med hensyn på β og så sette det deriverte uttrykket lik 0 gir:

$$0 = -2y'X + 2\beta X'X$$

Av dette uttrykket følger minste kvadraters estimatorene på vektorform:

$$\beta = (X'X)^{-1}X'y \quad (4.7)$$

Selv om dette uttrykket på matriseform ser forholdsvis enkelt ut, er det svært komplisert å løse. I [13] er det angitt løsninger i en regresjonsmodell med en forklaringsvariabel. Der går det frem at selv disse løsningene krever en god del regnearbeid. Det er først nå når vi har fått elektroniske regnemaskiner at vi har fått mulighet til enkelt å løse regresjonsmodeller med mange parametere. Regnearbeidet går i hovedsak ut på å beregne den inverse til matrisen $(X'X)$.

Det kan her oppstå spesielle problemer under estimeringen (beregningene) hvis noen av kolonnene i X tilnærmet er lineært avhengig av hverandre. Da vil determinanten til $(X'X)$ nærme seg 0, og vi vil få stor instabilitet i beregningene. Dette betyr at en ikke bør inkludere forklaringsvariable i modellen som er sterkt avhengig av hverandre. Dette problemet betegnes i regresjonsanalyse som multikolinearitet. Dersom dette problemet oppstår, kan en omgå det ved f eks å benytte "Ridge regresjon" der det adderes til en diagonalmatrise under estimeringen for å unngå instabiliteten, se kapittel 5.6.

5 Tester for å evaluere aktuelle modeller

Anvendelse av testene vil vi se nærmere på i kapittel 7 og 8 i tilknytning til modellbyggingen. De tester og kriterier som spesielt er viktig i regresjonsanalysen, er

- Multipl regressjonskoeffisient
- Signifikanstest på hver enkelt parameter
- Standardavviket til residualene
- Durbin-Watson test
- Autokorrelasjonsfunksjon
- Korrelasjonsmatrise for parametrene
- Prediksjons- og konfidensintervall.

I tillegg vil forskjellige plott av residualene være viktig både for å vurdere kvaliteten på modellen og for å få underlag for

forbedringer av modellen. Plott på tidsaksen av residualene er et svært viktig plott. Dette vil i mange fall angi strukturer som ikke er lagt inn i modellen. Se [12]. Normalfordelingsplottet er også viktig. Dette viser om støyleddet tilfredsstillende en Normalfordeling. Et variasjonsgjennomsnittsplott kan også være nyttig: Det viser i hvilken grad variansen/standardavviket til residualene øker med gjennomsnittsverdien. Ifølge forutsetningen skal vi ikke ha noen økning her. Se [11] hvor dette er anvendt. I mange tilfeller vil også plott av residualene som funksjon av tiden kunne avdekke dette.

5.1 Multipel korrelasjonskoeffisient

La oss se på tilfellet for regresjon med en forklaringfaktor. Bruker vi y som avhengig variabel og x som forklaringsvariabel, er den lineære modellen gitt ved

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad (5.1)$$

Vi kaller så de estimerte verdier for y_t for

\hat{y}_t , og disse er de verdiene vi får for y_t når vi setter inn de estimerte

verdier for α og β , dvs $\hat{\alpha}$ og $\hat{\beta}$, og den verdi vi har for forklaringsvariabelen.

$$\hat{y}_t = \hat{\alpha} + \hat{\beta} x_t \quad (5.2)$$

De realiserte verdier av støyen, dvs residualene, kaller vi e_t . Da har vi at

$$y_t = \hat{y}_t + e_t \quad (5.3)$$

De observerte y_t kan altså splittes i en modellforklart del \hat{y}_t og en uforklart del e_t . La oss så trekke fra gjennomsnittet av y -verdiene, \bar{y} , på hver side av likhetstegnet. Da får vi

$$(y_t - \bar{y}) = (\hat{y}_t - \bar{y}) + e_t \quad (5.4)$$

og dette uttrykker den samme oppdelingen, men nå som avvik fra gjennomsnittet, siden gjennomsnittet til de observerte y -ene og de estimerte y -ene er like (det kan vises) og gjennomsnittet av residualene er lik 0.

Vi tar nå kvadratet på begge sider og summerer over alle n observasjonene. Da får vi

$$\sum_{t=1}^n (y_t - \bar{y})^2 = \sum_{t=1}^n (\hat{y}_t - \bar{y})^2 + \sum_{t=1}^n e_t^2 \quad (5.5)$$

siden det kan vises (vi gjør ikke det her) at kryssproduktet mellom $(y_t - \bar{y})$ og e_t er lik null. (5.5) kan med ord uttrykkes som:

$$\begin{aligned} & \text{Total variasjon} \\ &= \text{Forklart variasjon} \\ &+ \text{Uforklart variasjon} \end{aligned}$$

Hvis vi dividerer (5.5) på begge sider med kvadratsummen på venstre side får vi:

$$\begin{aligned} & \frac{\sum_{t=1}^n (y_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \\ &= \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} + \frac{\sum_{t=1}^n e_t^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \end{aligned} \quad (5.6)$$

Her er det første ledd på høyre side i (5.6) som beskriver den forklarte variasjonen, og dette leddet blir kalt R^2 :

$$R^2 = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (5.7)$$

Den multiple korrelasjonskoeffisienten, R , er definert som kvadratroten til R^2 .

Utleddningen ovenfor lar seg lett utvide til det generelle tilfellet med flere forklaringsfaktorer.

Vi ser at venstre side i (5.6) er lik 1. Altså må R^2 alltid være mindre enn eller lik 1. Altså ligger R mellom -1 og 1. $R^2 = 1$ betyr at vi har perfekt tilpasning til data mens $R^2 = 0$ betyr at vi ikke har noen korrelasjon mellom modell og data. Er $R = 0,85$ betyr dette at 85 % av variasjonen i y_t er forklart ved hjelp av modellen.

I tillegg er det to egenskaper ved den multiple korrelasjonskoeffisienten som er viktige å huske:

1. Koeffisienten har ingen klar tolkning hvis konstantleddet utelates.
2. Koeffisienten har en tendens til å stige i verdi hvis vi øker antall forklaringsvariable.

5.2 Signifikanstest på hver enkelt parameter (t-verdi og/eller sannsynlighetsnivå)

Det er i utgangspunktet ønskelig at alle parameterestimerer er statistisk signifikante. En vanlig måte å teste dette på er ved de såkalte *t*-verdier, som er verdien på parameterestimatet dividert med det estimerte standardavvik. *En tommelfingerregel for den såkalte t-testen sier da:*

Aksepter hypotesen om at parameterestimatet er signifikant forskjellig fra 0 hvis t-verdien er større enn 2.

Teoretisk vil en test med 95 %-signifikansnivå (oppgis også ofte som 5 % signifikansnivå), dvs at den vil være riktig i 95 av 100 tilfeller, ha grenseverdien 2,06 ved 25 observasjoner. Når antall observasjoner går mot uendelig, avtar denne grenseverdien mot 1,96.

En annen måte å se denne testen på er å ta utgangspunkt i parameterens estimerte verdi og så trekke fra og legge til 2 ganger det estimerte standardavviket. Hvis det intervallet som da fremkommer ikke dekker verdien 0, er estimatet signifikant forskjellig fra 0.

5.3 Standardavviket til residualene

Standardavviket til residualene måler usikkerheten i den uforklarte delen av modellen. Denne usikkerheten påvirker også usikkerheten i de prognoser som modellen generer.

I utgangspunktet søker vi den modellen som har minst usikkerhet. Dermed blir standardavviket til residualene også et viktig mål for sammenlikning av ulike forslag til modell, selv om dette i praksis bare blir ett av flere mål.

5.4 Durbin-Watson test

En av de fire sentrale forutsetningene for støyleddene i en regresjon er at det ikke skal være noen avhengighet mellom de ulike støyleddene, se forutsetning iii) i kapittel 3. Hvis det er en slik avhengighet, sier vi at støyleddene er autokorrelerte. Dette medfører at alle varianser for parameterestimatene (og kovarianser mellom disse) undervurderes, noe som leder til overoptimistiske *t*-verdier og multiple korrelasjonskoeffisienter.

En observator som kan brukes til å teste hvorvidt det er autokorrelasjon i påfølgende residualer, også kalt 1. ordens autokorrelasjon, er den såkalte Durbin-Watson observatoren. Den er gitt ved

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (5.8)$$

For at vi skal kunne konkludere med at residualene ikke er autokorrelerte sier testen at d ikke skal ligge så langt unna 2. Imidlertid er det et problem at testen har

områder som ikke gir noen konklusjon. Dette ses av figur 5.1.

I figuren er d_N nedre grense mens d_0 er øvre grense. Grenseverdiene avhenger av hvor mange parametere som er estimert og hvor mange observasjoner vi har, og finnes i tabeller som ikke er gjengitt her. Se f eks [14] for en slik tabell.

5.5 Autokorrelasjonsfunksjonen

I de tilfellene vi ikke får noen konklusjon fra testen med Durbin-Watson observatoren, eller vi har mistanke om autokorrelasjon av høyere orden, dvs avhengighet mellom residualer med større avstand enn en, kan vi se på autokorrelasjonsfunksjonen, som måler korrelasjonen mellom residualene i så stor avstand som vi måtte ønske.

Den generelle autokorrelasjonsfunksjonen i avstand k , r_k , er definert som

$$r_k = \frac{\sum_{t=1}^n (e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2} \quad (5.9)$$

Autokorrelasjonsfunksjonen er en normalisert variabel som antar verdier mellom -1 og 1.

r_k testes mot 2 ganger standardavviket, σ , gitt ved (den forenklede formelen)

$$\sigma = \frac{1}{\sqrt{n}} \quad (5.10)$$

Verdier av r_k som er høyere enn 2σ anses som signifikante.

Se for øvrig [12] og [11] for en mer detaljert beskrivelse av autokorrelasjonsfunksjonen.

En test som sjekker hvorvidt det er autokorrelasjon i alle autokorrelasjonsfunksjonene samlet er basert på *Kji-kvadrat observatoren*:

$$Q = n \sum_{k=1}^k r^2(k) \quad (5.11)$$

hvor

k = antall autokorrelasjoner, dvs antall lag

n = antall støyledde

$r(k)$ = autokorrelasjonen for avstand k mellom støyleddene.

En tommelfingerregel gir følgende test:

Hvis Q er omtrent like stor som antall frihetsgrader, eller lavere, er det ingen indikasjoner på avhengighet i støyleddene.

Antall frihetsgrader er her gitt ved antall autokorrelasjoner som beregnes minus antall parametere som er estimert i modellen.

5.6 Korrelasjonsmatrise for parametrene

Det er ønskelig, men slett ikke alltid mulig, å oppnå estimerte parametere som ikke er avhengige av hverandre. Hvis noen av forklaringsvariablene utvikler seg tilnærmet likt sier vi at vi har *multikolinearitet*, og da har estimeringsmetoden minste kvadraters metode problemer med å skille effektene av de ulike forklaringsfaktorene. Resultatet er parameterestimater med svært stor usikkerhet.

Det klassiske tegnet på multikolinearitet er store standardavvik på parameterestimater kombinert med høy verdi på den multiple korrelasjonskoeffisienten. Dette betyr at forklaringsvariablene er høyt korrelert med den avhengige variabelen, men at vi ikke kan si hvilken. Hvis vi finner at en av to forklaringsvariable er signifikant når vi utelater den andre, men at begge er ikke-signifikante når begge er inkludert i regresjonslikningen, er dette også et tegn på multikolinearitet.

Løsninger på dette problemet kan være å øke antall observasjoner (hvis det er mulig) eller å utelate en av forklaringsfaktorene. "Ridge-regresjon" er en metode som korrigerer for høy grad av multikolinearitet. Vi går imidlertid ikke nærmere inn på dette her.

Korrelasjonsmatrisen er et hjelpemiddel som kan hjelpe oss å avgjøre om multikolinearitetsproblemer er til stede. Den viser korrelasjonen, målt ved kovariansen mellom parameterestimaterne.

5.7 Prediksjons- og konfidensintervall

Et *konfidensintervall* sier noe om hvor sikker beskrivelsen av den historiske utvikling er, mens et *prediksjonsintervall* sier noe om hvor sikker prognosen er.

Vi skal her konsentrere oss om usikkerheten i prognosene. Her er det to kilder til usikkerhet:

1. Alle β -parametrene i regresjonslikningen er estimerte, dvs vi kjenner ikke de eksakte verdier for disse. Imidlertid er det slik at jo flere observasjoner vi har, desto sikrere er de estimatene vi bruker.
2. I tillegg har vi i prognoseperioden en usikkerhet på grunn av støyledet. Selv om vi kjente de eksakte verdier for β -parametrene ville vi ikke være i stand til å gi perfekte prognoser for y_t pga støyledet.

I tilfellet med en forklaringsvariabel, x_p , kan det vises at den totale usikkerheten i prognosene, dvs summen av usikkerhetene i 1. og 2. ovenfor, gitt ved variansen til prognosene, s_p^2 , kan skrives som:

$$s_p^2 = s^2 \left(1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum_{t=1}^n (x_t - \bar{x})^2} \right) \quad (5.12)$$

Her er $s^2 = \frac{1}{n-2} \sum_{t=1}^n (y_t - \hat{y}_t)^2$,

dvs variansen til støyledet

n = antall observasjoner

\hat{y}_t = estimert (modelltilpasset) verdi av

y_t

p = prognoseperiode

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t$$

Positiv autokorrelasjon	Ingen konklusjon	Ingen autokorrelasjon	Ingen konklusjon	Negativ autokorrelasjon
0	d_N	d_0	2	$4-d_0$
			$4-d_N$	4

Figur 5.1 Grenser for test med Durbin-Watson observatoren

Vi ser av (5.12) at variansen til prognosene avtar når antall observasjoner, n , vokser. Dessuten øker variansen med fremtidige verdier av x_t , dvs x_p , som avviker fra det historiske gjennomsnittet, \bar{x} .

Et 95 % prediksjonsintervall for y_p er gitt ved:

$$I_{95} = (y_p - t_{0,025} * s_p, y_p + t_{0,025} * s_p) \quad (5.13)$$

hvor

y_p = prognose i periode p

$t_{0,025}$ = 0,025-fraktile i Students t -fordeling

s_p = standardavviket til prognosene, gitt ved kvadratrotten til s_p^2 , gitt i (5.12).

Et slikt intervall dekker prognose-verdien for y_t , dvs y_p , med 95 % prosent sannsynlighet. Det vil si at i 95 av 100 tilfeller vil prognosene ligge i dette intervallet.

(5.12) lar seg lett utvide til det generelle tilfellet for multipl regressjon, dvs med flere forklaringsvariable, men dette går vi ikke inn på her.

6 Trinnvis regresjon

Når regresjonsmodeller utvikles, prøver en seg ofte med ulike modellvarianter. For økonomiske tidsrekker kan dette komme av at en har ulike/konkurrerende teorier for hvordan markedet fungerer eller av at en fra statistikker kan velge flere forklaringsvariable som i utgangspunktet kan representere en/ flere av de variable som teorien angir. I begge tilfeller forsøker en å komme frem til den modellen som gir den beste forklaringen på historisk utvikling og som samtidig gir "fornuftige" prognoser.

Viktige kriterier for valg av modell er nevnt i kapittel 5: Det er i utgangspunktet ønskelig at alle forklaringsfaktorene er signifikante. Dessuten ønsker vi at forklaringsfaktorer som trekkes inn skal bidra til å gi en stadig bedre forklaring av den historiske utvikling. Dette måler vi ved hjelp av den multiple korrelasjonskoeffisienten. Imidlertid er det slik at den multiple korrelasjonskoeffisienten vokser uansett når vi inkluderer en ny forklaringsvariabel, og det uavhengig av om denne ikke bidrar til å forklare mer av utviklingen.

I mange statistikkpakker er det derfor inkludert en metode som kan hjelpe oss i

å avgjøre hvorvidt en forklaringsvariabel bør inkluderes i modellen. Denne metoden kalles *trinnvis regresjon*. Dette er en regresjonsanalyse som inneholder kriterier for hvilke forklaringsvariable som skal velges ut og hva som skal til for at de skal innlemmes i regresjonslikningen.

Det kan settes opp ulike kriterier for valg og innlemmelse av forklaringsvariable. Vi skal ikke gå inn på dette i detalj. Prinsippet er imidlertid at en forklaringsfaktor må bidra til signifikant forklaring av utviklingen i den avhengige variabelen ut over det de allerede inkluderte forklaringsfaktorer gjør.

I kapittel 7 og 8 lar vi programpakken Statgraphics utføre trinnvis regresjon med de data som er beskrevet i kapittel 2. Her er det en F-verdi som avgjør om den aktuelle forklaringsvariabel skal innlemmes i modellen. F-verdien gir uttrykk for en grenseverdi for en såkalt partiell F-test, som tester hvorvidt en forklaringsvariabel har signifikant forklaringskraft i forhold til den avhengige variabelen. Har den det, innlemmes den i modellen. Hvis ikke, holdes den utenfor. Er det flere aktuelle forklaringsvariable å velge mellom, velges den som har høyest F-verdi. Deretter søker metoden etter den neste variabel som har størst forklaringskraft i form av F-verdi og innlemmer denne, osv. Dette skjer trinnvis inntil de forklaringsvariablene som står igjen har en F-verdi som er lavere enn en på forhånd spesifisert grenseverdi. Disse inkluderes ikke i modellen. Samtidig tester også metoden alle de forklaringsvariablene som er inkludert i modellen som om hver enkelt var den sist inkluderte: Skulle signifikansnivået som følge av at nye variable er trukket inn i modellen endre seg og falle under den på forhånd spesifiserte F-verdi, utelates disse.

Det er i statistikkpakken Statgraphics satt en default-verdi for denne F-verdien på 4,0, og den benyttes i kapittel 7. Hvis de partielle F-testene som utføres skal ha 95 % signifikansnivå er nok dette en litt for lav verdi, fordi det i forhold til F-fordelingen krever langt flere observasjoner enn vi har til rådighet. Med et slikt krav til signifikansnivå burde F-verdien i henhold til F-fordelingen ha ligget mellom 5,99 og 4,45, avhengig av hvor mange forklaringsvariable vi til enhver tid har inkludert i modellen. Bruk av default-verdien vil altså medføre at vi risikerer å inkludere flere forklaringsvariable enn det vi strengt tatt burde. Imidlertid er det prinsippet for metoden som er det viktige

i denne sammenhengen, og vi velger derfor å bruke statistikkpakkens default-verdi for F-verdien.

7 Lineær prognosemodell

7.1 Utgangspunkt for regresjonsanalysen

Den lineære modellen er den mest brukte i regresjonsanalysen. Den forutsetter at de enkelte forklaringsvariablene påvirker den variabelen vi skal lage prognoser for – i dette tilfelle tellerskritt – autonomt, hvert med sitt bidrag.

Likning (2.1) viser dette ved at det står et plusstegn i likningen mellom bidraget til hver forklaringsvariabel. Det er denne likningen som er utgangspunktet for vår analyse. De forklaringsvariablene vi vil undersøke om bør inkluderes i vår modell, er:

- Realprisen (P_t)
- Realpris med ett kvartals forsinkelse (P_{t-1})
- Realpris med to kvartalers forsinkelse (P_{t-2})
- Realpris med tre kvartalers forsinkelse (P_{t-3})
- Realpris med fire kvartalers forsinkelse (P_{t-4})
- Antall abonnement (AB_t)
- Antall virkedager ($Vdag_t$)
- Tiden (T_t)
- Sesongutslag 2. kvartal i forhold til 1. kvartal (S_{2t})
- Sesongutslag 3. kvartal i forhold til 1. kvartal (S_{3t})
- Sesongutslag 4. kvartal i forhold til 1. kvartal (S_{4t}).

7.2 Regresjonsanalysen

Til sammen har vi her 11 forklaringsvariable. Sammen med konstanten β_0 har vi potensielt 12 parametere i modellen. Dette er som før sagt egentlig langt i overkant når vi kun har 18 kvartalsvise observasjoner, men analysen vil muligens forkaste noen av forklaringsfaktorene.

For å bestemme hvilke forklaringsfaktorer som skal inkluderes, benytter vi en trinnvis regresjon. Forklaringsvariablene tas da inn i modellen i prioritert rekkefølge avhengig av hvor mye de forklarer. Prosedyren stopper slik at variablene som har et signifikansnivå på mindre enn ca 5 % ikke tas med i modellen.

Følgende variable ble i prioritert rekkefølge tatt med i modellen:

- Antall abonnement (AB_t)
- Sesongutslag 3. kvartal ($S3_t$)
- Antall virkedager ($Vdag_t$)
- Sesongutslag 4. kvartal ($S4_t$)

De øvrige forklaringsvariablene gav i denne modellen ikke signifikant forklaring og inkluderes derfor ikke i modellen.

Eliminerer vi de variablene som ikke gav signifikant forklaring i (2.1) og inkluderer et støyledd, ε_t , med egenskaper som beskrevet i kapittel 3, blir modellen:

$$TS_t = \beta_0 + \beta_2 AB_t + \beta_4 Vdag_t + \beta_{53} S3_t + \beta_{54} S4_t + \varepsilon_t$$

Resultatene av parameter-estimeringene er gitt i tabell 7.1.

Tabellen viser med unntak av sesong4, svært sterke t-verdier, gitt i parentesene under hvert estimat. Resultatene er heller ikke overraskende. Vi ser av figur 2.2 at tidsrekken klart er sesongbetont. Dette har vi fått inkludert i modellen. Tidsrekken er voksende. Det samme er antall abonnement i figur 2.4. Det er da denne som forklarer veksten i tidsrekken. I tillegg nyanseres utviklingen ved at antall virkedager varierer fra kvartal til kvartal, se figur 2.5. Alle estimatene har dessuten "riktig" fortegn i forhold til det vi forventet i kapittel 2.1.

Konklusjonen er at abonnementsutviklingen har en sterk innvirkning på tellerskrittutviklingen, noe som absolutt ikke er uventet. Økning av antall abonnement gir økt tellerskritt volum (positiv parameterverdi). I tillegg er sesongutslaget i sommerkvartalet (juli, august og september) klart mindre enn nivået på tidsrekken på vinteren (negativ parameterverdi). Også siste kvartal – kvartal 4 – ligger lavere enn tidsrekken nivå på vinteren (negativ parameterverdi). At det er funnet signifikans på antall virkedager gir også god mening: Trafikkvolumet øker med økning i antall virkedager (positiv parameterverdi). På helligdagene

vil stort sett all forretningstrafikk være borte. Det har klar betydning for trafikkvolumet.

7.3 Vurdering av modellen og resultatene

Tabell 7.1 viser at standardavviket til støyen ($SE = 542$) ligger på omtrent 1 % av tidsrekken nivå. Vi må her være klar over at vi ikke opererer med aggregerte observasjoner. Dette er derfor et meget godt resultat. Den multiple korrelasjonskoeffisienten er beregnet til 0,9919. Dette indikerer også en svært god tilpasning. Dette betyr at 99,19 % av tidsrekken variasjoner forklares med forklaringsvariablene i modellen, den funksjonsform og de parametere som er tatt med i modellen. Samtidig vet vi at det ikke er skjevhet i residualene, da gjennomsnittsverdien for dem i regresjonsmodeller er lik 0.

Durbin Watson observatoren (DurbWat) er lik 1,16. Den skal helst ligge rundt 2. Den angir i hvilken grad det er autokorrelasjon i støyleddene i avstand 1 – altså mellom på hverandre følgende støyledd. I dette tilfellet er det ingen konklusjon for Durbin-Watson observatoren, fordi verdien faller i området mellom d_N og d_ϕ , se figur 5.1. Test ved hjelp av autokorrelasjonsfunksjonen gir imidlertid ingen signifikante autokorrelasjoner når vi tester mot to ganger standardavviket, se figur 7.1. Dette indikerer at det ikke er avhengighet mellom støyleddene i avstand en og heller ikke når det er større avstand.

Figur 7.2 viser residualene.

Det ses av figuren at residualene ikke hopper tilfeldig opp og ned over X-aksen. Det er med andre ord mulig det er en avhengighet mellom dem (autokorrelasjon). En test ved hjelp av kji-kvadrat observatoren, se kapittel 5, ville kunne bekrefte eller avkrefte dette. Vi har imidlertid ikke utført denne testen her.

Det kan også være en svak tendens til at residualene er noe større til å begynne med i forhold til i slutten av tidsrekken. Et slikt forhold kan muligens tyde på at standardavviket til støyen avtar noe med

tiden. Vanligvis kan et slikt mønster være omvendt særlig på grunn av at tidsrekken ofte øker i nivå med tiden. Datamaterialet er med 18 observasjoner imidlertid noe spinkelt til å trekke slike konklusjoner.

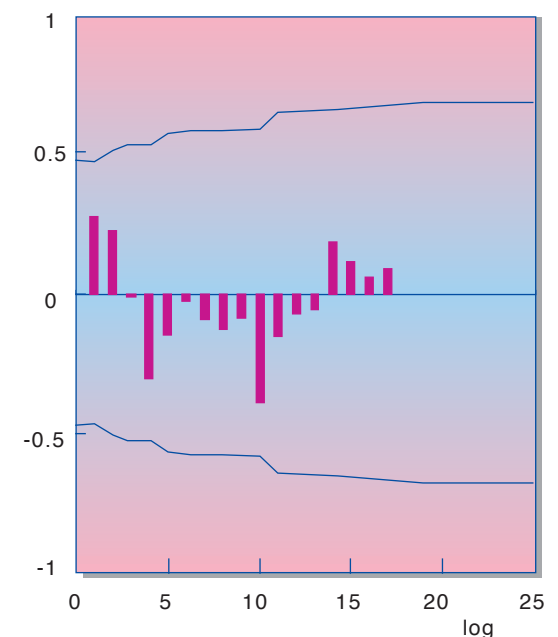
Figur 7.3 viser de aktuelle observasjoner og de tilpassede verdier som modellen gir i de samme punktene. Fremstillingen er ikke sortert etter tiden, men etter størrelse på de tilpassede verdier. Det ses av figuren at vi har en meget god tilpasning når det tas hensyn til at linjen starter ved ca 43 000.

Figur 7.4 viser Normalfordelingsplottet av residualene. Dette plottet er omtalt i [12]. Vi ser av figuren at residualene i høy grad ut fra plottet tilfredsstillende normalfordelingen.

Tabell 7.2 viser korrelasjonsmatrisen mellom de estimerte parametrene i modellen. Det ses av tabellen at det er en forholdvis, men ikke urovekkende, høy korrelasjon mellom estimert konstantledd og estimert parameter for virkedageeffekt. For øvrig er korrelasjonene lavere.

Til slutt vises to figurer av modelltilpasningen. Det er figur 7.5 der de observerte verdier – det vil si tidsrekken – sammenliknes med modellen og de tilpassede verdier. Deretter ser vi på samme type

Estimert autokorrelasjonsfunksjon



Figur 7.1 Autokorrelasjonsplott for residualene. Prikket linje angir grensene for to ganger standardavviket

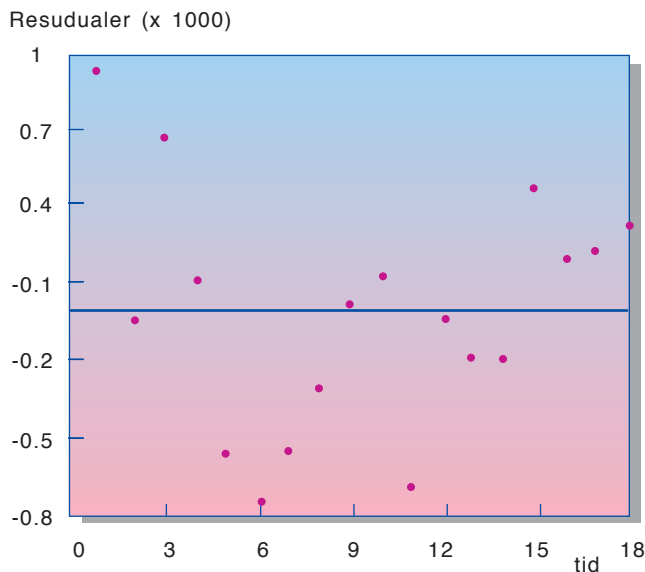
Tabell 7.1 Resultater for lineær prognosemodell

$$TS_t = -145276,3 + 288,4AB_t + 830,1Vdag_t - 7358,6S3_t - 782,3S4_t$$

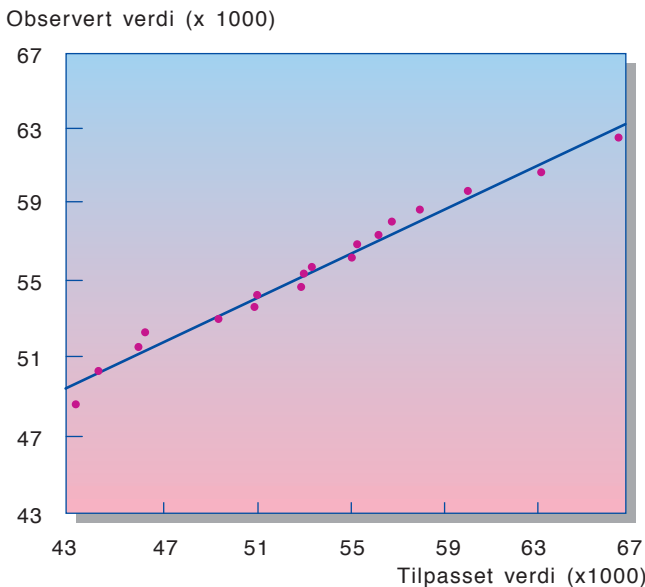
(-21,29)
(41,5)
(9,89)
(-15,2)
(-2,29)

R² = 0,9919 SE = 542,08 DurbWat = 1,16

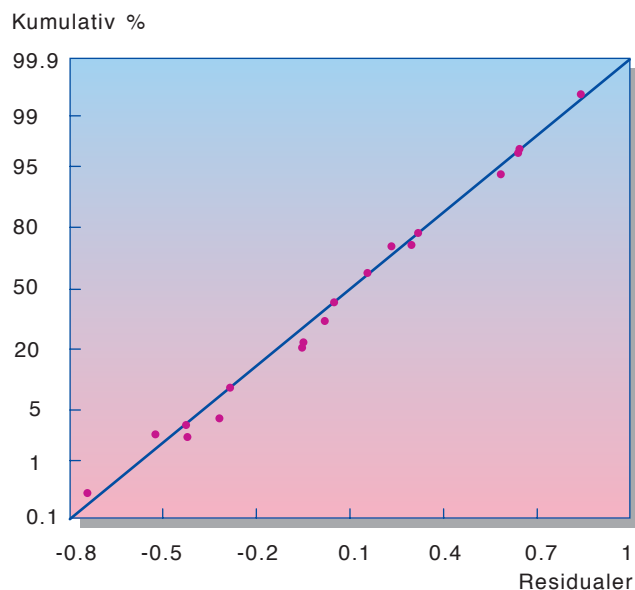
Figur 7.2 Residualer fra lineær modell plottet som funksjon av tiden



Figur 7.3 Observerte og tilpassede verdier fra den lineære modellen



Figur 7.4 Normalfordelingsplott av residualene i den lineære modellen



figur, men hvor det i hvert observasjonspunkt er tegnet inn et 95 % konfidensintervall for tilpasningen. Dette er vist i figur 7.6. Det ses av denne figuren at den relative usikkerheten til modellen er svært liten. Det må her tas i betraktning at tellerskrittaksen starter på 42 000. På grunn av identifikasjon av korrelasjon i støyleddene kan imidlertid konfidensintervallet være noe større enn det som her er satt opp.

Vi ser til slutt på estimatene som modellen gir. Estimert av β_2 er 288,4. Standardavviket (ikke vist i tabell 7.1) er 6,9 og må betraktes som lite. Vi har altså et relativt skarpt estimat for β_2 . Dette betyr at modellen anslår at hvert nytt abonnement gir ca 288 tellerskritt i kvartalet (+/- 14 tellerskritt). Dette er noe lavt, tatt i betraktning at abonnemene både inneholder privat- og forretningsabonnenter.

Estimert for β_4 er 830,1. Det angir hvor mange tellerskritt en ekstra virkedag bidrar med. Dette estimatet synes også svært lavt.

Tilsvarende gjelder også estimatene for β_{53} og β_{54} : Tallverdien til estimatene virker svært lave.

Siden variabelen T_t ikke er funnet signifikant i denne analysen, fikk vi ikke støtte for hypotesen om at nye abonnement genererer færre tellerskritt enn de gamle.

Heller ikke ble noen av prisvariablene funnet signifikante. Dette er ikke uvanlig i analyser hvor det er lite variasjon i prisdata. Her er det imidlertid brukbar variasjon, og vi hadde derfor ventet å finne at realprisene var signifikante.

Vi skal nå se nærmere på dette problemet ved å prøve den andre modelltypen som vi spesifiserte i kapittel 2.2 – den loglineære modellen.

8 Loglineær prognosemodell

8.1 Additiv og multiplikativ effekt av forklaringsvariablene

Vi har så langt sett på modeller med additiv effekt. Det betyr at hver forklaringsvariabel har gitt additive bidrag til modellen. I den modelltypen vi nå skal se på, gir forklaringsvariablene multiplikativ effekt. Det betyr at variablene multipliseres sammen. Vi får da en forsterket effekt av forklaringsvariablene, se kapittel 2.2.

Parametrene i denne modelltypen inngår som eksponenter for de variable. Vi starter med det samme utvalget forklaringsvariable som vist i 7.1.

8.2 Regresjonsanalysen

Det benyttes igjen trinnvis regresjon. Prosedyren tar inn forklaringsvariable i modellen i prioritert rekkefølge avhengig av hvor mye de forklarer, se kapittel 6. Følgende variable ble i prioritert rekkefølge tatt med i modellen:

- Antall abonnement (AB_t)
- Sesongutslag 3. kvartal ($S3_t$)
- Antall virkedager ($Vdag_t$)
- Tiden (T_t)
- Realpris (P_t).

Øvrige forklaringsvariable gav ikke signifikant bidrag til modellen. Nærmest var realpris med 2 kvartalers forsinkelse. Dersom denne variabelen skulle vært tatt inn i modellen, ville vi fått store avhengigheter mellom de to prisvariablene, noe som ville skape problemer under estimeringen av parametrene. Dette ville vi også avdekket i analyse og testene av residualene. Konklusjonen ble altså at den ene realprisvariabelen skulle tas med. Da ble realpris uten forsinkelse tatt med, da denne var den mest signifikante.

Modellen har da følgende form, når vi inkluderer et støyledd, ε_t , med egen-skaper som beskrevet i kapittel 3:

$$TS_t = \beta_0 * (P_t)^{\beta_{10}} (AB_t)^{\beta_2} (T_t)^{\beta_3} (Vdag_t)^{\beta_4} e^{\beta_{53} S3_t + \varepsilon_t} \quad (8.1)$$

En transformasjon over til loglineær form gir følgende modell:

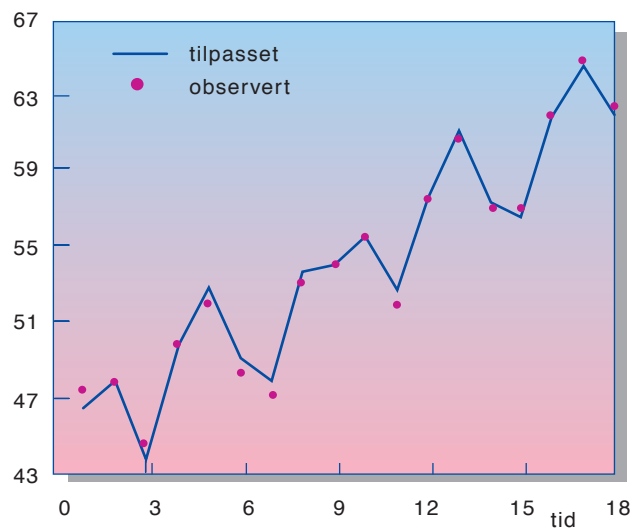
$$\begin{aligned} \ln TS_t = & \ln \beta_0 + \beta_{10} \ln(P_t) \\ & + \beta_2 \ln(AB_t) + \beta_3 \ln(T_t) \\ & + \beta_4 \ln(Vdag_t) + \beta_{53} S3_t \\ & + \varepsilon_t \end{aligned} \quad (8.2)$$

Ser vi bort fra sesongvariabelen som jo er en dummyvariabel som er 1 eller 0 (og støyleddet), er alle øvrige variablene logaritmiske transformasjoner. Regresjonsmodellen er nå transformert over på en lineær form slik at det er mulig å benytte minste kvadraters metode. I den videre analysen er det denne modellen som behandles. Etter at analysen og testene er utført og modellen akseptert, over-

Tabell 7.2 Korrelasjonsmatrise mellom de estimerte parametre i den lineære modellen

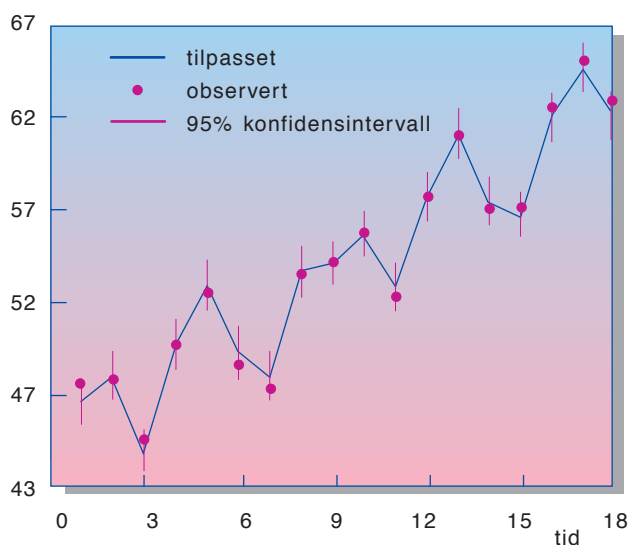
	Konstant	Ab	Vdag	S3	S4
Konstant	1.000	-.5261	-.8627	.60	.2903
Ab	-.5261	1.000	.0244	.03	-.0238
Vdag	-.8627	.0244	1.000	-.74	-.3421
S3	.6070	.0379	-.7469	1.00	.4325
S4	.2903	-.0238	-.3421	.43	1.000

Volum tellerskritt (x 1000)



Figur 7.5 Observerte og tilpassede verdier ved den lineære modellen

Volum tellerskritt (x 1000)



Figur 7.6 Observerte verdier, tilpassede verdier og 95 % konfidensintervall på den lineære modellen

føres resultatene til modell (8.1). For konstantleddet $\ln \beta_0$ betyr dette at den estimerte parameterverdien til konstantleddet skal opphøyres i e (invers til \ln) og

deretter benyttes i modell (8.1). Øvrige parameterverdier er like.

Vi ser så på resultatene av analysen som er gitt i tabell 8.1.

Tabellen viser at alle parameterverdiene til alle variablene med unntak av realpris er meget sterkt signifikante. Realpris er signifikant på et 5 % signifikansnivå. Statistisk sett betyr det 5 % sannsynlighet for feilaktig å inkludere variabelen i modellen. Dessuten har alle parametere "riktig" fortegn i forhold til hva vi forventet i kapittel 2.1.

Vi ser igjen at vi har positive parameterverdier på variable for antall abonnement og antall virkedager og en negativ parameterverdi på parameteren for sommerkvartalet. For øvrig ser vi at vi får en negativ parameter for tiden. Som nevnt i kapittel 2.1 er en rimelig tolkning av dette at nye abonnement ikke er like trafikksterke. Mange av de abonnementene som kommer inn nå vil generere mindre trafikk enn trafikk fra de eksisterende abonnementene. Den negative verdien på parameteren for tid kan derfor tolkes som en trafikkmessig svekking av de abonnement som kommer inn nå.

I denne artikkelen er det ikke foretatt noen dypere analyse av den effekt tele-torg-trafikk og sex-trafikk til Hong Kong, Nederlandske Antiller og Canada har. Dette har ført til en markant vekst i et nytt markedssegment de senere år, samtidig som tellerskrittprisen har falt

betydelig. Vi har altså fått endringer på tilbudssiden i dette markedet, og dette er ikke modellert inn i vår modell. En videre analyse her, f.eks ved å skille ut disse effektene, vil kunne forbedre prognosemodellen.

8.3 Vurdering av modellen og resultatene

Tabell 8.1 viser at standardavviket til støyen (*SE*) er lik 0,0075. Standardavviket i en loglineær modell er identisk med det relative standardavviket i en lineær modell. Det betyr at det relative standardavviket – det vil si standardavviket dividert med tidsrekken – er lik 0,0075 eller 0,75 %. Dette må sies å være meget bra. Det ses av kapittel 7 at den relative usikkerheten for den lineære modellen var ca 1 % ut fra et standardavvik på støyen på 542. Resultatet fra den loglineære modellen er altså bedre.

Den multiple korrelasjonskoeffisient er lik 0,9954 som også er meget bra. *Det betyr som nevnt at de forklaringsvariablene som er inkludert i modellen forklarer 99,54 % av tidsrekkenes variasjon.*

Durbin Watson observatoren (Durb Wat) er lik 1,96. Den skal ligge rundt 2, og testen indikerer da også "ingen autokorrelasjon", se kapittel 5. Durbin Watson observatoren tester imidlertid kun autokorrelasjonen i avstand 1. Vi kan undersøke øvrige avhengigheter i residualene ved å plote autokorrelasjonsfunksjonen slik som vist i figur 8.1.

Det ses av figur 8.1 at det praktisk talt ikke er autokorrelasjon i avstand 1 – altså mellom på hverandre følgende residualer. Alle de beregnede autokorrelasjoner i figuren ligger under usikkerhetsgrensen, slik at vi ikke har noen som er signifikante. Noen er litt høye, men det må tolkes som tilfeldige utslag. Forutsetning om uavhengighet i residualene, gitt i likning (3.3), er da tilfredsstillt. En test ved hjelp av kji-kvadrat observatoren ville muligens gjøre oss enda sikrere på denne konklusjonen. Vi har imidlertid heller ikke her utført denne testen.

Figur 8.2 viser et plott av residualene.

Det ses av figuren at residualene ligger forholdsvis spredt over og under X-aksen. Det er ikke mulig å se sesongmessige variasjoner i residualene til tross for at vi bare har estimert ett sesongledd. Det er heller ikke noen spesiell økning eller reduksjon i residualene som funksjon av tiden. Dermed er forutsetningen om konstant varians til residualene uavhengig av tiden – i likning (3.2) – tilfredsstillt.

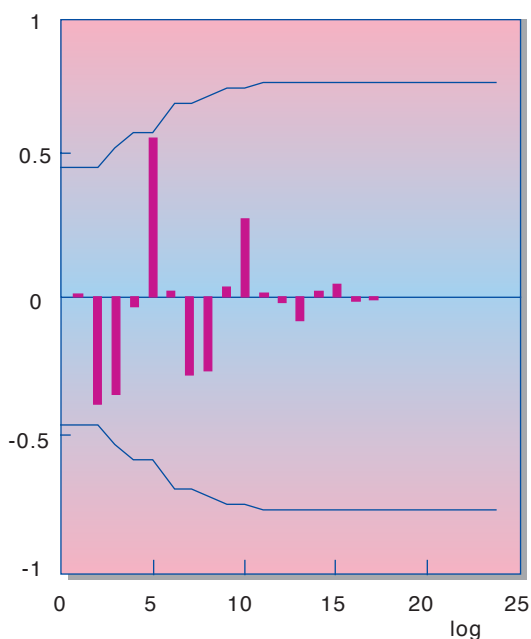
Når minste kvadraters metode benyttes til estimering av parametrene, vil som nevnt gjennomsnittet av residualene bli lik 0. Det vil si at forutsetningen om at den støyen som er tilbake i modellen har forventningsverdi 0, er tilfredsstillt slik som angitt i likning (3.1). Dette gjelder imidlertid for den loglineære modellen som vi nå analyserer. Når loglineære residualene transformeres tilbake til modell (8.1), vil de ha en viss skjevhet. Dette må en imidlertid ta med seg. I totalvurderingen av modellene må en være klar over dette poenget.

Figur 8.3 viser de aktuelle og de tilpassede eller predikerte verdier som modellen gir. Fremstillingen er ikke sortert etter tiden, men etter størrelse på de tilpassede verdier. Det ses av figuren at det er en meget god tilpassing relativt sett.

Figur 8.4 viser Normalfordelingsplottet av residualene. Det ses av figuren at residualene i høy grad tilfredsstiller Normalfordelingen. Dermed er det fjerde og siste krav gitt i kapittel 3 om støyen/residualene tilfredsstillt.

Tabell 8.2 viser korrelasjonsmatrisen mellom de estimerte parametrene i modellen. Det ses at det er en høy korrelasjon mellom de to første parametrene i modellen. Dette er også to av de mest signifikante parametre i modellen med verdier på henholdsvis 7 ganger standardavviket for konstantleddets del og 12 ganger standardavviket for antall abonnement. Eliminering av en av disse variable fører til dårligere tilpassing, og det introduserer også avhengigheter i støyleddstrukturen. Det er derfor valgt å opprettholde samme modelltype.

Estimert autokorrelasjon



Figur 8.1 Autokorrelasjonsfunksjonen for residualene i den loglineære modellen

Tabell 8.1 Resultater for loglineær prognosemodell

$\ln TS_t = -9,99 - 0,09\ln(P_t) + 2,68\ln(AB_t) - 0,03\ln(T_t) + 0,99\ln(Vdag_t) - 0,13S3_t$
(-7,25) (-2,18) (12,13) (-3,93) (13,14) (-20,87)
$R^2 = 0,9954$ $SE = 0,0075$ DurbWat = 1,96

Til slutt vises det to figurer av den loglineære modelltilpasningen. I figur 8.5 vises de observerte verdier som sammenliknes med modellen og de tilpassede verdier. I figur 8.6 vises samme figur, men hvor det i hvert observasjonspunkt er tegnet inn et 95 % konfidensintervall for tilpasningen. Det ses at den relative usikkerhet i den loglineære modellen er svært liten. Dette ses også av variasjonene i forhold til nivået.

La oss så se på de estimerte parametrene, som i tilfellet for alle variablene, bortsett fra for sesongvariabelen, uttrykker elastisiteter. Estimert for β_{10} er $-0,09$. Det betyr således at en økning i realprisen på 1 % gir en nedgang i etterspørselen etter volum tellerskritt på $0,09\%$ i samme periode. Dette kan virke noe lavt, ettersom internasjonale undersøkelser angir korttids priselastisiteter for bedriftsmarkedet i denne størrelsesorden. Korttids priselastisiteter for privatmarkedet angis noe høyere i slike undersøkelser, og vi har i våre data begge typer abonnement. Tar vi hensyn til usikkerheten, målt ved to ganger standardavviket, er den på $\pm 0,08\%$ -poeng. Da virker estimatet mer rimelig.

Tilsvarende angir modellen at 1 % økning i abonnementsmassen i et gitt kvartal medfører $2,68\%$ økning i volum tellerskritt. Dette kan virke rimelig.

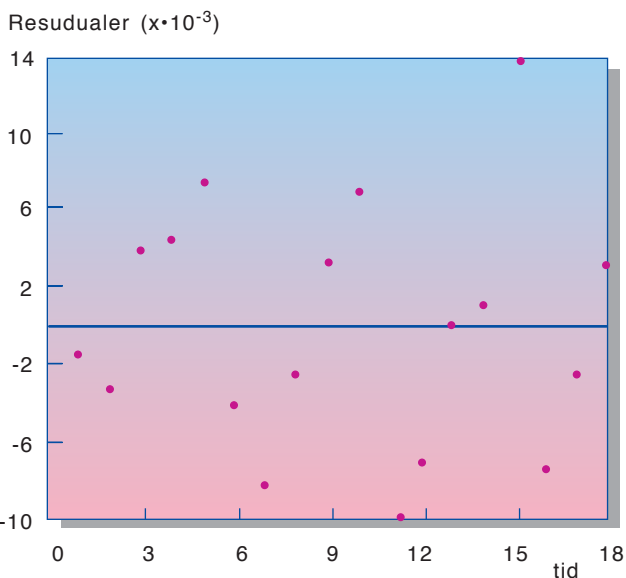
Denne modellen gir støtte til hypotesen om at det er en trend mot lavere antall tellerskritt. I modellen er imidlertid ikke denne trenden kraftig: Hvis vi går ca to perioder frem (1 % økning), medfører dette isolert sett en nedgang på $0,03\%$ i volum tellerskritt. Ut fra vår hypotese skal imidlertid denne variabelen ses i sammenheng med endringen i abonnementsmassen.

Tilsvarende anslår modellen at 1 % økning i virkedagene i et kvartal gir ca 1 % økning ($0,99\%$) i volum tellerskritt dette kvartalet. Dette virker også rimelig.

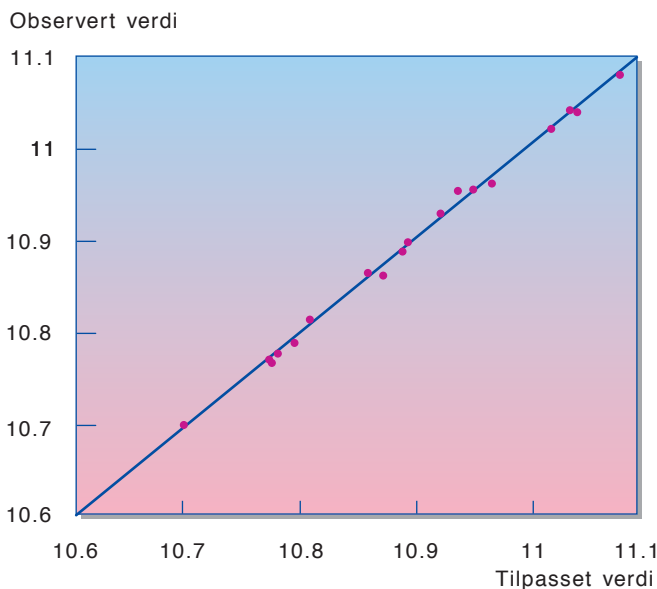
Parameterestimatet for S_{3t} , β_{53} er litt vanskelig å tolke siden denne variabelen ikke er på ln-form. Opphører vi derimot estimatet ($= 0,13$) med e som grunntall, får vi $0,878$, og det er dette tallet som inngår multiplikativt i den ikke-transformerte modell (8.1). Det betyr en reduksjon på $100\% - 87,8\% = 12,2\%$ i 3. kvartal. Dette synes ikke urimelig.

9 Prognoser

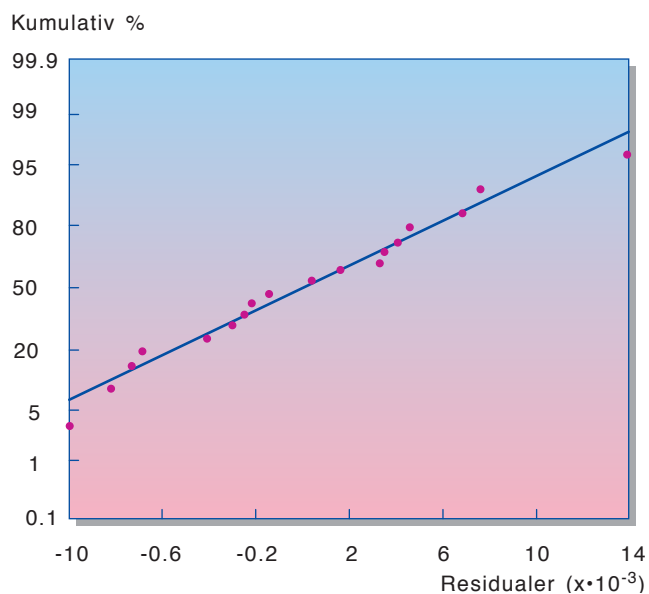
Hvis de to modellene vi har brukt som eksempler her er de to beste modellene vi kan komme frem til, må vi velge en av



Figur 8.2 Residualer fra loglineær modell plottet som funksjon av tiden



Figur 8.3 Observerte og tilpassede verdier fra den loglineære modellen



Figur 8.4 Normalfordelingsplott av residualene i den loglineære modellen

disse til å lage prognoser med. Som regel vil en da se hvor god hver enkelt modell er til å gi en troverdig fremskrivning. Skal vi dømme etter modellenes evne til å beskrive den historiske utvikling, synes i utgangspunktet modellen som benytter ln-verdier å være den beste av de to.

Vi ønsker å komme frem til en modell som vi synes gir en god beskrivelse av den historiske utviklingen og samtidig gir en troverdig fremskrivning. I praksis er det likevel slik at disse forholdene enkelte ganger ikke går overens: Modellen med den beste beskrivelse av den historiske utvikling kan av ulike grunner vise seg å produsere prognoser som ikke er troverdige/ sannsynlige. Da må en velge den "nest beste" modellen, hvis denne gir en mer troverdig prognose.

Hvis modellene tidligere ikke har vært brukt til å lage prognoser med, kan det uansett være lurt å ha en (eller flere) modell-varianter i bakhånd, i tilfelle den

valgte modellen får problemer når nye data kommer til og de reestimerte parametrene viser stor variasjon: Vi ønsker i utgangspunktet at parameter-estimatene skal være stabile, dvs lite påvirkelige av antall observasjoner vi har til rådighet for estimeringen.

Vi skal ikke gå inn på diskusjonen om hvilken av modellene som gir de mest troverdige prognosene her, men velger modellen for ln-transformerte data, dvs modell (8.2), til å lage prognoser for $k = 10$ perioder fremover. Når vi skal lage prognoser kan vi enten beholde modellen på denne formen og så transformere prognoseverdiene for $\ln TS_{t+k}$ tilbake til TS_{t+k} ($k = 1, 2, \dots, 10$) ved å opphøye grunntallet e med ln-verdiene. Eller vi kan velge den ikke-transformerte funksjonsformen (8.1) direkte. Vi velger den siste varianten. Da blir prognosemodellen:

$$TS_{t+k} = 0,0005 * (P_{t+k})^{-0,09} (AB_{t+k})^{2,68} (T_{t+k})^{-0,03} (Vdag_{t+k})^{0,99} e^{-0,13S_{t+k}} \quad (9.1)$$

Her har vi satt støyleddet $\epsilon_{t+k} = 0$, siden dette er støyleddets forventningsverdi.

For å lage prognoser må vi ha prognoseverdier for alle forklaringsvariablene som inngår i modellen. Variablene T og $Vdag$ er i så måte greie: Vi har eksakte verdier som vi setter inn for hver periode. ($Vdag$ kan likevel endre seg hvis vi får informasjon om at vektene for de ulike virkedagene bør endres.)

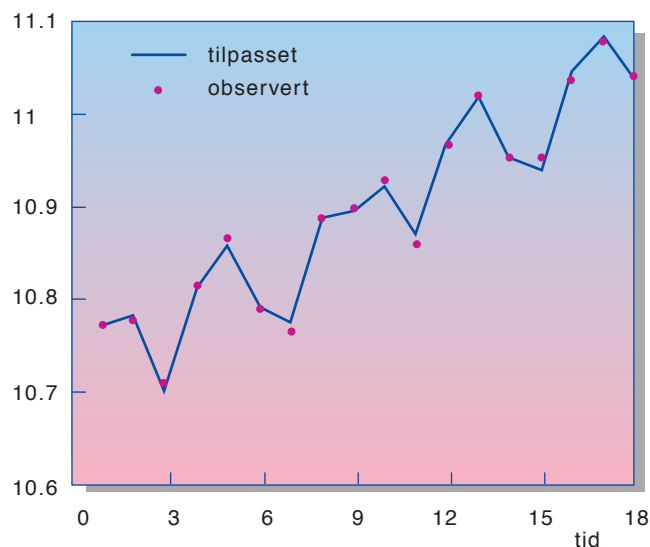
Prognoser for abonnementsutviklingen kan hentes fra prognosemodeller eller offisielle/vedtatte prognoser. Siden abonnementsutviklingen de siste årene har vært svært stabil og har endret seg lite fra år til år, og siden dette er et eksempel, velger vi en "naiv" prognose: Abonnementsutviklingen blir i prognoseperioden den samme som tilsvarende periode året før.

Prognoser for prisutviklingen kan hentes fra budsjettforslag, langtidsplaner eller uttalte prisstrategier, eventuelt fra egenproduserte prognoser/prognosemodeller. I dette eksempelet forutsetter vi at de nominelle prisene senkes med 10 % ved hvert årsskifte. I tillegg endres realprisene ved at konsumprisindeksen vokser med 0,2 % pr måned.

Tabell 8.2 Korrelasjonsmatrise mellom de estimerte parametere i den loglineære modellen

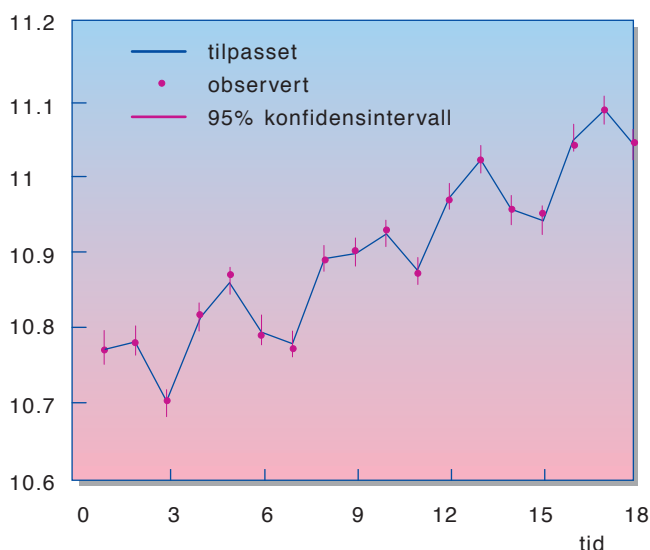
	Konstant	lnAb	lnVdag	lnT	lnP	S3
Konstant	1.000	-.9724	-.2574	.3661	-.7503	.0061
lnAb	-.9724	1.000	.0252	-.3493	.7909	.1550
lnVdag	-.2574	.0252	1.000	-.0832	-.0304	-.6773
lnT	.3661	-.3493	-.0832	1.000	.2582	-.1139
lnP	-.7503	.7909	-.0304	.2582	1.000	.0813
S3	.0061	.1550	-.6773	-.1139	.0813	1.000

log (volum tellerskritt)



Figur 8.5 Observerte og tilpassede verdier ved den loglineære modellen

log (volum tellerskritt)



Figur 8.6 Observerte verdier, tilpassede verdier og 95 % konfidensintervall fra den loglineære modellen

Med disse forutsetningene for forklaringsvariablene er prognoseverdiene for P_{t+k} , Ab_{t+k} , T_{t+k} og $Vdag_{t+k}$ for perioden 3/93 til 4/95 ($k = 1, 2, \dots, 10$) gitt i tabell 9.1.

Setter vi disse verdiene inn i modell (9.1) får vi prognoseverdiene som modellen produserer. Disse er vist i figur 9.1.

I tillegg til prognoseverdiene er det også nyttig å konstruere øvre og nedre 95 % prediksjonsintervall, se formel (5.13). I tilfellet med flere enn en forklaringsfaktor er imidlertid formelen mer komplisert, og vår programpakke beregner ikke dette. Vi har derfor ikke beregnet dette her.

Det bør også nevnes at en vanlig måte å skaffe seg tiltro til modellen er å holde igjen data over en gitt periode, for så å se hvor godt modellen treffer. Dette går vi ikke inn på her, men henviser til [15].

10 Avslutning

Ut fra en helhetsvurdering der det spesielt er lagt vekt på at den relative usikkerheten i den loglineære modellen var mindre enn i de lineære modellen og at avhengighetsstrukturen i residualene var fjernet, er konklusjonen at den loglineære modellen er den beste.

Likning (8.1) angir prognosemodellen. De variable som forklarer tellerskrittutviklingen er:

- Antall abonnement
- Antall virkedager
- Tiden
- Sesongeffekt
- Realpris.

Prognosemodellen gir en usedvanlig god tilpasning med lave usikkerheter i de tilpassede verdier. Årsaken til de gode prognosene ligger ikke bare i modellen i seg selv, men også i forutgående bearbeiding av underlagsdata samt en god regularitet i selve tidsrekken. Se for øvrig [10] og [16].

Det er i analysen ikke gjort forsøk på å skille ut effekter av teletorg og sex-samtaler til Hong Kong, Nederlandske Antiller og Canada. Dette vil kunne gi muligheter for ytterligere å forbedre prognosemodellen.

For brukeren av en slik prognosemodell er det viktig å sette seg inn i hva den forklarer og hva den ikke forklarer. I den innledende analysen i kapittel 2.1 satte vi

opp en rekke aktuelle forklaringsfaktorer. Vi sa imidlertid ikke noe om mulige andre forklaringsfaktorer. En grundig gjennomgang ville avdekket mange slike, hvorav de fleste ikke kan representeres ved hjelp av data. Disse er i vår modell summert opp i støyleddet.

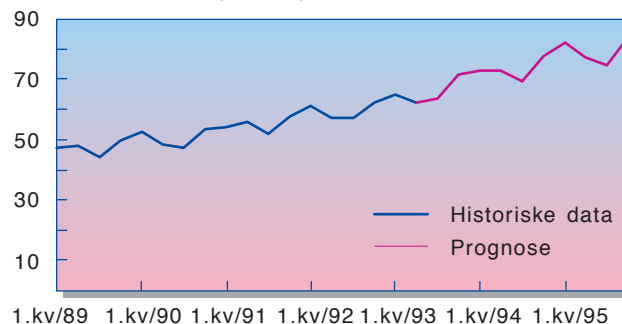
Referanser

- 1 Theil, H. *Principles of econometrics*. New York, J Wiley, 1971.
- 2 Mosteller, F, Tukey, J W. *Data analysis and regression*. London, Addison Wesley, 1977.
- 3 Draper, N R, Smith, H. *Applied regression analyses, 2nd ed*. New York, J Wiley, 1981.
- 4 Wonnacott, T H, Wonnacott, R J. *Regression*. New York, J Wiley, 1981.
- 5 Chow, G C. *Econometrics*. New York, McGraw Hill, 1983.
- 6 Malinvaud, E. *Statistical methods of econometrics, 3rd ed*. Amsterdam, North-Holland, 1980.
- 7 Haavelmo, T. The statistical implications of a system of simultaneous equations. *Econometrica*, 11, 1–12, 1943.
- 8 Haavelmo, T. The probability approach to econometrics. *Econometrica*, 12, Supplements 1–118, 1944.
- 9 Klein, L R. *A textbook of econometrics*. Englewood Cliffs, N.J., Prentice-Hall, 1974.
- 10 Bøe, J, Hjelkrem, C. Tellerskritt: grunnlagsdata og prognoser. *Teletronikk* 90(1), 145–155, 1994 (denne utgave).
- 11 Hjelkrem, C, Bøe, J. Box-Jenkins metode. *Teletronikk* 90(1), 110–124, 1994 (denne utgave).
- 12 Stordahl, K. Prognosemetoder: en oversikt. *Teletronikk* 90(1), 5–52, 1994 (denne utgave).
- 13 Olsen, B T, Stordahl, K. Prognoser for pris på nettkomponenter basert på utvidelse av lærekurvmodellen. *Teletronikk* 90(1), 166–172, 1994 (denne utgave).
- 14 Stewart, M B, Wallis, K F. *Introductory econometrics, 2nd edition*. Oxford, Blackwell, 1981.
- 15 Bøe, J, Hjelkrem, C. Sammenlikning av prognosemodeller fra ulike metoder. *Teletronikk* 90(1), 125–133, 1994 (denne utgave).

Tabell 9.1 Prognoseverdier for forklaringsvariablene

	P	Ab	T	Vdag
3.kv. 93	0,2582	529,4	19	73,8
4.kv. 93	0,2566	534,5	20	70,8
1.kv. 94	0,2295	539,3	21	69,9
2.kv. 94	0,2282	543,0	22	68,65
3.kv. 94	0,2268	545,4	23	73,8
4.kv. 94	0,2255	550,5	24	70,75
1.kv. 95	0,2017	555,3	25	72,45
2.kv. 95	0,2005	559,0	26	66,85
3.kv. 95	0,1993	561,5	27	73,1
4.kv. 95	0,1981	566,5	28	70,4

volum tellerskritt (x 1000)



Figur 9.1 Historiske data og prognoser for volum tellerskritt

- 16 Bøe, J, Stordahl, K. Tellerskrittene:
Televerkets viktigste inntektskilde.
Elektronikk 87(1), 78–92, 1991.

Vekstmodeller

AV JOHANNES BØE OG CARLO HJELKREM

Innledning

I forbindelse med utarbeiding av prognoser finnes det en rekke forskjellige metoder og modeller. Disse kan grovt deles inn i to grupper:

- Framskrivning på grunnlag av historiske data. Slike modeller er det som har vært benyttet i Televerket fram til i dag og vil bli benyttet i framtiden for de fleste tjenester.
- Prognoser som er utledet fra studier av markedets holdninger til produktene/tjenestene. Dette er en vanlig framgangsmåte ved introduksjon av nye tjenester.

Kombinasjoner av disse to hovedgrupper kan også forekomme. Ved å kombinere metodene er det mulig å gi sikrere anslag for framtiden/utviklingen enn en av metodene alene kan gi.

I de tilfeller man ikke har historiske data, må prognosene for etterspørsel etter (nye) tjenester lages på bakgrunn av andre former for erkjennelse, erfaring og informasjon. En slik erkjennelse er at etterspørselen etter nye produkter ofte følger en S-kurve over tid.

Det antas at etterspørselen etter et produkt / en tjeneste følger en voksende eksponentiell utvikling i en tidlig fase for så å følge en avtagende eksponentiell utvikling inntil den har nådd en potensialgrense, for deretter å avta inntil produktet er helt ute av markedet.

Figur 1 viser hvordan totaletterspørselen øker med tiden inntil den når en potensialgrense, for så å avta. Når totaletterspørselen avtar gradvis etter at potensialgrensen er nådd skyldes dette at produktet avløses av nye (og bedre) produkter. Når vi lager prognoser for nye tjenester, er det først og fremst utviklingen de første årene som er viktig. Disse modellen benevnes gjerne som vekstmodeller.

Vekstmodellene er en litt spesiell klasse modeller. Felles for disse er at de i utgangspunktet kun har én forklaringsvariabel, tiden t . Det er mulig å utvide med flere variable, men da har vi beveget oss over i de mer tradisjonelle regresjonsmodeller. Vanligvis inneholder de heller ikke mer enn tre parametre, noe som både er en styrke og en svakhet.

Vekstmodeller benyttes vanligvis til å lage prognoser når datagrunnlaget er spinkelt. Dette spinkle datagrunnlaget vil danne grunnlag for å estimere parametrene i modellen. Det vil imidlertid knytte

seg stor usikkerhet til estimatene når datagrunnlaget er spinkelt.

Man kan si at modellenes styrke ligger i at det kreves lite data i forhold til tradisjonelle modeller. I tillegg har de en evne til å lage gode prognoser for nye tjenester/produkter i en tidlig utviklingsfase. Da er det naturlig at det finnes lite erfaringsmateriale, og det er ikke mulig å benytte seg av historisk utvikling for å estimere sammenhenger i en komplisert modell. Erfaringer har vist at modellene er svært realistiske og passer godt til utviklingen av nye tjenester på et tidlig stadium. Svakheter er at de ofte ikke inkluderer andre forklaringsvariable enn tiden, t . Dette gjør at modellene ikke har samme fleksibilitet og tilpasningsevne som mer konservative og raffinerte modeller.

I figurene 2a–2d er det gitt eksempler på tre typiske vekstmodeller. Disse er

- Eksponentiell modell
- Gompertz modell
- Logistisk modell – S-kurve
- Weibull-modellen.

Figur 2a, 2b, 2c og 2d viser forskjellige vekstmodeller.

Figur 2a viser Eksponentiell modell:

$$y_t = \alpha e^{-\beta t}$$

Figur 2b viser Gompertz modell:

$$y_t = \alpha(e^{\beta})e^{t}$$

Figur 2c viser logistisk modell – S-kurve:

$$y_t = \frac{M}{(1 + e^{(\alpha - \beta t)})}$$

M er markedspotensial.

Figur 2d viser Weibull-modellen:

$$y_t = M - (M - m)e^{\alpha t^{\beta}}$$

M er markedspotensial

m er startnivå.

Det ses klart at alle unntatt den eksponentielle modellen konverterer mot en grenseverdi (metningsgrense). Starten på modellene er alle av eksponentiell natur. Dette kan være en naturlig vekst for en tjeneste eller et produkt i en tidlig fase av utviklingen.

Disse modellene kan med fordel benyttes til langtidsprognoser. Å lage gode langtidsprognoser er svært vanskelig. Mange

av forutsetningene som en langtidsprognose er basert på kan endres over tid, og dermed oppstår det feil i anslagene. Dette er/kan være et argument for å benytte seg av enkle modeller for langtidsprognoser. Dersom det antas at en tjeneste eller et produkt vil ha en metningsgrense – f.eks. lik antall husstander – en grense som kan estimeres ut fra andre kriterier, vil metningsmodellene være en aktuell modellklasse. Metningsgrensen vil da være grenseverdien modellen på sikt grenser opp til.

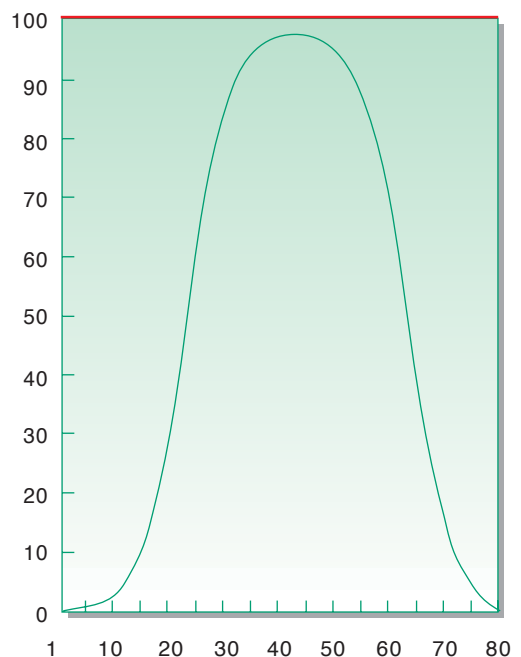
S-kurven

– Logistisk modell

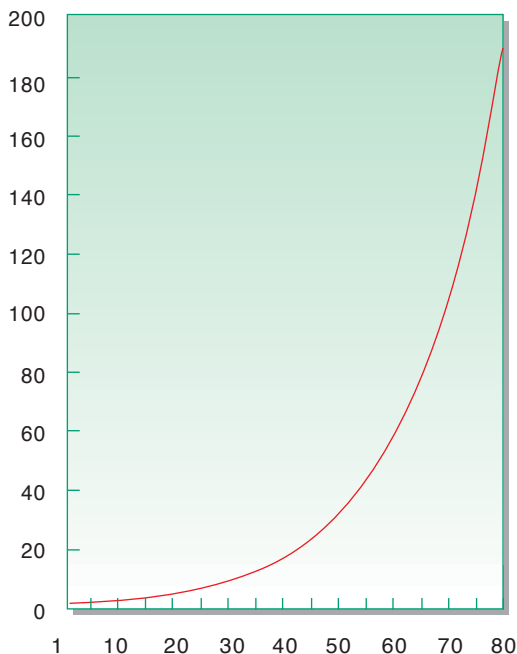
Vi skal nå konsentrere oss om S-kurver som er en videreutvikling av den enkle logistiske modell.

Vi vet at utviklingen av nye tjenester er noenlunde lik på enkelte områder uansett hvilken tjeneste det gjelder. Vi kan gjerne sammenlikne *teledata* med *mobilttelefon* og finne likhetstrekk i disse to så forskjellige tjenester med forskjellig utvikling. Ved å plote etterspørselen til disse to tjenester ser vi at de begge følger en tilnærmet S-kurve/metningsfunksjon. Forskjellen ligger i hvor bratt stigningen er og nivået på kurvene. Dette kan ses av kurvene i figur 3–6.

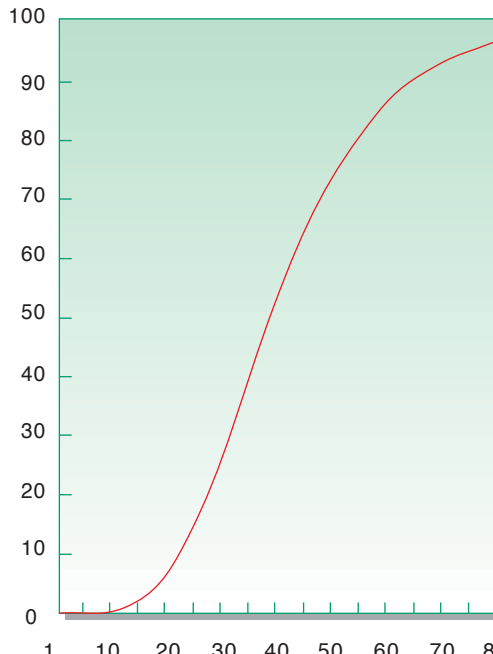
Figur 3 viser utviklingen på mobiltelefon sammenliknet med tilpasset S-kurve.



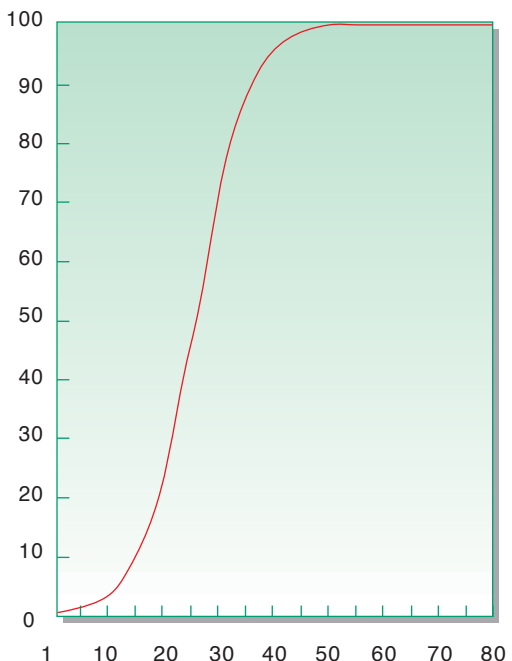
Figur 1 Vanlig utvikling av totaletterspørselen



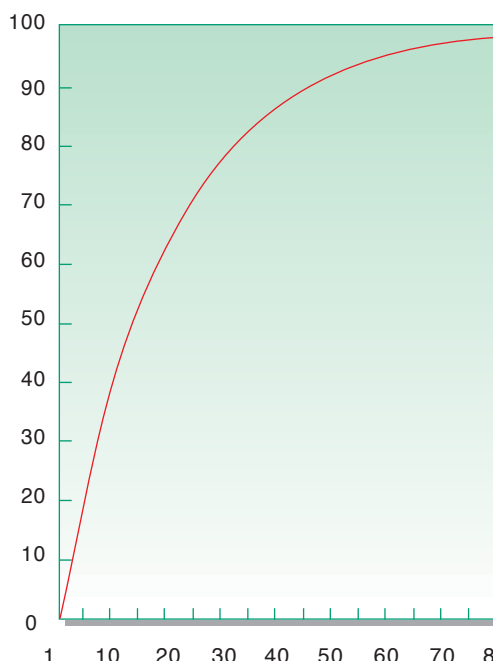
Figur 2a Ekspontsiell modell



Figur 2b Gompertz modell

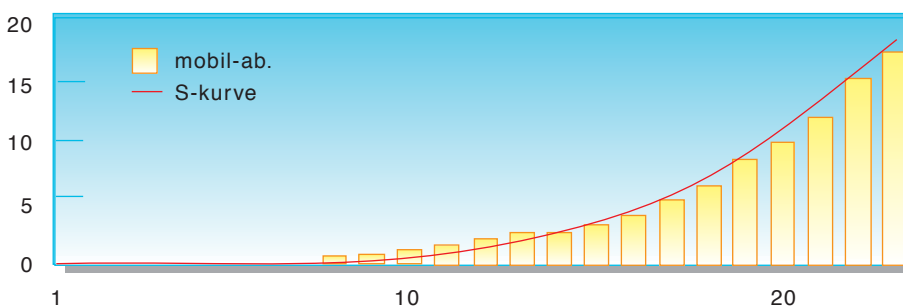


Figur 2c Logistisk modell - S-kurve



Figur 2d Weibull-modellen

(x10 000)



Figur 3 Utviklingen på mobiltelefon sammenliknet med tilpasset S-kurve

Figur 4 viser utviklingen på teleks sammenliknet med tilpasset S-kurve.

Figur 5 viser utviklingen på mobiltelefon sammenliknet med tilpasset S-kurve og forlenget S-kurve.

Figur 6 viser utviklingen på teleks sammenliknet med tilpasset S-kurve og forlenget S-kurve.

Vi beskriver her en S-kurve, det vil si etterspørselsfunksjon med følgende matematisk form.

$$S(t) = \frac{POT}{(1 + e^{(\alpha + \beta * t)})^c}$$

hvor:

- $S(t)$ er prognosen for etterspørselen etter en tjenesten i tidspunkt t

- POT er et uttrykk for markedspotensialet

- t er tidsfaktor med år som måleenhet

- α og β er koeffisienter som estimeres

- c er en fast konstant.

Modellen over består av to deler:

- En S-kurve som angir andel av potensialet som har anskaffet produktet/tjenesten. Variasjonsområde fra 0 til 1

- En beskrivelse av potensialet og eventuelt en utvikling av dette over tid.

En S-kurve kan uttrykkes på formen:

$$S_t = \tanh(t/2) = (e^t - 1)/(e^t + 1)$$

Figur 7 viser S-kurve med variasjonsområde $S = -1$ til $S = 1$.

For å få S-kurven til å variere mellom $S = 0$ og $S = 1$ legges det til 1 og deles på 2 altså:

$$S_t = 1/2\{(e^t - 1)/(e^t + 1) + 1\}$$

$$S_t = 1/2\{(e^t - 1)/(e^t + 1)$$

$$+ (e^t + 1)/(e^t + 1)\}$$

$$S_t = 1/2\{2e^t/(e^t + 1)\}$$

$$S_t = e^t/(e^t + 1)$$

Dette er det samme som

$$S_t = 1/(1 + e^{-t}) \quad (1)$$

Figur 8 viser likning (1) grafisk.

En forskyvning langs t -aksen oppnås ved innføring av parameteren a i likning (1):

$$S_t = 1/(1 + e^{a-t}) \quad (2)$$

Figur 9 viser likning (2) med forskjellige a -verdier. a -verdier større enn 0 fører til forskyvning av kurven mot større t -verdier.

Stigningen i kurven kan endres ved å innføre en parameter b i likning (2)

$$S(t) = \frac{1}{(1 + e^{(a-bt)})} \quad (3)$$

Figur 10 viser hvordan $S(t)$ endres når a holdes konstant og b varierer.

Utvider vi videre likning (3) med en utglidningskonstant c får vi følgende funksjon:

$$S(t) = \frac{1}{(1 + e^{(1-bt)})^c} \quad (4)$$

Av likning (4) ses det at c må være større enn 0.

Modellen varierer mellom 0 og 1.

Figur 11 viser hvordan $S(t)$ endres når a og b holdes konstant og c varierer.

For å få variasjonsområdet til å bli mellom 0 og et markedspotensielt multipliseres likning (4) med markedspotensielt POT .

$$S(t) = \frac{POT}{(1 + e^{(a-bt)})^c}$$

Estimering av parametrene i S-kurven – Logistisk modell

I S-kurven er det vanskelig å benytte minste kvadraters metode for å estimere parametrene direkte. En mulighet er å holde en parameter i tillegg til markedspotensielt fast for så å estimere ved hjelp av regresjonsanalyse.

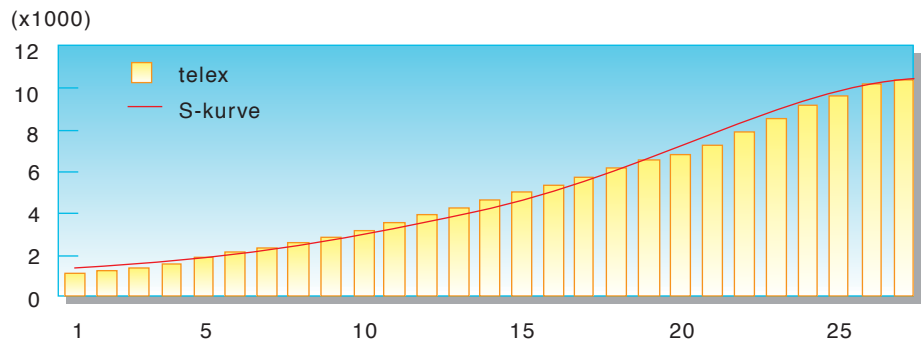
S-kurven (utvidet logistisk modell) består av fire parametre og er gitt ved likning (5):

$$y_t = \frac{M}{(1 + e^{(1-bt)})^c} \quad (5)$$

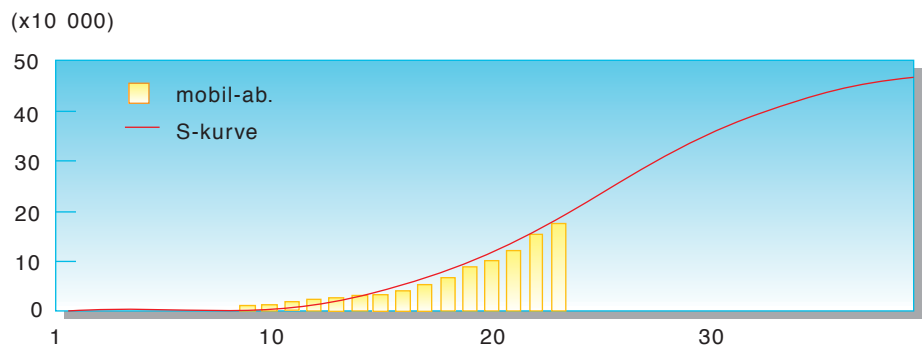
- M - markedspotensielt
- a - parameter som skal estimeres
- b - parameter som skal estimeres
- c - parameter som er en konstant.

Vi skal her tilpasse funksjonen

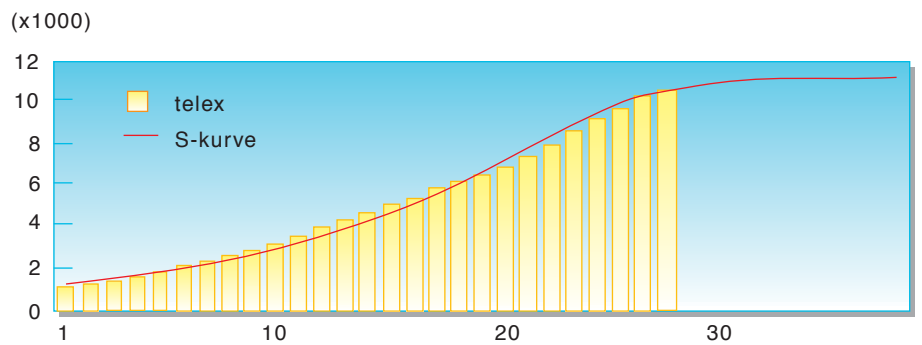
$$y_t = \frac{M}{(1 + e^{(1+bt)})^c} \quad (6)$$



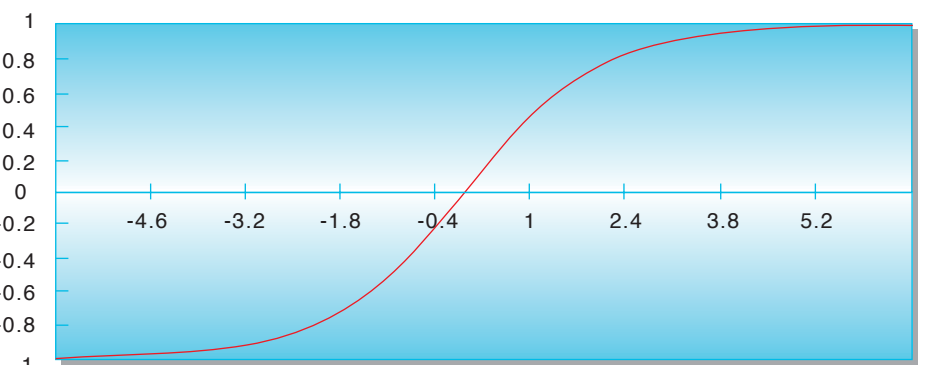
Figur 4 Utviklingen på teleks sammenliknet med tilpasset S-kurve



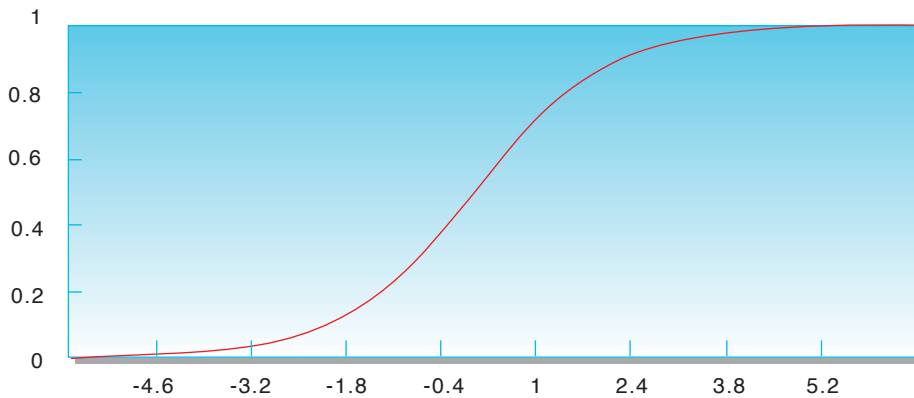
Figur 5 Utviklingen på mobiltelefon sammenliknet med tilpasset S-kurve og forlenget S-kurve



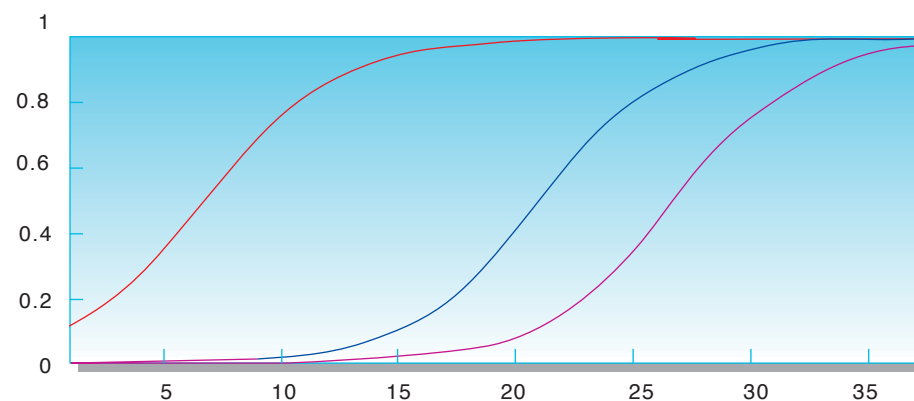
Figur 6 utviklingen på teleks sammenliknet med tilpasset S-kurve og forlenget S-kurve



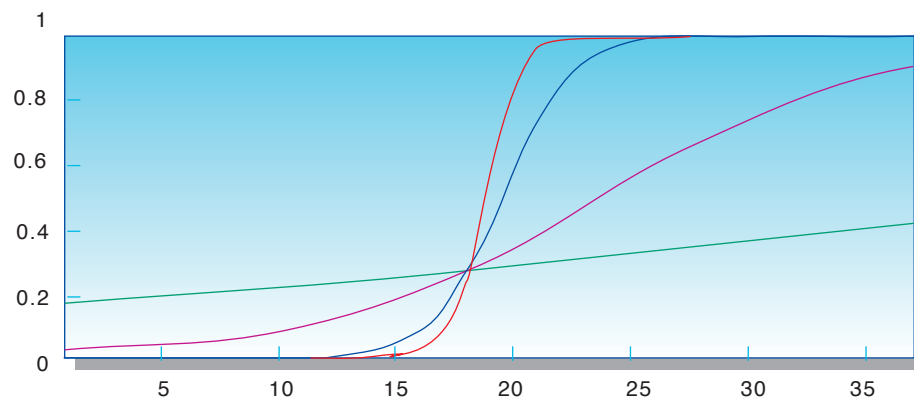
Figur 7 S-kurve med variasjonsområde $S = -1$ til $S = 1$



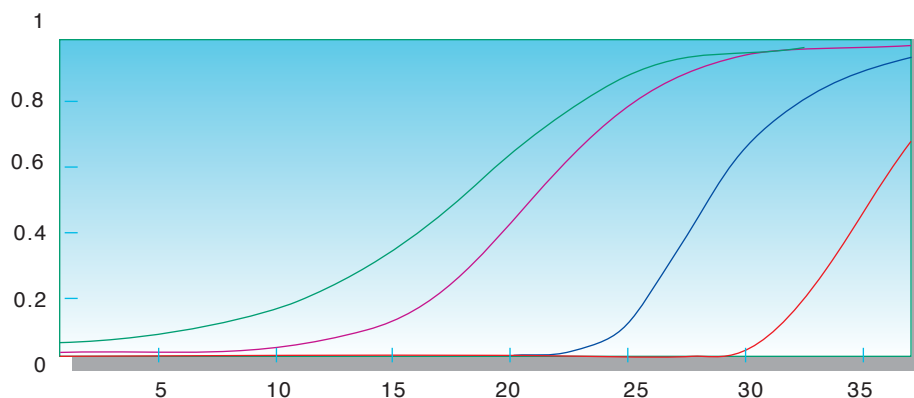
Figur 8 Likning (1) vist grafisk



Figur 9 Likning (2) med forskjellige a -verdier



Figur 10 Endring av $S(t)$ når a holdes konstant og b varierer



Figur 11 Endring av $S(t)$ når a og b holdes konstant og c varierer

slik at det er mulig å benytte lineær regresjon. Dette gjøres ved å transformere y_t over til lineær form y'_t .

Modellen på transformert form blir følgende:

$$LN\{(M/y_t)^{1/c} - 1\} = a + bt$$

$y'_t = a + bt$; hvor $y'_t = LN\{(M/y_t)^{1/c} - 1\}$

Ved å bruke vanlig regresjon estimeres a og b for enhver gitt M og c . Eventuelt at M (markedspotensialet) bestemmes ut fra en naturlig begrensning. Det må her nevnes at M bestemmes utenfor modellen og eventuelt ved hjelp av andre metoder. c er som i (5) satt opp som en konstant. Finnes det historiske data som modellen skal tilpasses bestemmes/estimeres c slik at den totale tilpasningen til data blir best. Hvis det ikke finnes data velges c til en verdi som tidligere erfaringsmessig har passet godt.

For bestemte M og c kan nå parametrene a og b finnes ved minste kvadraters metode.

Forutsetninger som må oppfylles er at det finnes minst to datapunkter/observasjoner. Hvis ikke kan disse to verdiene settes opp som "scenario"-anslag eventuelt at slike "scenario"-punkter settes opp i tillegg til historiske data for øvrig.

Bruk av markedsinformasjon

Det er mulig å analysere markedet for en tjeneste også før tjenesten introduseres. Ved å foreta demografiske og økonomiske vurderinger kan en gjøre seg opp en mening om markedspotensialet for en tjeneste eller produkt.

Eksempelvis kan antall boligabonnemang for telefoni være begrenset til antall husstander i Norge (med eventuell økning når abonnemang nr 2 blir kostnadseffektivt for husstandene). Slike vurderinger er det mulig å gjøre etter markedsundersøkelser. For helt nye tjenester er det viktig å finne abonnentenes innstilling til produkt i kombinasjon med pris. Dermed er det mulig å anslå etterspørselen etter en tjeneste i introduksjonsåret.

Etter hvert som etterspørselsdataene registreres – hvert år – gis det mulighet for å justere og forbedre de første prognosene. Dette er særdeles viktig da de første prognosene bygget på spinkelt grunnlag.

S-kurve benyttet som langtidsprognosemodell

Det interessante med S-kurven er at i tillegg til å beskrive utviklingen til en tjeneste i første fase, gir den gode estimater om en tjenestes siste fase. (I vekstfasen

Tabell 1 viser antall telefontilknytninger i Norge fra 1946 til 1991

År	Telefontilknytninger i 1000
1946	236,7
1947	252,8
1948	267,9
1949	279,4
1950	291,5
1951	304,8
1952	320,1
1953	337
1954	353
1955	368,6
1956	387,3
1957	407,4
1958	421,8
1959	434,7
1960	454,9
1961	486,6
1962	510
1963	530,9
1964	552,9
1965	582,6
1966	607,6
1967	637,1
1968	671,6
1969	707,8
1970	742,3
1971	782
1972	815,5
1973	847,8
1974	879,1
1975	913,9
1976	955,9
1977	1004,5
1978	1054,4
1979	1114,2
1980	1197,3
1981	1297,7
1982	1425,5
1983	1554,2
1984	1657,4
1985	1757,7
1986	1861,4
1987	1948,7
1988	2016,2
1989	2070,2
1990	2132,3
1991	2198,2

bør det benyttes andre modeller eller utvide S-kurven med flere forklaringsvariable.) S-kurven er formulert slik at de skal nærme seg et metningspunkt over tid og egner seg godt til tjenester hvor det kan defineres et metningsnivå. For tjenester hvor metningsnivået ikke kan defineres eller at det er lite hensiktsmessig å arbeide på akkumulerte data som f.eks. trafikkdata eller trafikkinntekter må det defineres andre modeller.

Det er generelt vanskelig å sette opp regler for hvordan metningsnivå/ potensial skal anslås. Forskjellene er store fra produkt til produkt og tjeneste til tjeneste. I tillegg kan forskjellige produkter konkurrere om samme potensial. Anslagene kan sies å være avhengig av hvor vanskelig det er å definere ordet "behov". På den andre siden er potensialvurderingene så viktige at man bør bruke tilstrekkelig tid til å tallfeste potensialet. Slike potensialer kan tallfestes på blant annet følgende måter:

1 Kunde for kunde

Liste opp alle kunder og mulige (potensielle) kunder. Dette gjøres når man har en begrenset kundekrets, og kundene telles opp i kombinasjon med et gitt behov.

2 Segmentering

Kunder og potensielle kunder grupperes inn i større utvalg/segmenter. Hvert utvalg eller segment blir vurdert med hensyn på potensialet.

3 Totalvurderinger

Vi kan beregne hvor mye hele befolkningen/ hele landet kan kjøpe/ forbruke ved å benytte oss av nøkkel- og normtall. (Som f.eks. fås fra offisielle statistikker samt godt kjennskap til produktet.) For forbruksvarer kan vi for eksempel benytte befolkningstall (prognoser) multiplisert med tidligere års gjennomsnittsforbruk. Et annet eksempel er antall eneboliger som bygges hvert år multiplisert med gjennomsnittlig antall dører pr. enebolig. Tilsvarende for vedlikehold av dører i andre eksisterende boliger. Med hensyn på markedet for vedlikehold av dører (utskifting) kan gjennomsnittlig levetid komme inn som en god hjelpefaktor i potensialvurderingen.

Eksempel på langtidsprognose ved hjelp av S-kurve for telefonabonnement

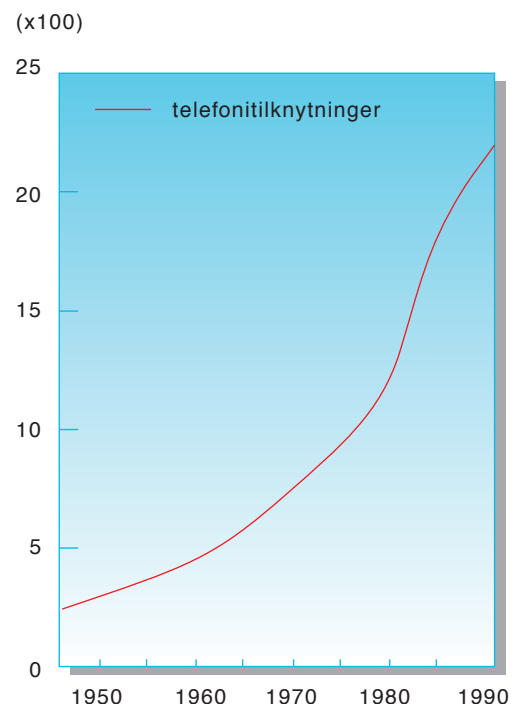
Vi har totaltall for innkoblede telefontilknytninger i Norge fra og med 1946 til og med 1991. Denne statistikken er utarbeidet av Teledirektoratet 1992 og heter *Statistikk 1991*.

Figur 12 viser tabell 1 grafisk.

Vi ser av dataene i tabell 1 at utviklingen er eksponentielt voksende. Samtidig kan vi intuitivt anta at antall telefontilknytninger før eller senere når en metning. Dette passer godt overens med det vi vet om en S-kurve og det er nå om å gjøre å estimere parametre som passer best til denne. Vi benytter etterspørselsfunksjonen:

$$S(t) = \frac{POT}{(1 + e^{(a-bt)})^c}$$

La oss anta at vi befinner oss i 1970 (det vil si med data fra 1946 til og med 1969). Vi skal nå lage en prognose ti år fram i tiden ved hjelp av S-kurven og regresjon for bestemmelse av parametrene a og b . Markedspotensialet M og konstanten c må vurderes først. I 1970 viste befolkningsprognosene at landets befolkning over tjue år lå relativt stabilt rundt



Figur 12 Tabell 1 vist grafisk

Tabell 2 viser de historiske data vi benytter, transformerte data, tilpasset S-kurve og kvadratavviket

År	Historiske data	Transformerte data	Tilpasset S-kurve	Kvadratavvik
1946	236,7	10,15825745	242	27,29
1947	252,8	9,895025198	253	0,35
1948	267,9	9,66295157	265	6,21
1949	279,4	9,494817245	278	1,99
1950	291,5	9,325221104	291	0,11
1951	304,8	9,146740414	305	0,03
1952	320,1	8,950806859	319	0,45
1953	337	8,744979303	335	5,89
1954	353	8,559406328	350	6,59
1955	368,6	8,386394168	367	2,42
1956	387,3	8,188394103	384	8,17
1957	407,4	7,985948103	403	22,43
1958	421,8	7,846954334	422	0,00
1959	434,7	7,726404789	442	49,49
1960	454,9	7,54463101	463	60,34
1961	486,6	7,275008154	485	4,04
1962	510	7,086991809	508	6,01
1963	530,9	6,926194304	532	0,48
1964	552,9	6,763607269	557	14,94
1965	582,6	6,554043561	583	0,28
1966	607,6	6,385719151	611	9,80
1967	637,1	6,195727709	640	6,40
1968	671,6	5,984303768	670	2,94
1969	707,8	5,773720648	702	38,96

Tabell 3 viser prognose for perioden 1970 – 1980 sammenliknet med faktisk utvikling

År	Siste observerte nivå	Forlenget S-kurve	Årlig vekst S-kurven	Prognose for 1970–1980	Faktisk utvikling	Avvik prognose og faktisk utvikling
1969	707,8	702			708	
1970		735	33	742	741	-1,35
1971		769	35	782	776	-6,36
1972		806	36	816	812	-3,55
1973		844	38	848	850	2,12
1974		883	40	879	890	10,54
1975		925	42	914	931	17,27
1976		968	43	956	975	18,68
1977		1014	45	1005	1020	15,43
1978		1061	47	1054	1067	12,89
1979		1110	49	1114	1117	2,51
1980		1162	52	1197	1168	-29,05

3 000 000 personer. Vi kunne da anta at markedspotensialet kunne sannsynliggjøres ved at hver voksne person ville ha tilgang på en telefontilknytning i gjennomsnitt. For en familie på to barn og to voksne, hvor de to voksne er over tjue er behovet én tilknytning på arbeidsplassen og én i boligen. Dette er ingen sofistikert måte å beregne et markedspotensial på, men som et eksempel for å sannsynliggjøre et potensial er dette godt nok. Vi setter markedspotensialet til 3 000 000 telefontilknytninger.

$$M = 3\,000\,000$$

Konstanten c estimeres ved å velge den c som gir det laveste kvadratavviket mellom historiske data og tilpasningen.

$MinQ$

$$= \sum_{t=1}^n (Historikk_t - Tilpasning_t)^2$$

Ved å minimere ovennevnte funksjon finner vi ut at c er et sted rundt 0,25.

Vi velger da $c = 0,25$

a og b parameteren blir ved hjelp av regresjon i EXCEL regneark beregnet til

a 370,6452702

b -0,185290001

Figur 13 viser de historiske data sammen med tilpasningen hentet fra tabell 2.

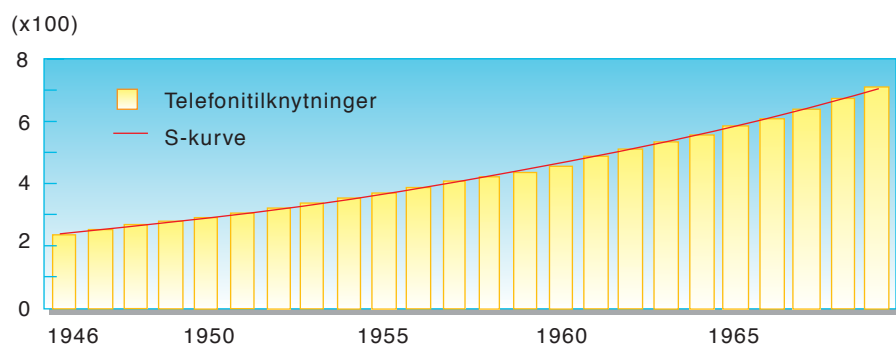
Figur 14 viser årlig vekst på historiske data og estimert årlig vekst i S-kurven.

Når vi nå skal lage prognose fram til og med 1980 for videre å sammenlikne med de eksisterende data for perioden, tas det utgangspunkt i nivået på siste observasjon (det vil si 707,8 i 1969) og legger på estimert årlig vekst hentet fra en forlenget S-kurve.

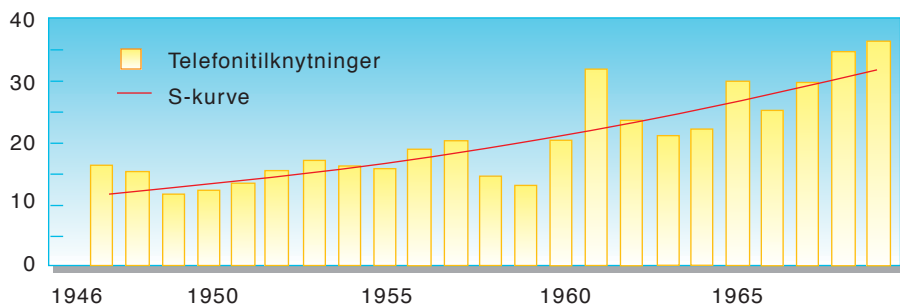
Vi kan se av tabell 3 at differansen mellom prognose og faktisk utvikling fram til og med 1973 inneholder relativt små avvik. Deretter har vi et voksende positivt avvik som fra og med 1977 avtar, for så å bli negativt. En nærmere analyse av avvikene vil kunne vise at disse i stor grad skyldes at den såkalte ventelisten på abonnement økte kraftig i 1974 og 1975, for så å avta etter 1979. Dette er forhold som vi ikke har tatt hensyn til i modellen som er benyttet.

Figur 15 viser prognosen for perioden 1970 – 1980 og den faktiske utviklingen i samme periode.

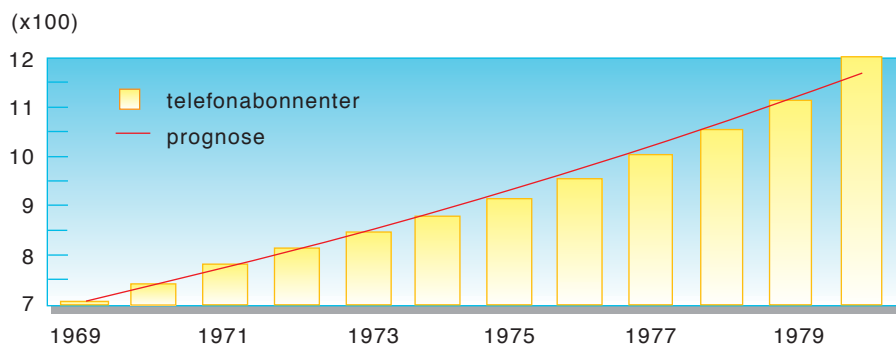
Modellen som er beskrevet her kan være et verktøy for å beskrive utviklingen av tjenester som er begrenset med et markedspotensial. For at modellen skal kunne gi brukbare resultater, må det arbeides med å anslå et fornuftig anslag for markedspotensialet. Dette bør gjøres med markedsanalyser i kombinasjon med realistiske anslag.



Figur 13 De historiske data sammen med tilpasningen hentet fra tabell 2



Figur 14 Årlig vekst på historiske data og estimert årlig vekst i S-kurven



Figur 15 Prognosen for perioden 1970 – 1980 og den faktiske utviklingen i samme periode

Box-Jenkins metode

AV CARLO HJELKREM OG JOHANNES BØE

001.18:621.39

1 Innledning

I mange tilfeller kan prognosemakeren være i en situasjon der prognosene ikke bygger på annet enn kunnskap om tidsseriens historiske utvikling. Såkalte glattingsmodeller, f.eks. eksponentiell glattning, Holts metode og Holt-Winters metode er metoder som genererer prognoser bare på bakgrunn av slik informasjon. Se for øvrig [7] og [8]. Intuitivt kan det reises spørsmål om metoder som ikke benytter seg av forklaringsvariabler vil gi gode prognoseresultater. Likevel er slike metoder viktige fordi

- i noen situasjoner, f.eks. ved planlegging av trafikk i ulike deler av telenettet eller ved produksjon av mange produkter, kan dette være den eneste praktiske framgangsmåten pga alle prognosene som skal lages
- i andre situasjoner kan det være umulig å finne data for relevante forklaringsfaktorer
- selv i situasjoner hvor relevante forklaringsfaktorer kan finnes, vil "envariabel-modellene" raskt gi enkle modeller som danner utgangspunkt mht forbedringer gjennom utarbeiding av mer sofistikerte modeller
- "envariabel-modellene" avdekker ofte store restledd, dvs. outliere, som har sammenheng med unormale hendelser (f.eks. streik) eller feil registrerte data. På denne måten vil slike modeller ofte være til hjelp i tidlige faser av en analyse.

Generelt vil vi likevel ikke forvente at slike modeller oppfører seg tilfredsstillende når vi skal lage prognoser over lengre perioder. Spesielt ikke hvis andre variabler som påvirker den variabel vi lager prognose for, utvikler seg annerledes i prognoseperioden enn de har gjort i fortid. Derfor er det generelt antatt at slike modeller egner seg best til "korte" prognoser.

Det var nettopp for kortsiktige prognoser at Box og Jenkins gjorde suksess med sine ARIMA-modeller (ARIMA-modeller og Box-Jenkins modeller brukes ofte synonymt). ARIMA står her for AutoRegressive Integrated Moving Average. Forklaring på disse uttrykkene følger nedenfor.

Selve grunnideen bak Box og Jenkins metode var ikke ny da den ble lansert: Å utvikle modeller som beskriver korrelasjonen, dvs. avhengigheten, mellom

observasjonene i en tidsserie. Dette gjøres generelt ved:

- Å minimere den del av tidsseriens utvikling som ikke kan forklares, og
- Å sørge for at det ikke er noen korrelasjon tilbake i den uforklarte delen (restleddet).

Dessuten var både autoregressive prosesser (AR) og glidende gjennomsnittsprosesser (MA) velkjente i statistisk teori.

Det nye i metoden lå i måten AR-prosesser og MA-prosesser ble sett i sammenheng: Det faktum at en AR-prosess med mange ledd tilnærmet kan beskrives ved en MA-prosess med ett (få) ledd, og omvendt, gjør ARIMA-modellene svært anvendbare når prinsippet om enkle modeller med få parametere (ledd) skal håndheves. Før vi går nærmere inn på dette, skal vi se på en kort beskrivelse av AR- og MA-prosessene.

Det vil i denne artikkelen føre for langt å gi en grundig matematisk/statistisk innføring i Box-Jenkins metode. Undervisning på slikt nivå foregår på høyskoler og over ett eller flere semestre. Likevel skal vi her forsøke å gi en ikke-teknisk innføring i ideene bak metoden med eksempler fra abonnementsutviklingen i Region Oslo¹. Noen få matematiske symboler og statistiske termer må vi likevel innom.

2 Autoregressiv prosess

En autoregressiv (auto = selv, regresiv = tilbakegående) beskrivelse av en tidsserie uttrykker den siste verdien av tidsserien i form av en veiet sum av de forutgående verdiene av tidsserien. Hvis tidsserievariabelen er kalt y , betyr dette at den siste verdien av y , y_t , beskrives ved hjelp av tidsserieverdier y_{t-1} , y_{t-2} , ..., osv. perioder før:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t \quad (1)$$

hvor $\phi_1, \phi_2, \dots, \phi_p$ er vekten som tillegges tidsserieverdier ved tidspunkt $t-1, t-2, \dots, t-p$ og a_t er den delen av tidsserien (ved tidspunkt t) som ikke kan beskrives ved AR-prosessen (støyleddet, residualen).

Vi forutsetter at støyleddet har forventning lik 0, at støyleddene er uavhengige, at de har konstant varians og at de er Normalfordelte.

Vektene, eller parametrene, representert ved ϕ -ene, vil her, som i tilfellet med regresjonsanalyse, være gjenstand for vår

interesse: Vi vil estimere dem ut fra de data som foreligger. Slik som (1) matematisk er formulert, beskrives y av p forutgående y -verdier. Vi sier da at y kan beskrives ved en AR-prosess av p -te orden og skriver dette kort som $AR(p)$.

I de fleste praktiske situasjoner vil antall verdier av forutgående tidsserieverdier, dvs. p , sjelden overstige 2. I så fall sier vi at vi har en AR-prosess av andre orden og skriver $AR(2)$. (1) blir da slik:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t \quad (2)$$

Tilsvarende vil en AR-prosess av første orden kun ha en forutgående y -verdi og bli betegnet med $AR(1)$.

I tidsserier med sesongvariasjoner kan den autoregressive beskrivelse lett utvides til å beskrive den siste tidsserieverdien som en funksjon av tidsserieverdier en, to, tre, ..., osv. sesonger før. Da benytter vi store greske bokstaver med fotskrift som refererer seg til antall sesonger for parametrene, f.eks. Φ_1 for sesong en, og fotskrift $t-s, t-2s, t-3s, \dots$, for observasjonene hhv. 1, 2, 3, ..., sesonger bakover. s er her sesonglengden.

3 Glidende gjennomsnittsprosess

En glidende gjennomsnittsbeskrivelse av en tidsserie uttrykker den siste verdien av tidsserien i form av en veiet sum av siste og tidligere verdier av støyleddene:

$$y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (3)$$

hvor $\theta_1, \theta_2, \dots, \theta_q$ er vekten som tillegges støyleddene på tidspunkt $t-1, t-2, \dots, t-q$ og symbolene ellers er som for AR-prosessen ovenfor. At det notasjonsmessig benyttes minustegn i stedet for plustegn, trenger vi ikke legge vekt på. På grunn av sammenhengen mellom AR- og MA-prosessene er det beleilig å skrive MA-prosessen med minustegn, mens det ellers ikke har noen betydning når vi skal estimere θ -ene.

Denne måten å beskrive en tidsserie på, kan ved første øyeblikk virke noe fremmed: Vi beskriver y som en funksjon av (siste og) tidligere feilledd. Sammenlikner vi imidlertid denne beskrivelsen

¹ Av praktiske grunner er tidsseriene av ulik lengde i de ulike eksemplene. Dette påvirker likevel ikke hovedpoengene i hvert tilfelle.

med en *servomekanisme*, lar den seg likevel forstå: I en servomekanisme korrigeres det for avviket (feilen) i forhold til den fastsatte kurs og ikke for hva kursen faktisk er.

Den matematiske formulering (3), beskriver y ved q forutgående a -verdier (og a_t). Vi sier da at y kan beskrives ved en MA-prosess av q -te orden og skriver dette kort som $MA(q)$.

I de fleste praktiske situasjoner vil antall verdier av forutgående restledd, dvs q , sjelden overstige 2. I så fall sier vi at vi har en MA-prosess av andre orden og skriver $MA(2)$. (3) blir da slik:

$$y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad (4)$$

Tilsvarende vil en MA-prosess av første orden kun ha en forutgående a - verdi ($+ a_t$) og bli betegnet med $AR(1)$.

I tidsreier med sesongvariasjoner kan den glidende gjennomsnittsbeskrivelse lett utvides til å beskrive den siste tidsserieverdien som en funksjon av restleddsverdier en, to, tre, ..., osv sesonger før. Da benytter vi store greske bokstaver med fotskrift som refererer seg til antall sesonger for parametrene, f eks Θ_1 for sesong en, og fotskrift $t-s$, $t-2s$, $t-3s$, ..., for støyleddene hhv 1, 2, 3, ..., sesonger bakover. s er her sesonglengden.

4 Autoregressiv-glidende gjennomsnittsprosess

I praksis finnes det også mange tilfeller av tidsrekker med både AR- og MA-prosesser (såkalte ARMA-prosesser) og disse refereres ofte til som miksede eller blandede prosesser. Matematisk uttrykkes dette ved en kombinasjon av (1) og (3):

$$y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (5)$$

hvor ϕ -ene og θ -ene er som beskrevet i kapittel 3 og 4. Ved å flytte alle y -ene i forutgående perioder over på høyre side av likhetstegnet, får vi

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (5)'$$

Vi ser altså at y er en funksjon av p tidligere observasjoner og støyleddene i q tidligere perioder. Vi sier da at y kan beskrives ved en ARMA-prosess av orden p og q . Som kort form skrives dette $ARMA(p,q)$.

I praksis vil vi også her oppleve at antall parametre sjelden overstiger 2 av hhv AR og MA. I så fall har vi sjelden mer enn en ARMA(2,2)-prosess. (5) ser da slik ut:

$$y_t - \theta_1 y_{t-1} - \theta_2 y_{t-2} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad (6)$$

Tilsvarende kan vi ha ARMA-prosesser med ARMA(1,2), ARMA(2,1) og ARMA(1,1). Ved lavere orden reduseres ARMA-prosessen naturlig nok til en ren AR- eller MA-prosess.

5 Stasjonærhet

For å beskrive avhengigheten i en tidsserie ved hjelp av AR-, MA- eller ARMA-prosesser, må tidsserien først være *stasjonær*. Et annet uttrykk for dette kan være "statistisk likevekt". Hvis tidsserien i utgangspunktet ikke er stasjonær, må den først transformeres eller omformes slik at den blir det.

Uten å gå inn på den statistiske definisjon, kan vi si at en stasjonær tidsrekke har disse egenskapene:

- Konstant varians
- Ingen trend: tidsserien skal svinge rundt et konstant nivå
- Ingen regulær sesongkomponent.

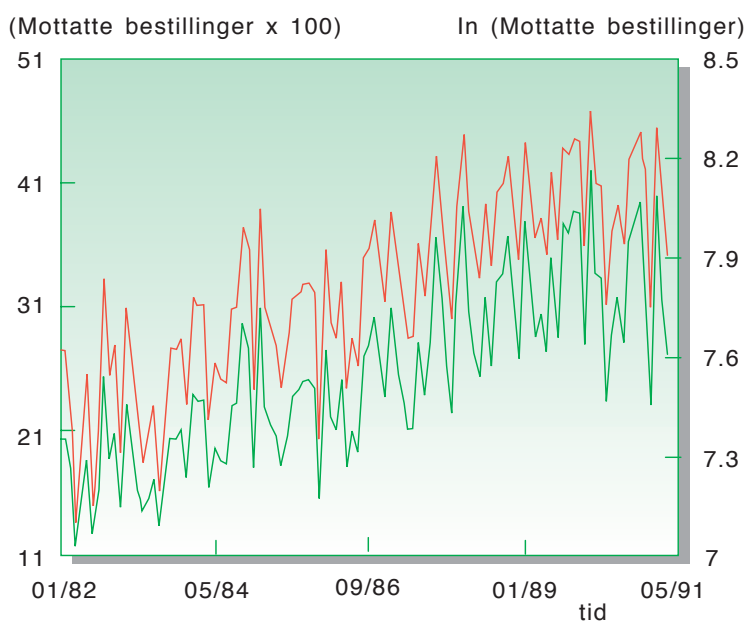
Vi skal nå se på hvordan vi behandler en tidsrekke slik at den blir stasjonær. Vari-

ans-stabilitet oppnås ved en passende transformasjon av data, mens fjerning av trend og regulær sesongkomponent gjøres ved differensiering av tidsrekken. (Metoden kan strengt tatt også benyttes uten at sesongkomponenten er fjernet.)

5.1 Stabilisering av variansen

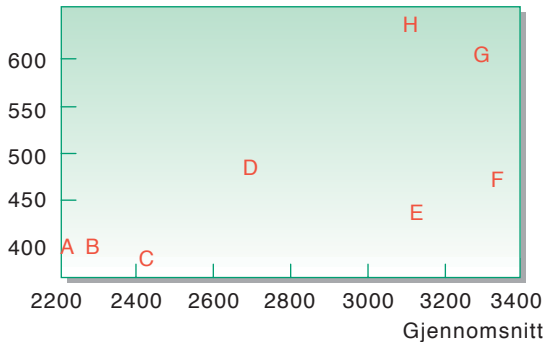
Ofte har tidsrekker varians som ikke er stabil over hele måleperioden. Et eksempel er akkumulerte data; vi får i mange tilfeller økende varians med økende nivå. I så fall må vi foreta en transformasjon av data, slik at variansen stabiliseres. Det mest vanlige er å ta logaritmen til tidsserien. Dette "krymper" variasjonen i observasjonene, men da på en slik måte at store verdier blir krympet mer enn små. Eksempel på hvilken effekt en slik transformasjon har på et valgt datasett, er vist i figur 1.

Vi ser her tidsserien for mottatte bestillinger pr måned for hovedabonnement (HA) i Region Oslo i perioden januar 1982 – april 1991. Den er markert med den heltrukne linjen, og verdiene kan avleses på venstre vertikale akse. Den stiplede linjen viser den samme tidsserien, men nå transformert til logaritmiske verdier som kan avleses på den høyre vertikale akse. Vi ser at den logaritmiske transformasjonen har medført at tidsseriens variasjonsområde er krympet fra verdier mellom ca 1150 og ca 4150 på den vertikale akse til venstre, til verdier mellom ca 7 og 8,2 på den vertikale akse



Figur 1 Mottatte bestillinger for HA (heltrukken) og \ln til mottatte bestillinger (stiplet) januar 1982 til april 1991. Region Oslo

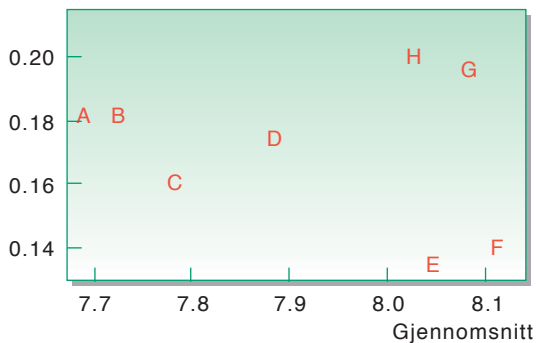
Standardavvik



A: Undersett for 1984 H: Undersett for 1991

Figur 2 Variasjons-gjennomsnitts-plott for 12 og 12 verdier av mottatte bestillinger HA pr måned 1984 – 1991. Region Oslo

Standardavvik



Figur 3 Variasjons-gjennomsnitts-plott. Logaritme-transformerte mottatte bestillinger HA pr måned 1984 – 1991. Region Oslo

til høyre. Samtidig er “topografien”, dvs elementer som sesongmønster, trendutvikling og andre mulige egenskaper ved den opprinnelige tidsserien, beholdt i den logaritmiske versjonen.

Andre transformasjoner kan også være aktuelle for å stabilisere variansen, mens det i enkelte tilfeller kan være vanskelig å finne en transformasjon som løser dette problemet². I praksis vil imidlertid ln-transformasjonen være tilstrekkelig i de fleste tilfeller.

En metode for avdekking av hvordan variansen utvikler seg er *variasjons-gjennomsnitts-plottet*: I en graf plotter vi gjennomsnittet av grupper av observasjoner langs en akse, mens spredningen, f eks gitt ved variansen eller standardavviket, til de samme observasjonene plottes langs den andre akse. Hvis spredningen til gjennomsnittene øker når gjennomsnittene vokser, er dette en klar indikasjon på behov for stabilisering av variansen.

Siden vi i vårt eksempel opererer med månedsdata, er det mest praktisk å gruppere observasjonene 12 og 12. I figur 2 er variasjons-gjennomsnitts-plottet for mottatte bestillinger på HA i perioden 1982 (A) til 1991 (H) vist.

I figur 2 er årsgjennomsnittene rangert etter stigende verdi. Vi ser tydelig at standardavviket stiger når årsgjennomsnittene vokser.

I figur 3 er variasjons-gjennomsnitts-plottet vist for ln-verdien av de samme

observasjonene som i figur 2. Ut fra figur 3 er det ikke grunnlag for å si at standardavviket er en stigende funksjon av gjennomsnittet.

5.2 Fjerning av trend og den regulære sesongkomponent

For å “nøytralisere” den trend som en tidsserie måtte ha, kan tidsserien *differensieres*. Dette betyr å omforme tidsserien til en ny tidsserie ved å trekke to påfølgende observasjoner fra hverandre:

$$v_t = y_t - y_{t-1} \quad (7)$$

Antar vi at $y_t, y_{t-1}, \dots, y_{t-n}$ er våre opprinnelige observasjoner, er v_t den nye observasjonen som framkommer når differensieringen er foretatt. Den nye og differensierte tidsserien blir nå $v_t, v_{t-1}, \dots, v_{t-n-1}$. Hvis den nye tidsserien fremdeles har trend, kan den differensieres en gang til, dvs:

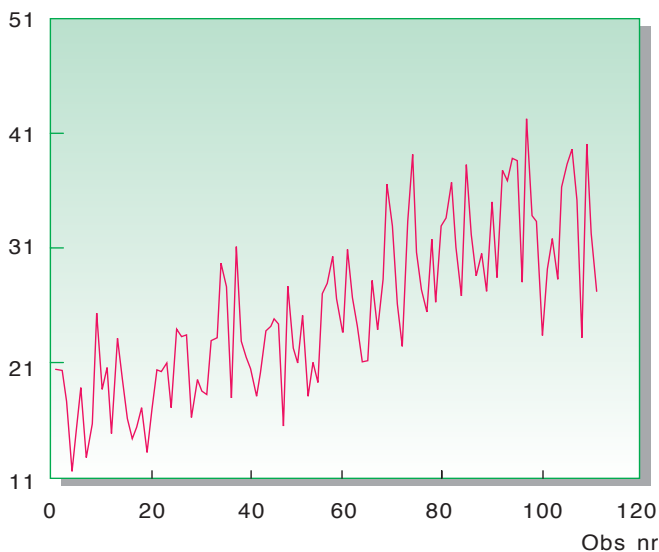
$$w_t = v_t - v_{t-1} \quad (8)$$

² En vid klasse av transformasjoner er den såkalte Box-Cox transformasjonen. Den er gitt ved:

$$g(y_t) = \frac{y_t^\lambda}{\lambda}$$

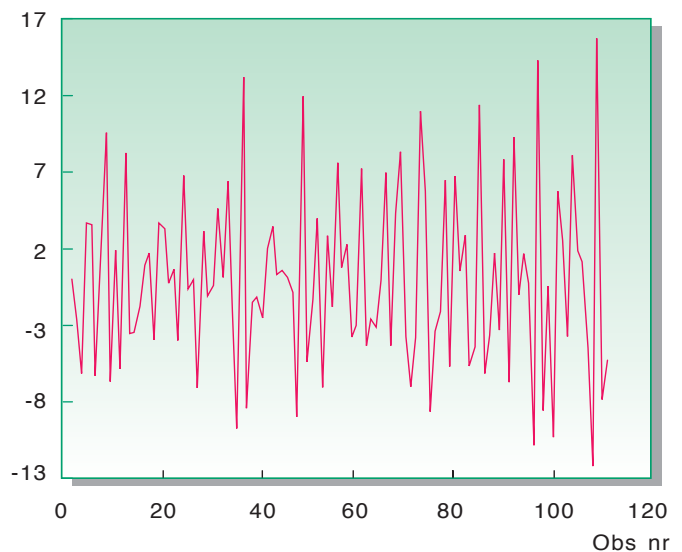
$l = 1$ tilsvarende $g(y_t)$ de opprinnelige observasjonene, mens $l = 0$ gir ln-transformasjon.

(Mottatte bestillinger x 100)



Figur 4 Mottatte bestillinger, HA, januar 1982 – april 1991. Region Oslo

(Mottatte bestillinger x 100)



Figur 5 Mottatte bestillinger, HA, differensiert en gang. Januar 1982 – april 1991. Region Oslo

hvor w er den nye observasjonen som framkommer etter at vi har differensiert for andre gang. Den to ganger differensierte tidsserien blir nå $w_t, w_{t-1}, \dots, w_{t-n-2}$.

På denne måten kan vi fortsette helt til tidsserien blir kvitt trenden.

Vær imidlertid oppmerksom på at hver gang vi differensierer, får vi en observasjon mindre tilbake. Dette fordi vi hele tiden trekker sammen (trekker fra) to og to observasjoner: Anta at vi har en tidsserie bestående av 66 observasjoner. Hvis vi differensierer denne tidsserien en gang og betegner den nye tidsserien med v -er, får vi:

$$\begin{aligned} v_1 &= y_1 - y_2 \\ v_2 &= y_2 - y_3 \\ v_3 &= y_3 - y_4 \\ &\vdots \\ v_{65} &= y_{65} - y_{66} \end{aligned}$$

Hvis vi har få observasjoner til rådighet, kan gjentatte differensieringer redusere antall observasjoner slik at analysen blir usikker.

La oss se på tidsserien i figur 4, dvs mottatte bestillinger i Region Oslo for perioden januar 1982 – april 1991. I figur 5 er de samme data differensiert en gang.

Vi ser at den differensierte tidsserien ikke vokser slik som den opprinnelige i figur 4. Det er rimelig å anta at den differensierte tidsserien er tilnærmet stasjonær når det gjelder trenden. Det er imidlertid ikke nok å forsøke å avgjøre dette med det blotte øyet. Testing ved hjelp av bl a autokorrelasjonsfunksjonen (se kapittel 6) kan hjelpe oss til en statistisk bedømmelse av dette.

På samme måte som ovenfor kan vi også differensiere mhp sesong, for å nøytralisere den regulære sesongkomponent. Da trekker vi to observasjoner med sesongens lengde fra hverandre:

$$v_t^s = y_t - y_{t-s} \quad (9)$$

Her er s sesongens lengde, og den brukes nå som topp-skrift på den nye sesongdifferensierte tidsserien:

$$v_1^s, v_2^s, \dots, v_{n-s}^s$$

s kan her f eks være 12 hvis vi har månedsdata eller 4 hvis vi har kvartalsdata.

Også her kan vi fortsette differensieringen hvis den regulære sesongkomponent ikke er nøytralisert.

På samme måte som ved differensieringen som ikke gjaldt sesong, mister vi også her observasjoner hver gang vi sesongdifferensierer. Imidlertid er tapet mhp antall observasjoner større her, fordi en hel sesong blir borte for hver differensiering: Anta at vi har en tidsserie med 24 observasjoner av kvartalsdata. Vi differensierer tidsserien en gang mhp sesong og får den nye tidsserien som vi betegner med v^A -er:

$$\begin{aligned} v_1^A &= y_5 - y_1 \\ v_2^A &= y_6 - y_2 \\ v_3^A &= y_1 - y_3 \\ &\vdots \\ v_{20}^A &= y_{24} - y_{20} \end{aligned}$$

Her blir altså 4 observasjoner borte. Differensierer vi en gang til mhp sesongen blir enda 4 observasjoner borte, osv. Tilsvarende vil en sesongdifferensiering av månedsdata redusere antall observasjoner med 12, to ganger sesongdifferensiering av månedsdata reduserer antall observasjoner med 24, osv. Igjen gjelder det at antall observasjoner vi i utgangspunktet har til rådighet vil være avgjørende for sikkerheten i analysen fordi sesongdifferensiering skjærer kraftig ned på antall observasjoner.

Figur 6 gir et eksempel på effekten av sesongdifferensiering. Igjen er det data fra figur 4 som er (sesong-)differensiert.

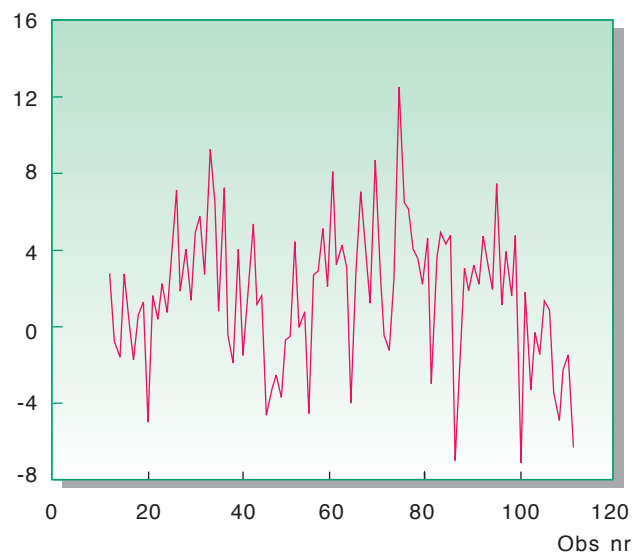
Går vi tilbake til figur 4 og sammenlikner alle januarobservasjoner, alle februarobservasjoner osv, vil vi finne et klart sesongmønster (bortsett fra enkelte uregelmessigheter). Dette er en (visuell) indikasjon på at det er behov for sesongdifferensiering. Ser vi på figur 6 er det her vanskelig å bedømme om vi har behov for flere sesongdifferensieringer. Her kan igjen autokorrelasjonsfunksjonen være til hjelp for å bedømme dette, uten at vi går inn på dette her.

I figur 6 er det ingen jevn trend å spore, slik som i figur 4. Likevel vet vi at den opprinnelige tidsserien har en klar positiv trend. Den sesongdifferensierte tidsserien er imidlertid svært "uroelig" mht nivå – den svinger ikke rundt et konstant nivå. I figur 7 er det foretatt begge

typer differensiering på de samme data. Resultatet er en tilsynelatende stasjonær tidsserie.

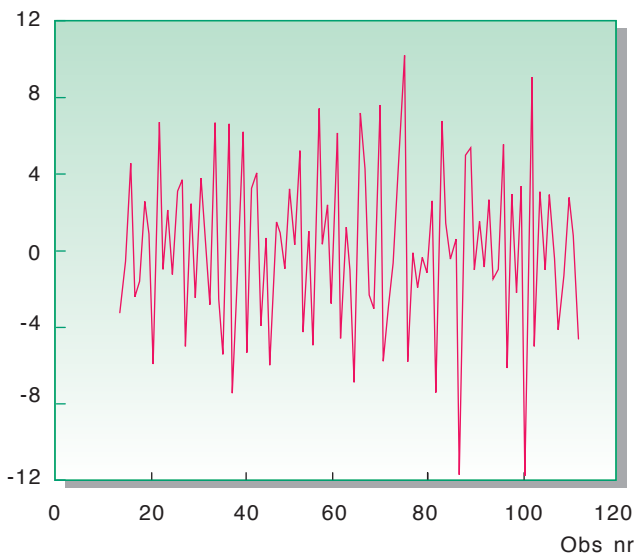
Det er gjennom differensieringene at I-en (Integrated) i ARIMA kommer inn. Den beskriver at variabelen må integreres for å komme tilbake til den opprinnelige variabelen når prognosene skal lages. Den generelle modell (uten sesongbeskrivelse) med d differensieringer, p

(Mottatte bestillinger x 100)



Figur 6 Sesongdifferensiering av mottatte bestillinger, HA, januar 1982 til april 1991. Region Oslo

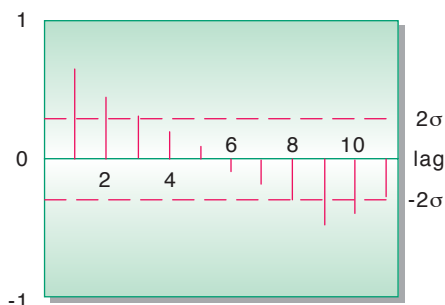
(Mottatte bestillinger x 100)



Figur 7 Differensiering og sesongdifferensiering av mottatte bestillinger, HA, januar 1982 – april 1991. Region Oslo

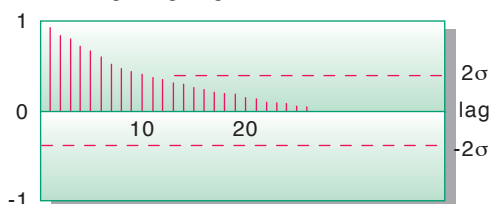
AR-parametre og q MA-parametre beskrives kort slik: $ARIMA(p,d,q)$.

Hvis det i tillegg er behov for å modellere sesongsvingninger, er den generelle kortversjonen: $ARIMA(p,d,q)X(P,D,Q)$. Her referer p , d og q til den sesonguavhengige delen av modellen, slik som i avsnittet ovenfor. P , D og Q i den andre parenteser beskriver den sesongmessige delen av modellen: P AR-parametere for beskrivelse av sesong, D sesongdifferensieringer og Q MA-parametere for beskrivelse av sesong.

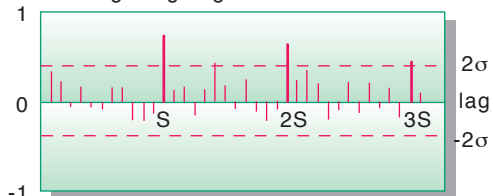


Figur 8 Eksempel på autokorrelasjonsplott

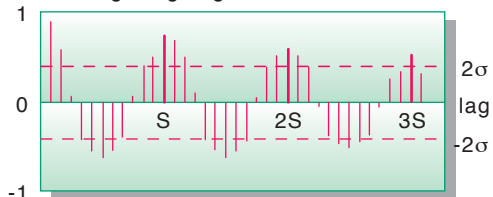
a) ikke-stasjonær tidsrekke uten sesongsvingninger



b) tidsrekke med ikke-stasjonære sesongsvingninger



c) tidsrekke med ikke-stasjonære sesongsvingninger



Figur 9 Teoretisk form på autokorrelasjonsfunksjonen for ikke-stasjonære tidsrekker

6 Autokorrelasjonsfunksjonen

Vi har i kapittel 5 vist hvordan vi kan få en tidsrekke til å bli stasjonær ved hjelp av transformasjon av data og differensiering. Eventuelt behov for transformasjon kan identifiseres vha variasjonsgjennomsnitts-plottet. Ut over en visuell betraktning av data, så vi imidlertid ingenting om hvordan behovet for differensiering kunne avdekkes presist.

Bruk av *autokorrelasjonsfunksjonen* er her sentral. Den måler avhengigheten mellom observasjonene i ulik tidsavstand, fra avstand 1 til k , og kan beregnes som følger:

1. Beregn autokovariansen mellom alle observasjonene i tidsavstand 1, gitt ved

$$C_1 = \frac{1}{N} \sum_{t=1}^N (y_t - \bar{y})(y_{t+1} - \bar{y}) \quad (10)$$

hvor

N = antall observasjoner vi har til rådighet

y = den variabel vi analyserer (f eks mottatte bestillinger)

t = tidspunkt

$$\bar{y} = \frac{1}{N} \sum_{t=1}^N y_t$$

2. Beregn variansen til tidsserien, gitt ved

$$C_0 = \frac{1}{N} \sum_{t=1}^N (y_t - \bar{y})^2 \quad (11)$$

3. Beregn autokorrelasjonsfunksjonen i avstand 1 (lag 1), gitt ved

$$r_1 = \frac{C_1}{C_0} \quad (12)$$

4. Gjenta steg 1. og 2. for observasjoner i avstand 2, 3, 4, ..., k , dvs beregn verdier for $r_2, r_3, r_4, \dots, r_k$. Autokorrelasjonsfunksjonens generelle form blir:

$$r_k = \frac{C_k}{C_0} = \frac{\sum_{t=1}^N (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^N (y_t - \bar{y})^2} \quad (13)$$

5. Beregn standardavviket til autokorrelasjonsfunksjonen, gitt ved

$$\sigma = \frac{1}{\sqrt{n}} \quad (14)$$

Autokorrelasjonsfunksjonen er en normalisert variabel som antar verdier mellom -1 og 1.

For identifikasjon og testing av meningsfylte autokorrelasjoner testes disse mot 2 ganger standardavviket (σ)³.

Det vanlige er å plote verdiene til autokorrelasjonsfunksjonen langs en akse, med økende avstand mellom observasjonene. Figur 8 viser et eksempel på et slikt autokorrelasjonsplott.

Vi ser her at autokorrelasjonsfunksjonen er plottet opp til avstand 11. I tillegg er verdien til 2 ganger standardavviket til autokorrelasjonene markert ved den stiplede linjen.

6.1 Identifikasjon av ikke-stasjonærhet

Ikke-stasjonære tidsrekker kan avdekkes ved hjelp av "formen" på autokorrelasjonsfunksjonen:

1. For en ikke-stasjonær tidsrekke uten sesongsvingninger vil autokorrelasjonsfunksjonen

- ha positive verdier
- være stor i tallverdi

- avta sakte i tallverdi ettersom avstanden mellom observasjonene øker.

2. For en tidsrekke med ikke-stasjonære sesongsvingninger vil autokorrelasjonsfunksjonen i avstander som refererer seg til sesongene

- ha positive verdier
- være stor i tallverdi
- avta sakte i tallverdi ettersom avstanden mellom sesongene øker.

Den teoretiske formen på autokorrelasjonsfunksjonen for disse tilfellene er vist i figur 9.

S tilsvarer sesong-lengden i figur 9, dvs 12 (måneder) slik figuren er tegnet.

Ofte kan det være vanskelig å tolke et autokorrelasjonsplott med henblikk på å avdekke ikke-stasjonærhet fordi:

³ Standardavviket (σ) gitt i (14) er en tilnærming til det statistisk korrekte uttrykk for dette. (14) underestimerer den sanne verdien av σ , men dette er en fordel når identifikasjon og modellsjekking skal utføres.

- Både ikke-stasjonæritet pga sesongsvingninger og utenom sesongsvingningene kan være til stede, noe som gir opphav til en blanding av de typene av autokorrelasjonsfunksjonene som er beskrevet i figur 9. I så fall vil autokorrelasjonsfunksjonen ha form som i figur 10a).
- Signifikante autokorrelasjoner (dvs autokorrelasjoner større enn 2 ganger standardavviket) kan være til stede utenom sesong-lengden. Dette kompliserer tolkningen av ikke-stasjonære sesongsvingninger, se figur 10b).

Etter at ikke-stasjonæritet er identifisert gjennom formen på autokorrelasjonsfunksjonen, bringes tidsrekken over på en stasjonær form ved å differensiere den en eller flere ganger, slik vi har vist i kapittel 5.2. I praksis bør en være oppmerksom på følgende:

- Autokorrelasjonsfunksjonen bør beregnes for så lang avstand mellom observasjonene at dens ikke-stasjonæritet blir avdekket. Dette kan gjøres ved å beregne den opp til avstand 15–20 for tidsrekker uten ikke-stasjonære sesongsvingninger og 3 ganger sesong (+1 observasjon) for tidsrekker med ikke-stasjonære sesongsvingninger.
- Det er i teorien ingen grenser for hvor mange ganger en tidsrekke bør differensieres for å bringe den over på stasjonær form. I praksis er det likevel uvanlig med mer enn to ikke-sesongavhengige differensieringer og en sesongavhengig differensiering.
- Mistolkning av autokorrelasjonsfunksjonen kan lede til over-differensiering. En tidsserie som går mot stasjonær tilstand vil redusere sin varians, mens over-differensiering øker variansen.

6.2 Eksempel

I kapittel 5.1 så vi hvordan vi kunne identifisere hvorvidt mottatte bestillinger i Region Oslo hadde konstant varians. Vi så på variasjons-gjennomsnitts-plottet og fant at en ln-transformasjon av data stabiliserte variansen.

I dette eksempelet er denne transformasjonen foretatt. For å identifisere eventuelt behov for differensiering starter vi med å plote autokorrelasjonsfunksjonen for den ln-transformerte tidsrekken, se figur 11.

Vi ser klart at formen på autokorrelasjonsfunksjonen likner på figur 9a), bortsett fra at den etter hvert kryper under 0.

Dette er et klart tegn på at tidsrekken har en trend, slik dette ble beskrevet under kapittel 5.2, og at det er behov for differensiering. I tillegg ser vi at formen på autokorrelasjonsfunksjonen er noe buet, slik som i figur 10a). Dette er et tegn på at det også kan være behov for sesong-differensiering. I figur 12 er tidsserien differensiert en gang.

Vi ser tydelig at indikasjonene på behov for differensiering pga trend er borte. Tilbake står signifikante (de er større enn 2 ganger standardavviket) autokorrelasjoner i avstand 12, dvs slik som i figur 9b). I tillegg ser vi enkelte andre store autokorrelasjoner. Noen av disse har klart form av figur 10b). Det er altså behov for å differensiere tidsserien en gang til, men nå mhp sesong. I figur 13 er tidsserien også differensiert en gang mhp sesong.

Vi ser nå tydelig at formen på autokorrelasjonsfunksjonen ikke har noen av de "mønstre" som vi beskrev i figur 9 og 10. Det er m a o store muligheter for at tidsserien er stasjonær. Tilbake står noen få autokorrelasjoner som er større enn 2 ganger standardavviket. Disse kommer vi tilbake til under modelleringen nedenfor. Resten har karakter av å være tilfeldig.

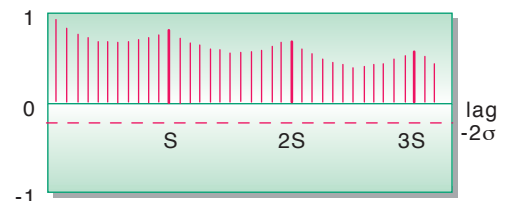
6.3 Identifikasjon av AR- og MA-prosesser i en stasjonær tidsserie

En stasjonær tidsserie vil som regel fremdeles være autokorrelert, ref de signifikante autokorrelasjonene i figur 13. Derfor har vi behov for å "fjerne" denne korrelasjonen ved å modellere den med AR- og MA-prosesser, slik at vi står igjen med støyledd som er uavhengige, dvs at det ikke er noe autokorrelasjon tilbake, ref forutsetningene om støyleddene i kapittel 2.

Før vi analyserer de signifikante autokorrelasjonene i en stasjonær tidsrekke bør vi også huske at selv i en tilfeldig tidsrekke, dvs en tidsrekke uten noe autokorrelasjon, vil vi vanligvis finne at 1 av 20 (5 %) autokorrelasjoner er signifikante når grensen for signifikans er 2 ganger standardavviket⁴. Det vi må spørre oss om er om autokorrelasjonene er meningsfylte:

- Signifikante autokorrelasjoner for observasjoner med liten avstand er en indikasjon på ikke-sesongavhengig eller periode til periode-prosesser.
- Signifikante autokorrelasjoner med avstand lik sesongperiodene eller 1, 2,

a) tidsrekke med ikke-stasjonæritet pga sesongsvingninger og utenom sesongsvingninger

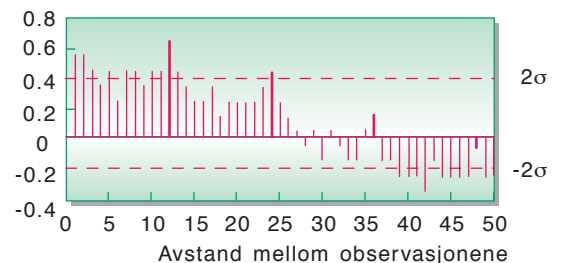


b) tidsrekke med ikke-stasjonære sesongsvingninger og signifikante autokorrelasjoner utenom sesonglengden



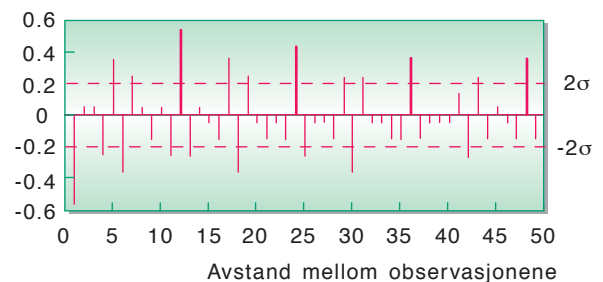
Figur 10 Teoretisk form på autokorrelasjonsfunksjonen for ikke-stasjonære tidsrekker. Spesielle tilfeller

Autokorrelasjonsfunksjonen



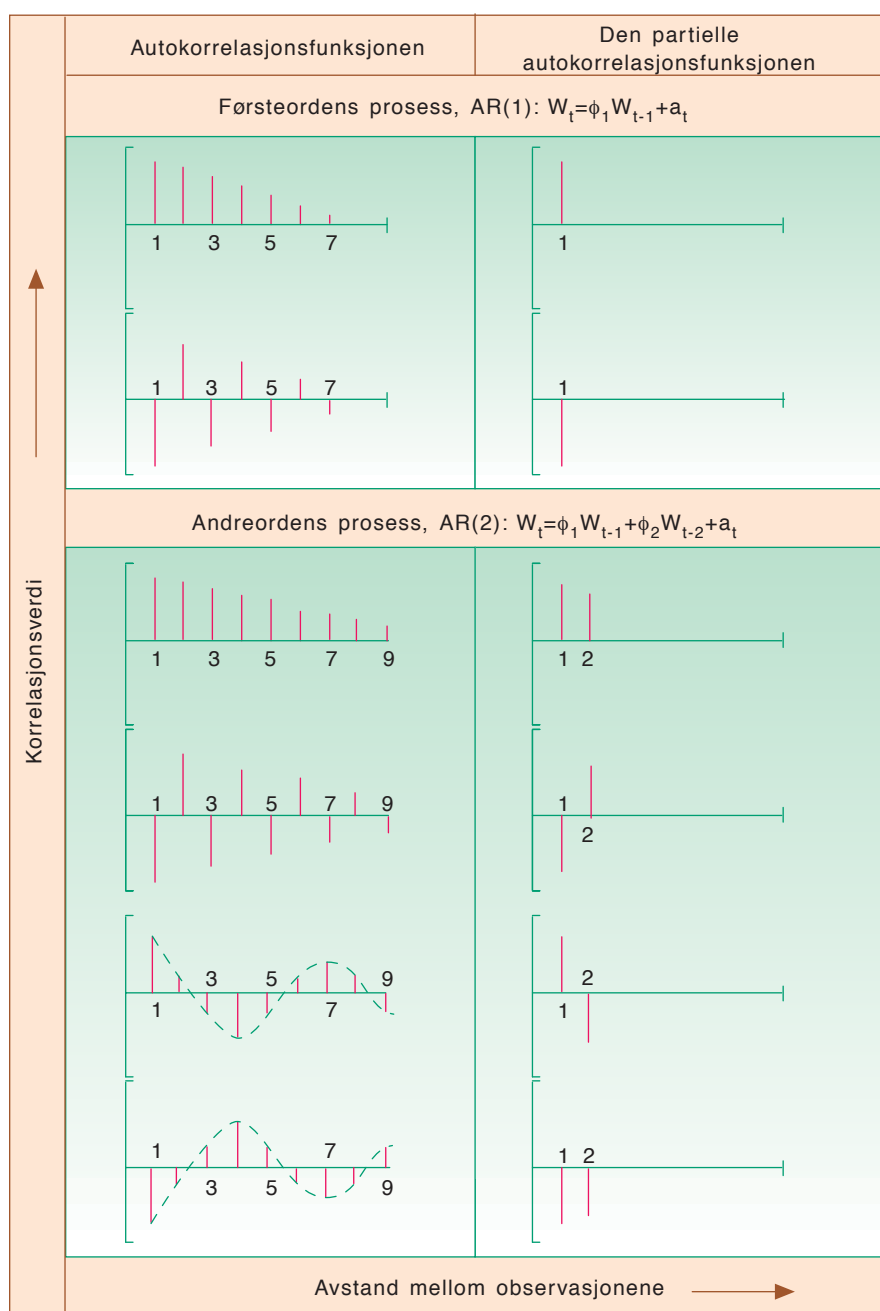
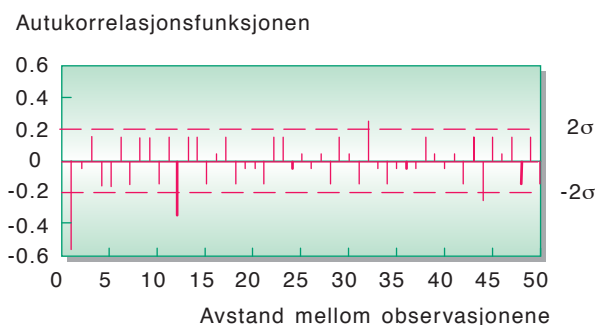
Figur 11 Autokorrelasjonsfunksjonen for ln-transformerte verdier av mottatte bestillinger, januar 1984 – desember 1991. Region Oslo

Autokorrelasjonsfunksjonen



Figur 12 Autokorrelasjonsfunksjonen for ln-transformerte verdier av mottatte bestillinger, januar 1984 – desember 1991. Region Oslo. Differensiert en gang

Figur 13 Autokorrelasjonsfunksjonen for ln-transformerte verdier av mottatte bestillinger, januar 1984 – desember 1991. Region Oslo. Differensiert en gang og sesongdifferensiert en gang



Figur 14 Teoretiske korrelasjonsfunksjoner for AR-prosesser uten sesongsvingninger

3, ..., ganger sesongperiodene er en indikasjon på sesong-prosesser.

Signifikante autokorrelasjoner utenom dette betraktes normalt ikke som meningsfylte, og vi ser derfor bort fra slike.

I tillegg er det viktig å huske at hvis data inneholder "unormale observasjoner" kan dette tilsløre det autokorrelasjonsmønster som vi forsøker å identifisere, eller (i beste fall) gi opphav til tilfeldige, ikke meningsfylte autokorrelasjoner.

Med unormale observasjoner menes her observasjoner som har spesielt høye eller lave verdier, f eks pga streik, definisjonsendringer, målefeil eller liknende. Slike observasjoner blir ofte kalt "outliere".

Før vi går videre definerer vi en annen viktig funksjon som benyttes ved identifikasjon av AR- og MA-prosesser: Den partielle autokorrelasjonsfunksjonen.

Den partielle autokorrelasjonsfunksjonen måler forholdet mellom to observasjoner med en gitt tidsavstand, etter at autokorrelasjonen ved alle tidsavstander mindre enn dette er fjernet.

Den partielle autokorrelasjonsfunksjonen estimeres ved å ta regresjonen⁵ av den stasjonære tidsrekken, dvs den differensierte og transformerte tidsrekken, ved tidspunkt t mot verdier på tidspunkt $t-1$, $t-2$, $t-3$, osv. Hvis w_t er den stasjonære tidsrekken, kan vi sette opp

$$w_t = \phi_{11} w_{t-1} + u_t \quad (15)$$

$$w_t = \phi_{21} w_{t-1} + \phi_{22} w_{t-2} + u_t \quad (16)$$

$$w_t = \phi_{31} w_{t-1} + \phi_{32} w_{t-2} + \phi_{33} w_{t-3} + u_t \quad (17)$$

·

·

·

osv

hvor u_t er støyleddet. Her er regresjonskoeffisientene ϕ_{11} , ϕ_{22} og ϕ_{33} de partielle autokorrelasjonsfunksjonene i hhv avstand 1, 2 og 3.

6.3.1 Den teoretiske autokorrelasjonsfunksjon og partielle autokorrelasjonsfunksjon for en AR-prosesser

I figur 14 er den teoretiske autokorrelasjonsfunksjonen og den teoretiske partielle autokorrelasjonsfunksjonen skissert

⁴ Dette tilsvarer en test på Normalfordelte variable med 95 %-signifikansnivå.

⁵ For en innføring i estimeringsmetoden regresjon, se [1].

for en 1. ordens og 2. ordens AR-prosess uten sesongsvingninger.

Vi ser av figur 14 at autokorrelasjonsfunksjonen kan beskrives ved at den

1. er signifikant, dvs har store tallverdier, når det er liten avstand mellom observasjonene⁶, og at den
2. har form som
 - avtagende eksponentiell utvikling
 - avtagende eksponentiell utvikling med skiftende fortegn
 - dempet sinus-kurve (bølgeform).

Vi ser også at det i tilfellet med avtagende eksponentiell utvikling og avtagende eksponentiell utvikling med skiftende fortegn er umulig å si om autokorrelasjonsfunksjonen beskriver en AR(1) eller en AR(2). Det er her den partielle autokorrelasjonsfunksjonen hjelper oss:

Vi ser av figur 14 at den partielle autokorrelasjonsfunksjonen kan beskrives ved at den

1. bare er signifikant i avstand 1 for en AR(1)-prosess og bare er signifikant i avstand 1 og 2 for en AR(2)-prosess. Dessuten at den
2. ikke følger noe fast mønster.

Med andre ord kan den partielle autokorrelasjonsfunksjonen hjelpe oss til å skille mellom en AR(1)-prosess og en AR(2)-prosess.

I figur 15 er den teoretiske autokorrelasjonsfunksjonen og den teoretiske partielle autokorrelasjonsfunksjonen skissert for en 1. ordens og 2. ordens AR-prosess for sesongsvingninger.

Vi ser at autokorrelasjonsfunksjonen og den partielle autokorrelasjonen i tilfellet med sesongsvingninger kan beskrives på den samme måten som i tilfellet hvor det ikke var sesongsvingninger, bortsett fra at det er i avstander som tilsvarer sesongperiodene.

Figur 14 og 15 beskriver kun prosesser opp til 2. orden. Dette er fordi det i de fleste praktiske situasjoner er enten 1. eller 2. ordens prosesser vi finner i tidsrekker. I tillegg er det slik at 3. ordens prosesser og høyere kan beskrives ved hjelp av en kombinasjon av 1. ordens

og/eller 2. ordens autoregressive prosesser.

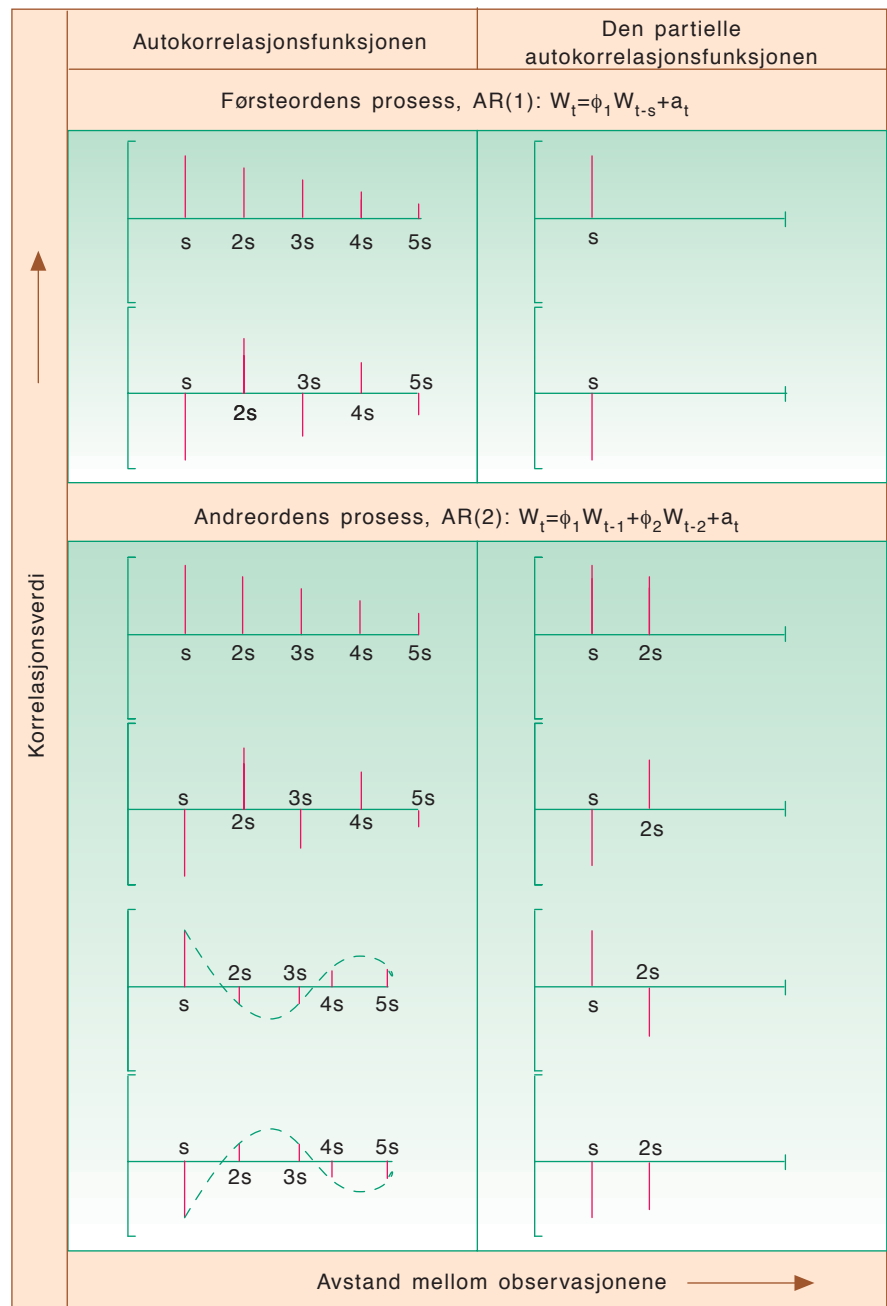
6.3.2 Den teoretiske autokorrelasjonsfunksjon og partielle autokorrelasjonsfunksjon for en MA-prosess

I figur 16 er den teoretiske autokorrelasjonsfunksjonen og den teoretiske partielle autokorrelasjonsfunksjonen skissert

for en 1. ordens og 2. ordens MA-prosess uten sesongsvingninger.

Vi ser av figur 16 at autokorrelasjonsfunksjonen kan beskrives ved at den

1. er signifikant bare for avstand 1 for 1. ordens prosess og
2. er signifikant bare for avstand 1 og 2 for 2. ordens prosess.



Figur 15 Teoretiske korrelasjonsfunksjoner for AR-prosesser som beskriver sesongsvingninger

⁶ Grensene for 2 ganger standardavviket er ikke tegnet inn her, siden de avhenger av antallet av faktiske observasjoner, ref formel (14).

Vi ser av figur 16 at den partielle auto-korrelasjonsfunksjonen kan beskrives ved at den har form som

1. avtagende eksponentiell utvikling for 1. ordens prosess og
2. avtagende eksponentiell utvikling eller dempet sinus-kurve (bølgeform) for 2. ordens prosess.

Den partielle autokorrelasjonsfunksjonen kan igjen hjelpe oss til å skille mellom en MA(1)-prosess og en MA(2)-prosess.

I figur 17 er den teoretiske autokorrelasjonsfunksjonen og den teoretiske partielle autokorrelasjonsfunksjonen skissert for en 1. ordens og 2. ordens MA-prosess for sesongsvingninger.

Vi ser at autokorrelasjonsfunksjonen og den partielle autokorrelasjonen i tilfellet med sesongsvingninger kan beskrives på den samme måten som i tilfellet hvor det ikke var sesongsvingninger, bortsett fra at det er i avstander som tilsvarer sesong-periodene.

6.3.3 Den teoretiske autokorrelasjonsfunksjon og partielle autokorrelasjonsfunksjon for en ARMA-prosess

I praksis kan tidsrekker ofte beskrives ved en eller flere autoregressive og glidende gjennomsnittsprosesser på samme tid. I slike situasjoner vil begge korrelasjonsfunksjonene ha form av avtagende eksponentiell utvikling og/eller dempet sinus-kurve, som igjen vil være brutt av verdier som ikke følger dette mønster i avstand 1 og/eller 2. Tilsvarende gjelder også for beskrivelse av sesongsvingninger.

Den vanligste framgangsmåten for å identifisere en ARMA-prosess er å modellere og estimere den prosessen som er mest framtreddende. Deretter kan tester og modell-evaluering avgjøre hvilke prosesser som er nødvendige i tillegg.

Det samme prinsippet gjelder for øvrig også for skillet mellom modellering av ikke-sesongavhengige og sesongavhengige prosesser.

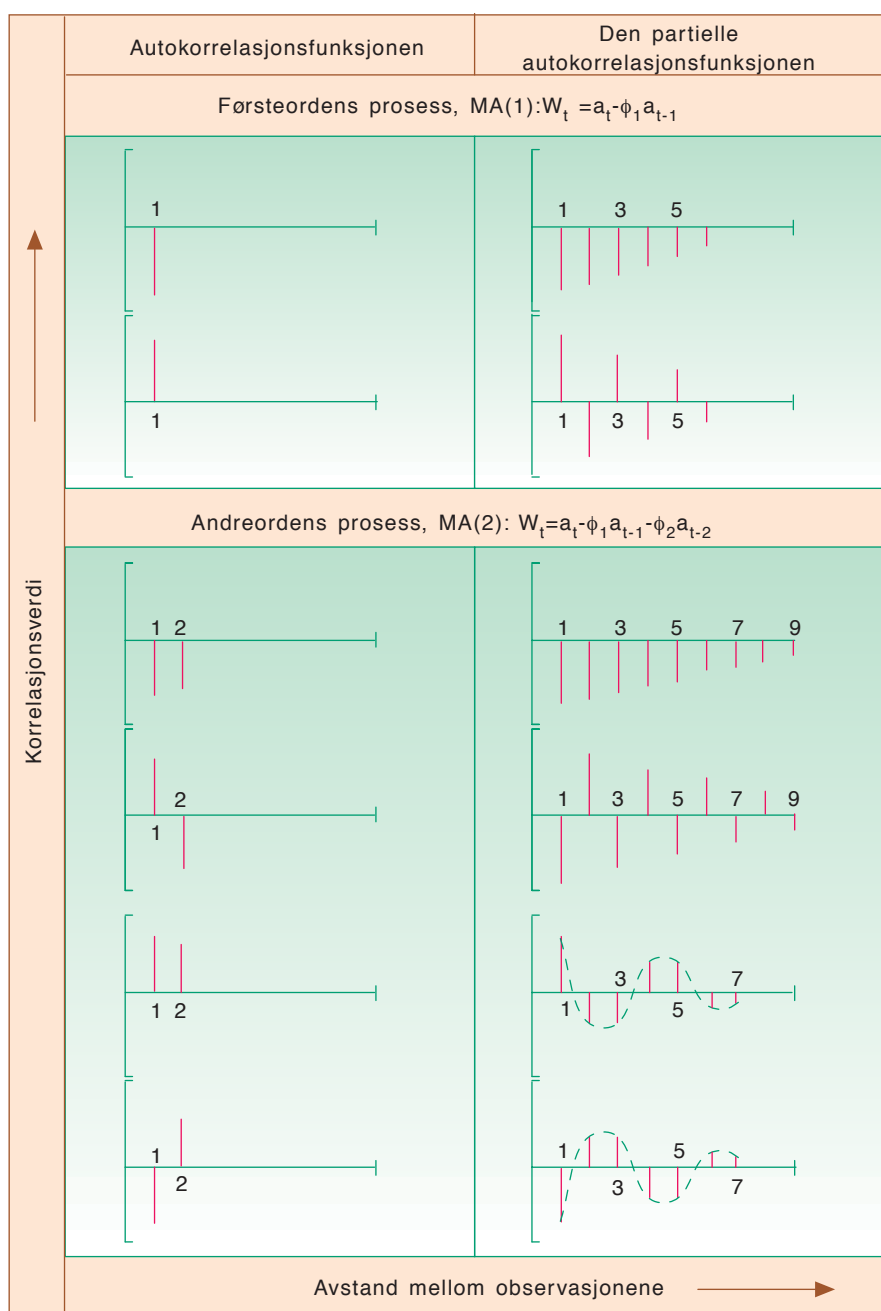
6.3.4 Eksempel

I kapittel 5.1 så vi hvordan vi stabiliserte variansen ved å ln-transformere tidsrekken mottatte bestillinger i Region Oslo. I kapittel 6.2 så vi hvordan tidsrekken ble differensiert slik at den ble stasjonær. Den trengte en ikke-sesongavhengig og en sesongavhengig differensiering. Figur 13 viste autokorrelasjonsfunksjonen etter at disse differensieringene var utført.

Det som tydelig trer fram i figur 13 er to sterkt signifikante autokorrelasjoner, en i avstand 1 og en i avstand 12. Formen på disse gir klare indikasjoner på en MA(1)-prosess for den ikke-sesongavhengige delen av tidsrekken og i tillegg en MA(1)-prosess for den sesongavhengige delen av tidsrekken. Bruker vi notasjonen i kapittel 5.2 gir dette forslag til en modell av formen ARIMA (0,1,1)X(0,1,1). Alternativt kan dette skrives som:

$$w_t = a_t - \theta_1 a_{t-1} - \Theta_1 a_{t-12} \quad (18)$$

Her er w_t den transformerte og differensiererte utgaven av mottatte bestillinger.



Figur 16 Teoretiske korrelasjonsfunksjoner for MA-prosesser uten sesongsvingninger

I tillegg ser vi to andre signifikante autokorrelasjoner, i avstand 32 og 44. Disse må i imidlertid anses som tilfeldige. Husk at normalt vil 1 av 20 være tilfeldige, se kapittel 6.3. Dette tillater ca 2 – 3 tilfeldige signifikante autokorrelasjoner når vi ser på 50 av dem, slik som i figur 13.

I figur 18 ser vi den partielle autokorrelasjonsfunksjonen for den samme tidsrekken som i figur 13. Vi ser tydelig den avtagende eksponentielle formen i avstandene 1 – 8. En slik “kort” avtagende eksponentiell form er en indikasjon på en MA(1)-prosess, se figur 16.

Går vi til sesong-verdiene er det ikke helt klart at disse indikerer en MA(1)-prosess slik som vi så i autokorrelasjonsfunksjonen. I dette tilfellet ville det derfor være lurt å modellere og estimere en modell uten prosesser for sesongbeskrivelse, og deretter se på autokorrelasjonsfunksjonen og den partielle autokorrelasjonsfunksjonen for de “nye” restleddene. I mange tilfeller vil eventuelle tilslørte prosesser da tre tydeligere fram.

7 Estimering

Estimeringen finner optimale⁷ verdier til de identifiserte parametrene i en modell. I Box-Jenkins metode er dette en iterativ prosess som minimerer kvadratsummen av restleddene, dvs a_t . For å utføre dette brukes en ikke-lineær minste kvadraters algoritme. Den iterative prosessen stopper når endringen i parameterverdiene oppnår en på forhånd satt minimumsverdi.

Vi har ikke til hensikt å gå inn på selve metoden her. Den interesserte leser henvises til [2].

7.1 Eksempel

I kapittel 6.4 identifiserte vi modellen (18) ved hjelp av autokorrelasjonsfunksjonen. Bedømmelse av den partielle autokorrelasjonsfunksjonen gav imidlertid ikke helt klarhet i hvorvidt sesongparameteren θ_1 burde innlemmes i modellen. Vi estimerer derfor først modellen

$$w_t = a_t - \theta_1 a_{t-1} \quad (19)$$

Estimeringsresultatet gir verdien 0,8545 for θ_1 , med et estimert standardavvik på

0,0567. Innsatt i (16) gir dette følgende modell:

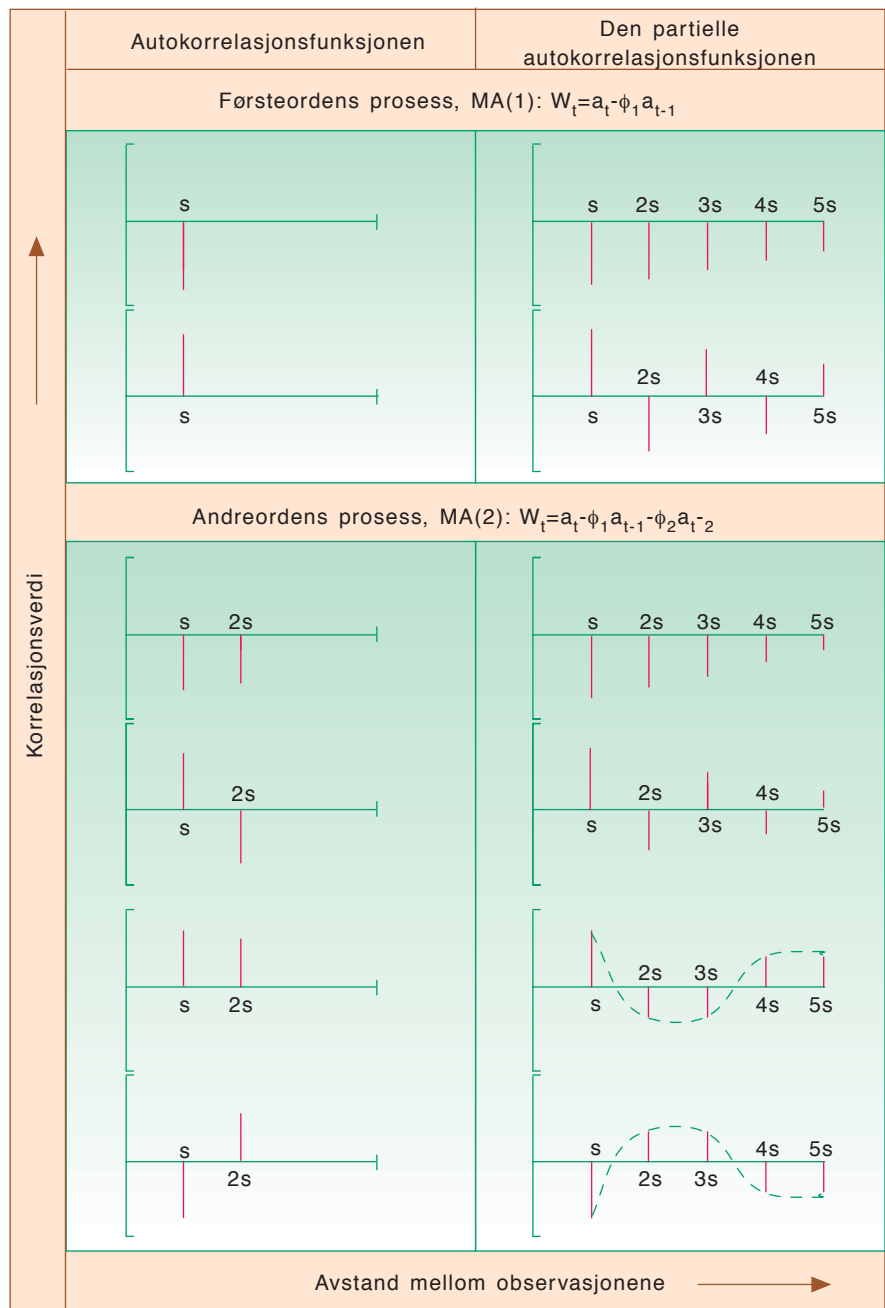
$$w_t = a_t - 0,8545a_{t-1} \quad (20)$$

8 Modellsjekkning

Målet med modellsjekkning er å avdekke hvorvidt den estimerte modellen gir en god beskrivelse av tidsrekken utvikling og hvor god den er til å lage prognoser med. Vi skal her se på de viktigste.

8.1 Parameterestimertene og de estimerte standardavvik

Det er i utgangspunktet ønskelig at alle parameterestimater er statistisk signifikante. En vanlig måte å teste dette på er ved de såkalte *t-verdier*, som er verdien av parameterestimertet dividert med det estimerte standardavvik. *En tommelfingerregel for den såkalte t-testen sier da:*



Figur 17 Teoretiske korrelasjonsfunksjoner for MA-prosesser som beskriver sesongsvingninger

⁷ Optimale i forhold til den tidsrekken vi forsøker å modellere.

Aksepter hypotesen om at parameterestimaten er signifikant forskjellig fra 0 hvis t -verdien er større enn 2^8 .

En annen måte å se denne testen på er å ta utgangspunkt i parameterens estimerte verdi og så trekke fra og legge til to ganger det estimerte standardavviket. Hvis det intervallet som da framkommer ikke dekker verdien 0, er estimatet signifikant forskjellig fra 0.

I noen tilfeller kan det være ikke-statistiske grunner til at en eller flere parametre beholdes i modellen selv om estimatene ikke er signifikante. Et eksempel på dette kan være at vi forsøker å modellere en syklisk utvikling. Dette kan gjøres ved hjelp av sesongavhengig AR(2)-prosess når autokorrelasjonsfunksjonen viser en dempet sinus-kurve. Parameterestimaten er da gitt ved komplekse røtter. Dette går vi ikke nærmere inn på her. Leseren henvises til [2] og [3].

8.1.1 Eksempel

La oss se på resultatene i modell (20). Dividerer vi parameterestimaten (0,8545) med det estimerte standardavviket (0,0567) får vi 15,1, dvs en t -verdi langt over 2. Vi kan nå være trygge på at estimatet er signifikant forskjellig fra 0.

8.2 Støyleddenes korrelasjonsfunksjoner

En av forutsetningene for støyleddene, a_t , var at disse skulle være uavhengige av hverandre, ref kapittel 2. Det er to måter å teste hvorvidt dette er tilfellet:

1. Ved hjelp av autokorrelasjonsfunksjonen og den partielle autokorrelasjonsfunksjonen. Hvis modellen beskriver avhengigheten i den stasjonære tidsrekken, skal det ikke være noen indikasjon på autokorrelasjon i restleddene. Unntaket her er tilfeldige signifikante autokorrelasjoner (1 av 20) og autokorrelasjoner som har stor tallverdi pga store støyledd, se kapittel 6.3.

2. Ved hjelp av en test på alle autokorrelasjonene samlet. Den er basert på kvadratsummen av støyleddenes autokorrelasjoner og heter *Kji-kvadrat observatoren*⁹:

$$Q = n \sum_{k=1}^k r^2(k) \quad (21)$$

hvor

k = antall autokorrelasjoner, dvs antall lag

n = antall støyledd

$r(k)$ = autokorrelasjonen for avstand k mellom støyleddene.

Under null-hypotesen om at alle autokorrelasjonene er null, vil Q være *kji-kvadrat* fordelt med k -s frihetsgrader. s er her antall estimerte parametre. I stedet for å gå inn på detaljer mht hvordan denne testen korrekt utføres, gir vi også her en *tommelfingerregel*:

Hvis Q er omtrent like stor som antall frihetsgrader, eller lavere, er det ingen indikasjoner på avhengighet i støyleddene¹⁰.

8.2.1 Eksempel

La oss gå tilbake til modell (20) i kapittel 7.1. Vi husker at den partielle autokorrelasjonsfunksjonen ikke gav noen klar indikasjon på hvorvidt det burde innlemmes et MA(1)-ledd for beskrivelse av sesongmønster. Derfor estimerte vi modellen (16). Figur 19 viser autokorrelasjonsfunksjonen for støyleddene.

Vi ser nå meget tydelig den store autokorrelasjonen i avstand 12. Denne indikerer en MA(1)-prosess for sesongsvingningene.

Den partielle autokorrelasjonsfunksjonen viser nå også meget tydelig en avtagende eksponentiell form i sesong-avstand, dvs en indikasjon på en MA(1)-prosess for beskrivelse av sesongsvingningene.

Kji-kvadrat observatoren Q beregnes til 40,8 over de 48 autokorrelasjonene som vi har brukt. Med en estimert parameter gir det 47 frihetsgrader. Siden Q er mindre enn 47, gir altså ikke denne testen indikasjoner på avhengighet i støyleddene.

Vi innlemmer nå parameteren θ_1 , dvs modellen blir $w_t = a_t - \theta_1 a_{t-1} - \theta_1 a_{t-12}$, og estimerer denne. Vi får da

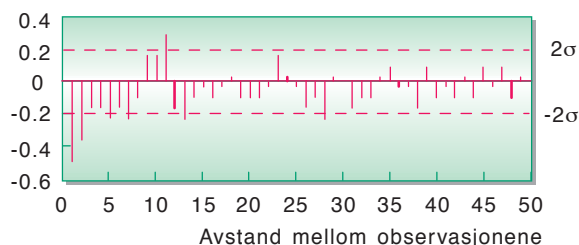
$$w_t = a_t - 0,7966a_{t-1} - 0,8421a_{t-12} \quad (22)$$

(12,0) (21,1)

Tallene i parentes under estimatene er t -verdiene: Begge estimatene er signifikant forskjellig fra null med god margin.

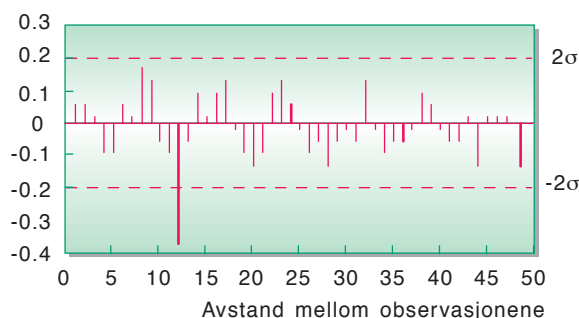
Figur 21 og 22 viser hhv autokorrelasjonsfunksjonen og den partielle autokorrelasjonsfunksjonen til støyleddene i

Den partielle autokorrelasjonsfunksjonen



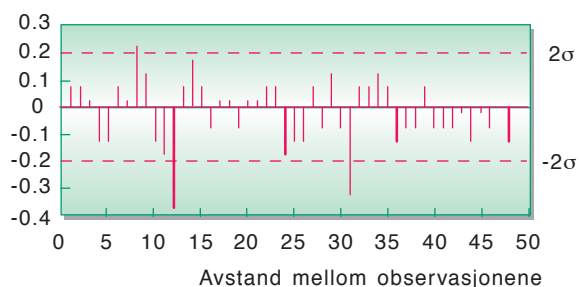
Figur 18 Den partielle autokorrelasjonsfunksjonen for \ln -transformerte verdier av mottatte bestillinger, januar 1984 – desember 1991. Region Oslo. Differensiert en gang og sesong-differensiert en gang

Autokorrelasjonsfunksjonen



Figur 19 Autokorrelasjonsfunksjonen for restleddene i modell (20)

Den partielle autokorrelasjonsfunksjonen



Figur 20 Den partielle autokorrelasjonsfunksjonen for restleddene i modell (20)

⁸ For en test med 95 %-signifikansnivå, dvs at den vil være riktig i 95 av 100 tilfeller, er grenseverdien 2,06 når vi har 25 observasjoner. Når antall observasjoner går mot uendelig, avtar denne grenseverdien mot 1,96.

⁹ Et annet navn som ofte brukes er *Box-Pierce Portmanteau* observatoren.

¹⁰ Dette tilsvarer 50 % av *kji-kvadrat* fordelingen. Dermed er grenseverdien for hvorvidt vi skal akseptere null-hypotesen lavere enn om vi brukte f.eks. 95 % signifikansnivå. *Tommelfinger-testen* er derfor strengere enn den korrekte testen.

(22). Vi ser ingen indikasjoner på signifikante autokorrelasjoner.

Kji-kvadrat observatoren Q beregnes til 21,0. Med 48 autokorrelasjoner og to estimerte parametere har vi 46 frihetsgrader. Testparameteren er altså styrket i forhold til forrige utgave av modellen.

I praksis vil vi ofte også forsøke å estimere modeller på bakgrunn av mønstre i autokorrelasjoner som ikke er signifikante. Vi forfølger imidlertid ikke denne tankegangen her.

8.3 Korrelasjons-matrisen

Det er ønskelig, men slett ikke alltid mulig, å oppnå estimerte modellparametere som ikke er avhengige av hverandre. Korrelasjonsmatrisen er et hjelpemiddel som kan hjelpe oss å avgjøre om dette er tilfellet. Den viser korrelasjonen, målt ved kovariansen mellom parameterestimaterne.

Vi skiller mellom sterk avhengighet og liten avhengighet. I teorien kan det være vanskelig å vite hvor grensen skal settes mellom disse, men [3] opererer med en grense på 0,6. Er korrelasjonen sterkere enn dette vil det virke negativt inn på de parametrene som er korrelerte, og dermed også på modellen.

I tilfeller med korrelasjon mellom parametrene innen en AR- eller en MA-prosess er den mest vanlige løsningen å fjerne en av parametrene – i praksis den med høyest orden – for så å konsentrere informasjonen i data inn i en modell med færre parametere. Dette selv om alle parametrene i utgangspunktet er statistisk signifikante.

I tilfeller med korrelasjon mellom parametrene innen en ARMA-prosess er den mest vanlige løsningen å forsøke å modifisere modellen, noe som kan medføre å fjerne AR- eller MA-delen fullstendig, selv om også i dette tilfellet alle parametere i utgangspunktet er statistisk signifikante. Dette tilfellet av parameter-korrelasjon er for øvrig ofte forbundet med modeller som har for mange parametere i forhold til hva som er nødvendig.

8.3.1 Eksempel

Korrelasjonsmatrisen viser svært lav korrelasjon mellom θ_1 og Θ_1 (-0,01), se tabell 1. Det er altså ikke grunnlag for å endre på modellen ut fra dette.

8.4 Modellens stabilitet

Hvis en modell skal være til praktisk nytte, må den være stabil. Stabilitets-egenskapene ligger i de estimerte parameterverdiene.

For at en modell skal være stabil må parameterverdiene i 1. og 2. ordens prosesser være større enn -1 og mindre enn 1.

At dette gir mening kan vises ved å ta utgangspunkt i f eks en MA(1)-prosess med parameterverdi på 1,1:

$$y_t = a_t - 1,1a_{t-1} \quad (23)$$

Ved å omskrive (20), har vi at

$$a_t = y_t + 1,1a_{t-1}$$

Ved å tilbakedatere a_t en periode har vi at

$$a_{t-1} = y_{t-1} + 1,1a_{t-2} \quad (24)$$

Setter vi inn (21) i (20) får vi

$$y_t = a_t - 1,1(y_{t-1} + 1,1a_{t-1})$$

$$y_t = a_t - 1,1y_{t-1} - 1,1^2a_{t-1} \quad (25)$$

Fortsetter vi med å sette inn for a_{t-1} , a_{t-2} , osv får vi

$$y_t = a_t - 1,1y_{t-1} - 1,1^2y_{t-2} - 1,1^3y_{t-3} - 1,1^4y_{t-4} \dots \quad (26)$$

Det er lett å skjønne at dette er en dårlig modell: Modellen sier at vi skal legge mer og mer vekt på observasjonene jo lengre bakover i tidsrekken vi går, – og det kan umulig være lurt.

8.5 Standardavviket til støyleddene

Det estimerte standardavviket til støyleddene måler usikkerheten i den uforklarte delen av modellen. Denne usikkerheten påvirker usikkerheten i de prognoser som modellen genererer:

Hvis det lages prognoser for en periode fram i tiden, men fra tre ulike startpunkter, vil i gjennomsnitt to av tre prognoser ha en feil som er mindre enn det estimerte standardavviket til støyleddene.

I utgangspunktet søker vi den modellen som har minst usikkerhet. Dermed blir det estimerte standardavviket til støyleddene også et viktig mål for sammenlikning av ulike forslag til modell for en gitt tidsrekke, selv om dette i praksis bare blir ett av flere mål.

8.5.1 Eksempel

Vi går tilbake til modell (20) i kapittel 7.1. I denne ARIMA(0,1,1)X(0,1,0)-modellen har støyleddene et estimert standardavvik på 0,145. Siden variabelen er ln-transformert er det estimerte standardavviket gitt som prosent: I gjennomsnitt vil prognosefeilen i to av tre tilfeller være mindre enn 14,5 % når vi lager prognoser en periode fram.

Til sammenlikning har modell (22), som er en ARIMA(0,1,1)X(0,1,1), et estimert standardavvik på 11,5 %.

Vi oppnådde altså å redusere usikkerheten for prognosene en periode fram med 3 % ved å beskrive sesongutviklingen ved en MA(1)-prosess.

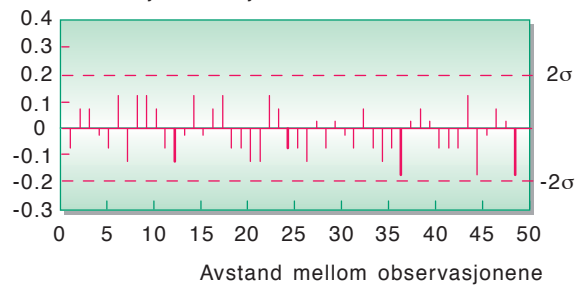
8.6 Modellens støyledd

Under forutsetningene for støyleddet i kapittel 2 har vi til nå konsentrert oss

Tabell 1 Korrelasjonsmatrise for θ_1 og Θ_1 i modell (22)

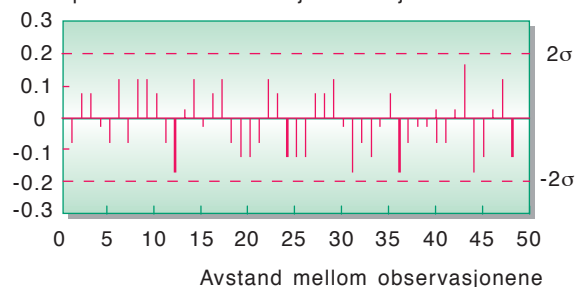
	θ_1	Θ_1
θ_1	1.00	
Θ_1	-0.01	1.00

Autokorrelasjonsfunksjonen



Figur 21 Autokorrelasjonsfunksjonen for støyleddene i modell (22)

Den partielle autokorrelasjonsfunksjonen



Figur 22 Den partielle autokorrelasjonsfunksjonen for støyleddene i modell (22)

mest om å finne ut av hvorvidt disse er uavhengige. I tillegg skal de imidlertid ha forventning lik 0, ha konstant varians og være Normalfordelt.

1. At støyleddene har forventning lik 0 kan beskrives ved at de i gjennomsnitt er lik null. Hvis dette ikke er tilfellet vil modellen generere prognoser som er skjeve i positiv eller negativ retning.

Gjennomsnittet selv er et estimat for forventningsverdien, med en viss usikkerhet. Derfor vil det i praksis ikke være akkurat lik 0. En test for å avdekke om gjennomsnittet er signifikant forskjellig fra 0 er gitt ved:

Forkast hypotesen om at støyleddenes gjennomsnitt ikke er signifikant forskjellig fra 0 hvis dette gjennomsnittet i tallverdi er større enn

$$2 \frac{\sigma}{\sqrt{N}} \quad (27)$$

der

σ = standardavviket til støyleddene

N = antall støyledd.

2. Som regel er variansen stabilisert under prosessen med å bringe tidsserien over på stasjonær form. I kapittel 5.1 så vi hvordan dette kunne gjøres ved å ln-transformere data. Skulle støyleddene vise seg å ha kontinuerlig økende eller avtagende varians, er dette m a o vanligvis et tegn på uriktig eller manglende datatransformasjon. I praksis bør en da gå tilbake og finne den riktige transformasjon. Alternativt er det i noen programpakker mulig å re-estimere modellen med en transformasjons-parameter, slik at den optimale transformasjonen framkommer.

Det er mulig å utføre tester på støyleddene for å avdekke hvorvidt variansen statistisk kan sies å være konstant. Dette går vi ikke inn på her. En enklere metode er å plote ut støyleddene og visuelt avgjøre hvorvidt variasjonen i støyleddene kan sies å være konstant.

3. Å avgjøre hvorvidt støyleddene er Normalfordelte kan også avgjøres ved hjelp av tester. En enklere metode er å lage et histogram som samler støyleddene i frekvenssøyler på hver side av gjennomsnittet. Hvis formen på histogrammet ikke avviker vesentlig fra en Normalfordelingskurve, er dette en indikasjon på at vi kan se bort fra problemet.
4. I tillegg til disse forutsetningene var vi i kapittel 6.3 inne på at såkalte *outliere* kunne virke forstyrrende på estimerte verdier av parametrene og i verste fall påvirke hele modellstrukturen. Slike unormale observasjoner avdekkes lett ved å plote alle støyleddene i et residualplott. Her plottes verdien til alle støyleddene på en akse med null som basisverdi. I tillegg viser plottet også grensene for to ganger estimert standardavvik – grensen for test av hvorvidt vi har flere store støyledd enn det som kan betraktes som tilfeldig. I praksis må da noe gjøres for at prognosemodellen ikke skal opptre med for stor usikkerhet. Det vanlige er å justere ekstreme observasjoner slik at støyleddet blir omtrent to ganger standardavviket.

8.6.1 Eksempel

Vi tar igjen utgangspunkt i modell (22) og ser først på gjennomsnittet til støyleddene. Det er beregnet til -0,0119, dvs svært nær null. Ser vi på testen som gir uttrykk (27), har vi at for støyleddene

(de transformerte og differensierte) er $\sigma = 0,105$ og $N = 83$. Dette gir

$$2 \frac{\sigma}{\sqrt{N}} = 0,023 \quad (28)$$

Siden støyleddenes gjennomsnitt i tallverdi er mindre enn 0,023 kan vi altså ikke forkaste hypotesen om at dette gjennomsnittet ikke er signifikant forskjellig fra 0. Vi

godtar derfor at -0,0119 er så nær null at forskjellen er neglisjerbar.

I figur 23 er støyleddene plottet. I tillegg er grensene for to ganger støyleddenes estimerte standardavvik avmerket. På bakgrunn av dette er det ikke grunnlag for å si at støyleddene har kontinuerlig økende eller avtagende varians.

Vi ser at det er 4 støyledd som er lik eller større enn to ganger standardavviket. I forhold til de 83 støyleddene er dette innenfor de akseptable 5 %. I praksis bør en likevel undersøke slike observasjoner. Spesielt gjelder dette når støyleddene er så store som støyledd nr 86 (-3,3 σ). Hvis slike støyledd vesentlig påvirker verdiene til de estimerte parametrene eller i verste fall hele modellstrukturen når de reduseres til ca to ganger σ , bør dette vurderes. Vi går ikke nærmere inn på dette her.

Histogrammet for støyleddene viser at disse følger den typiske klokkeformen til en Normalfordeling. Det er m a o ingen grunn til å si annet enn at støyleddene oppfyller forutsetningen om Normalfordeling.

9 Prognoser

Vi tar igjen utgangspunkt i modell (22):

$$w_t = a_t - 0,7966a_{t-1} - 0,8421a_{t-12} \quad (22)$$

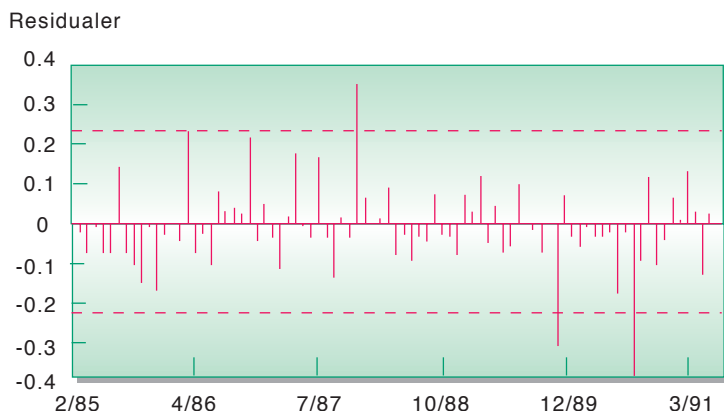
La oss anta at vi skal lage prognose for periode $t+1$. Da har vi at

$$w_{t+1} = a_{t+1} - 0,7966a_t - 0,8421a_{t-11} \quad (29)$$

Vi ser her at det eneste leddet som er ukjent på likningens høyre side er a_{t+1} . Siden vi ikke kjenner verdien på framtidig støyledd må vi legge inn en forutsetning om verdien på a_{t+1} . Under analysen som ledet fram til (22) forsikret vi oss om at alle støyleddene var tilfeldig fordelt (dvs ingen avhengighet mellom dem), at de hadde forventning lik null og at de var Normalfordelt. Skal vi gjette på en verdi på framtidige støyledd må det beste anslaget derfor bli 0, gitt at disse også i framtiden beholder de samme egenskaper. Prognosen en periode fram, innsatt $a_{t+1} = 0$, blir derfor generert av:

$$w_{t+1} = -0,7966a_t - 0,8421a_{t-11} \quad (30)$$

Ønsker vi å lage prognoser for flere perioder fram er framgangsmåten tilsvarende: Vi setter inn forventet verdi lik null for alle framtidige støyledd.



Figur 23 Residualplott for modell (22)

I modell (22), som kun har MA-parametere, ser vi lett at prognosen for periode $t+13$ reduserer seg til

$$w_{t+13} = 0 \quad (31)$$

Det betyr ikke at prognosen for periode 13 og høyere er lik null: Når prognosene skal lages, må selvfølgelig w i alle perioder "nøstes" tilbake til y , dvs at differensieringene føres tilbake til den opprinnelige og ikke-differensierte form og ln-transformasjonen transformeres tilbake til de opprinnelige data. Sistnevnte gjøres ved å opphøye de ln-transformerte data med $e = 2,7183 \dots$ som grunntall. Dette regnestykket utføres normalt av statistikkpakkene.

For en modell med AR-parametere vil vi også komme i den situasjon at framtidige verdier av forklaringsvariabelen inngår. Dette er imidlertid løst av det faktum at disse alltid vil være generert i prognosen for 1 eller 12 perioder (for sesongparametere) tidligere:

Anta at vi har modellen

$$w_t = 0,5w_{t-1} + a_t \quad (32)$$

Når vi setter $a_{t+1} = 0$, $a_{t+2} = 0$, osv, er

$$w_{t+1} = 0,5w_t \quad (33)$$

$$w_{t+2} = 0,5w_{t+1} \quad (34)$$

·
·
·
osv.

Selve modellen ARIMA(0,1,1)x(0,1,1) har for øvrig fått navnet AIRLINE-modellen. Dette er fordi den første gang ble benyttet i en analyse som Box og Jenkins utførte, der de modellerte passasjertrafikk med fly.

9.1 Usikkerhet i prognosene

Standardavviket til støyleddene er et mål på usikkerheten i den modellen vi har. I ARIMA-modeller er dette samtidig også et mål for usikkerheten i prognosen en periode fram. Ved å bygge videre på dette er det også mulig å utlede og beregne den forventede usikkerheten i prognosene 2, 3, 4, ..., osv perioder fram. Tilsvarende kan den samlede usikkerheten i f eks summen av 12 og 12 perioders-prognosen også beregnes.

Ut fra dette kan deretter såkalte prediksjonsintervaller beregnes. Disse angir, som en viss sansynlighet, øvre og nedre grense for prognosene.

Det vil imidlertid føre for langt å gå inn på det teoretiske fundamentet for dette i denne artikkelen.

9.2 Simulering

For å få tiltro til en prognosemodell kan det i tillegg være lurt å simulere med modellen. Dette gjøres ved at vi lar modellen lage prognoser for observasjoner som vi allerede kjenner, f eks det siste året. Etterpå kan vi måle hvor godt modellen treffer med prognosene.

Framgangsmåten er parallell med den måten som ulike prognosemodeller basert på ulike metoder blir sammenliknet. Det henvises til [9] som diskuterer dette.

9.3 Eksempel

Vi lar modell (22) lage prognoser for 12 perioder framover, dvs for perioden januar 1992 – desember 1992. Prognosepakken MICRANAL, som for øvrig også er brukt under identifikasjon og estimering i eksemplene ovenfor, skriver også ut øvre og nedre grense for prognosenes usikkerhet, dvs prediksjonsintervall med 80 % signifikansnivå. Dette er referert i tabell 2.

I tillegg gir tabell 2 estimert standardavvik for den usikkerheten som prediksjonsintervallene gjenspeiler. Hvis forutsetningen om Normalfordelte støyledd er oppfylt, viser dette det intervallet hvor vi forventer at 68 % av prognosefeilene vil ligge. Siden prognosevariabelen i (22) er ln-transformert er dette usikkerhetsmålet gitt i prosent. Vi ser tydelig at usikkerheten vokser ettersom vi går lengre og lengre ut i prognoseperioden: Vi forventer, med 68 % sannsynlighet, at prognosefeilen for januar vil ligge +/- 11,5 % fra prognosen på 4023 mottatte bestillinger på HA. Når prognosen for desember gis, er denne usikkerheten vokst til +/- 13,8 %. Usikkerheten vokser altså ikke dramatisk innenfor den valgte prognosehorisonten.

I bunnen er prognosen for summen av prognosene vist med 80 % prediksjonsintervall. Det estimerte standardavvik for prognosefeilen er her gitt i absolutte tall.

10 Avsluttende merknader

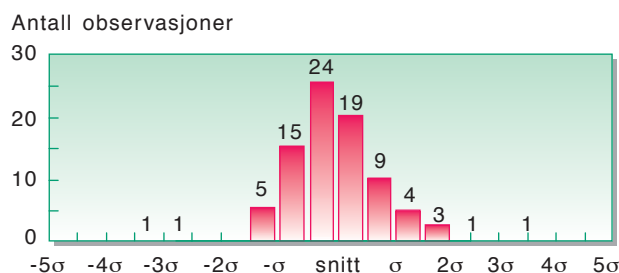
Innledningsvis ble det sagt at ARIMA-modellene er et godt redskap når de baseres på prinsippet om at få parametere skal anvendes. Dette er faktisk et viktig prinsipp i

alle prognosemodeller: En modell som adekvat beskriver data med et minimum av parametere. Sammenhengen mellom AR- og MA-prosessene som ble nevnt i innledningen kommer altså til sin fulle rett i denne sammenheng. Når vi i tillegg kan beskrive en tidsrekke ved hjelp av kombinerte prosesser (ARMA-prosesser), utgjør disse modellene et meget fleksibelt og slagkraftig verktøy mht beskrivelse av avhengighetene i en tidsrekke.

Det har ikke vært mulig å gå inn på alle finesser og sider ved ARIMA-modellene. Det henvises til gode lærebøker som [3], som går lengre enn det som er mulig i denne artikkelen. Dessuten har Televerkets Forskningsinstitutt utført et forskningsprosjekt hvor prognoser etter denne metoden var tema, se [5].

Det skal også nevnes at Box-Jenkins metode er videreutviklet slik at forklaringsvariable kan trekkes inn. Dette kalles på engelsk "Transfer function"-modeller. Slike modeller muliggjør forbedringer både i nøyaktighet og i forlengelse av prognosehorisonten. En annen videreutvikling er såkalte "Intervention"-modeller, som takler problemer i form av outliere, nivåskift o.l.

ARIMA-modellene er teoretisk svært tiltalende. Dessuten har de i praksis ofte gitt svært gode resultater. Likevel har nok metodens matematisk/statistiske kompleksitet til en viss grad bremset utbredelse og bruk av den. Før statistikkpakker tok metoden inn som standard analyseverktøy, krevde den at brukeren var skolert i relativt komplekse statistisk-metodiske sammenhenger.



Figur 24 Histogram for støyleddene

Referanser

- 1 Stordahl, K, Hjelkrem, C, Bøe, J. Regresjonsmodeller. *Teletronikk* 90(1), 87–102, 1994 (denne utgave).
- 2 Box, G E P, Jenkins, G M. *Time series analysis: forecasting and control*. San Fransisco, Holden Day, 1970.
- 3 McLeod, G. *Box Jenkins in Practice*. Lancaster, Time Series Library, GJP Publications, 1983.
- 4 Makridakis, S, Wheelwright, S C, McGee, V E. *Forecasting: methods and applications, 2nd ed*. New York, John Wily & Sons, 1983.
- 5 Stordahl, K et al. *Korttidsmodeller basert på tidsrekkeanalyse*. Kjeller, Televerkets Forskningsinstitutt, 1979. (TF-rapport 10/79, del I og II.)
- 6 Jenkins, G M, McLeod, G (eds). *Case studies in time series analysis. Volum 1*. Lancaster, Time Series Library, GJP Publications, 1982.
- 7 Bøe, J, Hjelkrem, C. Glattingsmodeller. *Teletronikk* 90(1), 76–86, 1994 (denne utgave).
- 8 Stordahl, K. Prognosemetoder – en oversikt. *Teletronikk* 90(1), 5–52, 1994 (denne utgave).
- 9 Bøe, J, Hjelkrem, C. Sammenlikning av prognosemodeller fra ulike metoder. *Teletronikk* 90(1), 125–133, 1994 (denne utgave).

Tabell 2 Prognoser, prediksjonsintervall og estimert standardavvik for prognosefeilen for mottatte bestillinger av HA i Region Oslo, generert av modell (22)

		Obs. nr	Nedre 80 % prediksjons-grense	Prognose	Øvre 80 % prediksjons-grense	Estimert standardavvik til prognosefeil for transformert variabel
1992	Jan	97	3473,70	4022,57	4658,16	0,114608
	Feb	98	2757,82	3203,18	3720,46	0,116956
	Mar	99	2667,59	3107,51	3619,99	0,119256
	Apr	100	2480,03	2897,39	3384,98	0,121514
	Mai	101	2299,60	2694,22	3156,56	0,123730
	Jun	102	2712,47	3186,81	3744,10	0,125907
	Jul	103	2505,99	2952,30	3478,09	0,128047
	Aug	104	3085,96	3645,37	4306,18	0,130152
	Sep	105	3349,21	3966,84	4698,37	0,132223
	Okt	106	3451,14	4098,25	4866,71	0,134263
	Nov	107	2974,20	3540,98	4215,76	0,136272
	Des	108	2290,01	2733,32	3262,46	0,138252
Sum	1992		36251,5	40048,7	43846,0	2966,62

Sammenlikning av prognosemodeller fra ulike metoder

AV JOHANNES BØE OG CARLO HJELKREM

001.18:621.39

1 Innledning

Det er ingen enkel sak å se forskjell på gode og dårlige prognoser, prognosemodeller eller prognosemetoder. Spesielt vanskelig kan dette være når resonnementene bak prognosemodellene er teoretisk velfunderte. Vi skal her vise hvordan det er mulig å kvalitetsbestemme kvantitative prognoser som er utarbeidet ved hjelp av modeller basert på ulike kvantitative prognosemetoder.

Prognosemodeller bør kvalitetsbestemmes etter tre viktige hovedkriterier, og en god prognose bør være fundamentert i en modell som har høy karakter på alle tre.

1. *Prognosemodellen skal inneholde et teoretisk og statistisk grunnlag.* Dette vil si at prognosemodellen bør tilfredsstillende anerkjente teorier for den utvikling som modellene skal beskrive og dessuten tilfredsstillende krav til statistisk signifikans som prognosemetoden er basert på. Uten slikt grunnlag er god treffsikkerhet ofte bare tilfeldig.
2. *Prognosemodellen skal ha gode historiske resultater å vise til.* Det vil si at den må gi en god beskrivelse av den historiske utvikling og dessuten ha god treffsikkerhet på tidligere prognoser. Sistnevnte kan enten erfares eller simuleres.
3. *Prognosemodellen må være tilpasset brukerne av prognosene.* Dette betyr at prognosemodellen må gi svar på det vi ber om. Samtidig må modellen være "operasjonell" for brukeren, dvs lett å anvende i forhold til tolkning, antall prognoser som skal lages, tid til disposisjon, brukerens kvalifikasjoner m m.

I praksis vil det ofte oppstå konflikter mellom disse hovedegenskapene, og en må derfor inngå kompromisser, selv om dette ikke er ønskelig.

Vi skal her ta utgangspunkt i egenskap nr 2 og se på hvordan vi kan måle ulike prognosemodellers treffsikkerhet når vi skal lage den samme prognosen. Vi skal tenke oss en situasjon hvor vi har utviklet prognosemodeller ved hjelp av ulike metoder og ikke har noen erfaring fra hvor gode prognoser de ulike modellene genererer. Vi må da *simulere treffsikkerheten*.

Vi har valgt å lage prognose for volum tellerskritt i Region Oslo. Tilrettelegging av data for volum tellerskritt med de justeringer som må gjøres, er omtalt i detalj i artikkelen "Tellerskritt – grunnlagsdata

og prognoser" i denne utgaven av Telektronikk. Kort skissert er problemstillingen at tellerskrittens lengde endres fra tid til annen. Dette er en av to måter å endre tellerskrittprisen på, – den andre er å endre selve prisen på tellerskrittet. Endringer i tellerskritt lengden gir imidlertid tydelige hopp i data-rekken for tellerskritt, fordi tellerskrittene i ulike takstklasser som regel er forlenget (dvs at prisen er redusert) uten at dette er kompensert med tilsvarende økning i antall tellerskritt. Slike hopp kan i prinsippet modelleres inn i en prognosemodell ved hjelp av teknikker for beskrivelse av såkalte nivåskift. Imidlertid vil stadige nivåskift gi en så stor modell med så mange parametere at dette reduserer sikkerheten betraktelig. Vi har derfor valgt å regne om alle historiske tellerskrittdata til "volum tellerskritt", med tellerskritt lengdene i 1993 som enhet. Vi får da en tidsserie som er sammenliknbar over tid og uavhengig av historiske endringer i tellerskritt lengden. Omregningen fra tellerskritt til volum tellerskritt tar hensyn til ulike takstklassers vekt, fordeling mellom full og redusert takst og samtale-tidsfordelingen for de ulike takstklasser.

2 Modelltyper og inndata

Vi skal diskutere sju forskjellige modelltyper, hvor tre går under klassen glattingsmodeller og to er rene tidsrekkemodeller. Disse modellene benytter kun den variabel som det skal lages prognoser for som inndata. De siste to er forklaringsmodeller, dvs modeller som i tillegg til den variabel som det skal lages prognoser også benytter forklaringsvariable som inndata. Modelltypene er som følgende:

Modell 1:	Ekspontentiell glatting	Glattingsmodell
Modell 2:	Holts metode	Glattingsmodell
Modell 3:	Holt-Winters metode	Glattingsmodell
Modell 4:	Box-Jenkins metode	Tidsrekkemodell
Modell 5:	Kalmanfilter	Tidsrekkemodell
Modell 6:	Regresjon	Forklaringsmodell
Modell 7:	Transfermodell	Forklaringsmodell

Treffsikkerheten til disse modellene simuleres ved å holde tilbake noen av de siste dataene som er observert og ikke

bruke disse på estimeringen av modellene. Siden det i Televerket som regel lages årsprognoser og vi benytter månedssdata, holder vi tilbake 12 observasjoner og ser hvordan de forskjellige modelltypene klarer seg i forhold til hverandre. Tabell 1 viser data som testingen skal foretas på.

I tillegg til volum tellerskritt og antall abonnement viser tabell 1 også antall virkedager pr måned. Telefonbruken er forskjellig på virkedager i forhold til lørdager og helligdager. Derfor tillegges forskjellig type dager ulik vekt. Vi har målt antall tellerskritt som etterspørres på de ulike hverdager, lørdager, søn- og helligdager og inneklemt virkedager (f eks romjul og dager fra Palmesøndag til Skjærtorsdag). Dette gir de ulike dagene vekt, gitt i tabell 2.

Dette er for øvrig også omtalt i artikkelen "Tellerskritt – grunnlagsdata og prognoser" i denne utgaven av Telektronikk. Vi går derfor ikke nærmere inn på dette her.

3 Hvordan sammenlikne de forskjellige modellene basert på enkle mål for føyning

Vi vil altså simulere modellenes treffsikkerhet ved å holde igjen data (de siste 12 observasjonene) når modellene lager prognose for totalt volum tellerskritt pr måned for perioden august 1992 til og med juli 1993.

I praksis er dette ikke en tilfredsstillende metode for å evaluere prognosenes treffsikkerhet fordi prognosene som regel også er avhengig av inndata som også er prognoser med en gitt usikkerhet. Dette gjelder spesielt forklaringsmodellene, men berører i vårt tilfelle også glattingsmodellene idet disse modellene benytter volum tellerskritt pr abonnement pr virkedag som prognosevariabel: For envariabelmodeller er dette en måte å nyttiggjøre informasjon om forklaringsvariable som vi vet påvirker tellerskrittutviklingen. Deretter fremkommer prognosen for volum tellerskritt ved å multiplisere med abonnement og virkedager i de respektive perioder.

I vår simuleringsperiode (august 1992 – juli 1993) er data for abonnement og virkedager sikre, idet vi benytter observerte data. Dette gjør disse prognosene mer treffsikre enn i et reelt tilfelle, hvor abonnement og virkedager (sikker) selv er prognoser med en gitt usikkerhet.

Tabell 1 Volum tellerskritt, abonnement og virkedager i Region Oslo

Periode	Abonnement	Virke-dager	Volum tellerskritt	Volum ts/ab/vdag
1989	46508	24,65	16949	14,78
februar	46673	22,4	15303	14,64
mars	46815	21,95	15334	14,92
april	46893	23	16911	15,68
mai	46990	22,3	15129	14,44
juni	46970	24,4	15909	13,88
juli	46877	24	12677	11,27
august	47051	25,4	16000	13,39
september	47233	23,7	15803	14,12
oktober	47347	24,7	17397	14,88
november	47544	24,4	17086	14,73
desember	47658	21,6	15324	14,89
1990	47846	24,8	18069	15,23
februar	48045	22,4	16260	15,11
mars	48206	24,7	18049	15,16
april	48296	20,25	14754	15,09
mai	48402	23,15	17119	15,28
juni	48432	22,85	16621	15,02
juli	48437	24,7	13317	11,13
august	48510	25,4	17370	14,10
september	48622	23	16759	14,99
oktober	48763	25,4	18842	15,21
november	49099	24,4	18618	15,54
desember	49172	20,6	15971	15,77
1991	49289	24,8	19702	16,12
februar	49461	22,4	17443	15,74
mars	49604	21,2	17004	16,17
april	49688	23,75	19165	16,24
mai	49824	22,3	18266	16,44
juni	49881	23	18387	16,03
juli	49862	25,4	14864	11,74
august	49952	24,7	18232	14,78
september	50089	23,7	19054	16,05
oktober	50274	25,4	20828	16,31
november	50435	23,7	19677	16,46
desember	50528	21,3	17239	16,02
1992	50697	24,8	20754	16,51
februar	50839	22,7	18899	16,38
mars	50913	24,7	21398	17,02
april	50974	20,95	18076	16,93
mai	51082	22,35	19431	17,02
juni	51150	23,55	19814	16,45
juli	51117	25,4	16892	13,01
august	51322	24	19279	15,65
september	51571	24,4	21025	16,71
oktober	51743	24,7	21457	16,79
november	51845	23,7	21278	17,32
desember	51943	22,2	19540	16,95
1993	52052	23,4	21225	17,43
februar	52359	22,4	20404	17,40
mars	52565	25,4	23376	17,51
april	52648	20,95	19482	17,66
mai	52714	21,45	20375	18,02
juni	52720	24,4	22630	17,59
juli	52721	24,7	17690	13,58

Eliminerer vi denne usikkerheten står vi igjen med den treffsikkerhet som utelukkende skyldes modellen.

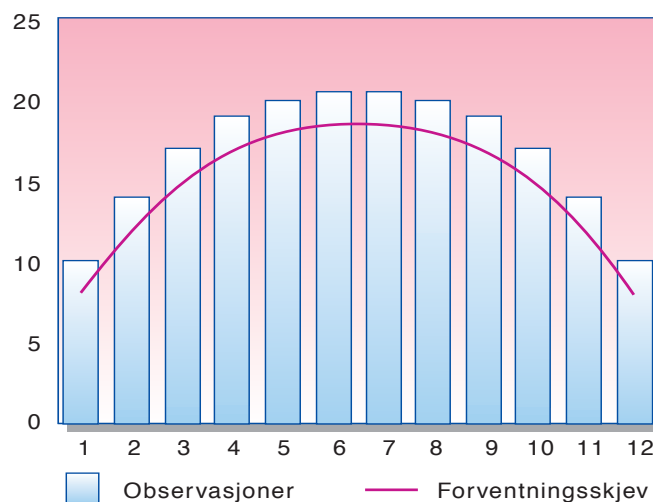
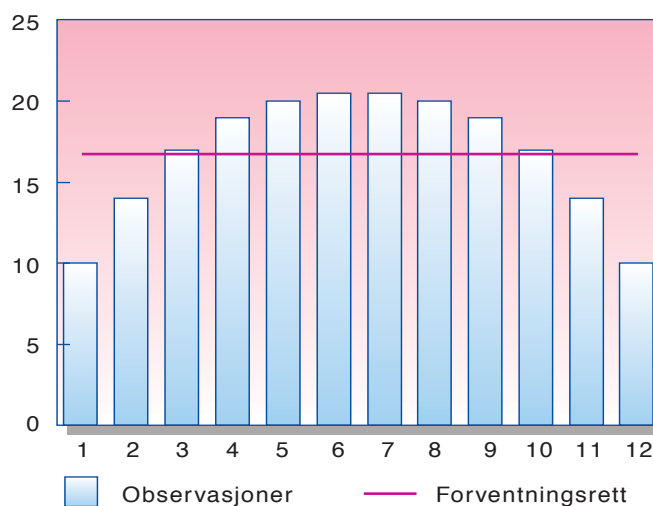
Det kan benyttes mange og ulike mål på modellens treffsikkerheter. Vi skal benytte 6 mål som er vanlige å bruke.

Vi ønsker å finne ut om modellen gjennomsnittlig treffer for lavt eller for høyt. I så fall kan vi si at modellen er forventningsskjev. Dette kan beregnes ved å måle gjennomsnittlig avvik eller gjennomsnittlig prosentvis avvik. Disse er vanligvis forkortet ME (Mean Error) eller MPE

(Mean Percentage Error). Er de positive avvikene i prognoseperioden utlignet av negative avvik, vil summen av avvikene være lik 0. Da er modellen forventningsrett. I motsatt fall er modellen forventningsskjev med positive eller negative avvik.

Tabell 2 Vekting av virkedager

	Vekt:
Lørdag og søndag	0,3
Helligdager og innklemtede virkedager	0,2 – 0,7
Resten	1



At en modell er forventningsrett betyr imidlertid ikke at den er god. Det betyr bare at summen av avvikene er lik 0. Figur 1a viser en forventningsrett prognose. Vi ser der at prognosen har store negative avvik i begynnelsen av perioden og store positive avvik i slutten. Summen av alle avvikene er lik 0. Figur 1b viser en forventningsskjev prognose. Vi kan ut fra tilpasningen anta at prognosen i figur 1b er den beste av disse to. Dette viser at gjennomsnittlig avvik alene ikke er tilstrekkelig til å bestemme hvorvidt en prognose er god eller dårlig.

Et måltall som beskriver det gjennomsnittlige avviket uavhengig av fortegn er MAE (Mean Absolute Error), gjennomsnittlig absolutt feil. Her vil ikke positive og negative avvik utligne hverandre fordi vi tar absoluttverdien av feilene. MAE viser hvor gal prognosen er i gjennomsnitt, uten hensyn til fortegn.

Den gjennomsnittlige prosentvis absolute feil, forkortet MAPE (Mean Absolute Percentage Error) angir den gjennomsnittlige absoluttverdien av feilen i forhold til nivået på data.

Et alternativ til MAE er MSE (Mean Squared Error), gjennomsnittlig kvadratfeil, som kvadrerer avvikene før gjennomsnittet beregnes. Siden kvadratet av både positive og negative tall blir positivt, utligner heller ikke her positive og negative avvik hverandre. Den store forskjellen mellom MSE og MAE er behandlingen av store avvik. Når et stort avvik kvadreres blir resultatet svært stort. Dette betyr at MSE får forholdsmessig større utslag når avvikene er store.

For å bøte på de store avvikenes betydning kan man ta roten av MSE og vi får et nytt måltall, nemlig RMSE (Root Mean Squared Error), kvadratroten av gjennomsnittlig kvadratfeil. RMSE er vanligvis mer meningsfylt enn MSE fordi det er målt i samme enhet som MAE og ME. Som en grov tommelfingerregel kan det (med 95 % sannsynlighet) forventes at et avvik like stort som RMSE forekommer en av tre ganger, et avvik like stort som $2 * RMSE$ forekommer en av tjue ganger.

Tabell 3 viser måleparametrene og hvordan disse beregnes for å kunne skille forskjellige modeller og metoder og kategorisere disse som gode eller dårlige sett i forhold til hverandre.

Her er:

$$y_t = \text{volum tellerskritt på tidspunkt } t$$

$$\hat{y}_t = (\text{modellens}) \text{ prognose for volum tellerskritt på tidspunkt } t$$

Forklaring til verdsettene:

ME: Gjennomsnittlig feil i prognose som modellen gir i forhold til observerte verdier. Benyttes for å vise om det er en skjevhet i prognosene. Bør være så nær 0 som mulig. Positiv feil gir for lave prognoser, mens negativ feil gir for høye prognoser.

MPE: Gjennomsnittlig prosentvis feil i prognoser i forhold til observerte verdier. Ellers lik ME.

MSE: Gjennomsnittlig kvadratavvik mellom prognose og observerte verdier. Et mål på nøyaktigheten. Bør være så lavt som mulig, jo større verdier desto mer unøyaktig er modellprognosene. Store avvik veier tungt.

RMSE: Kvadratroten av gjennomsnittlig kvadratavvik mellom prognoser og observerte verdier. Store avvik veier ikke tungt. Ellers lik MSE.

MAE: Gjennomsnittlig absolutt feil i estimerte verdier i forhold til observerte verdier. Også et mål på nøyaktigheten. Bør være så lavt som mulig, jo større verdier desto mer unøyaktig er modellprognosene.

MAPE: Gjennomsnittlig prosentvis absolute feil i estimerte verdier i forhold til observerte verdier. Ellers lik MAE.

4 Modeller og prognoser

Her refereres kort de ulike modellene og hvilke prognoser disse genererer. Alle modeller utnytter den samme informasjon, gitt i tabell 1. Alle glattingsmodellene og tidsrekkemodellene benytter således volum tellerskritt pr abonnement pr virkedag som prognosevariabel, mens transfer- og regresjonsmodellen benytter volum tellerskritt som avhengig variabel og abonnement og virkedager som forklaringsvariable.

For eksponentiell glatting, Holts metode og Holt-Winters metode må det settes såkalte glattings-parametere. Disse sier hvordan nye og gamle observasjoner skal vektlegges i modellen. Disse glattings-parametrene er valgt ut fra skjønn. Kriteriet har vært balansegangen mellom en god beskrivelse av historisk utvikling og fornuftige prognoser.

For transfer- og regresjonsmodellen er det ikke lagt vesentlig vekt på å utvikle de "best mulige modeller" innen sin klasse. Modellene er således valgt ut fra

at de gir en rimelig god beskrivelse av den historiske utvikling og samtidig tilfredsstillende de mest grunnleggende krav til statistisk signifikans.

Alle metodene, unntatt Transfer-modeller, er behandlet i egne artikler i dette nummer av Teletronikk. Vi velger derfor kun å presentere modellene og prognoseresultatene uten å gå inn på den beregnings-tekniske delen av modellene.

4.1 Eksponentiell glatting

Prognosemodellen er gitt ved en oppdateringslikning (1) og en prognoselikning (2):

$$\hat{y}_t = 0,9 \hat{y}_{t-1} + 0,1 y_t \quad (1)$$

$$F_{t+k} = \hat{y}_t \quad (2)$$

Her er:

y_t = observert verdi av volum tellerskritt pr abonnement pr virkedag, tidspunkt t

\hat{y}_t = estimat for y_t

\hat{y}_{t-1} = estimat for y_{t-1}

F_{t+k} = prognose k perioder frem ($k = 1, 2, \dots, 12$)

Konstanten 0,9 i (1) kalles glattingsparameter, og er den vekt som legges på historikken, mens konstanten 0,1 er den vekt som legges på siste observasjon. En annen tolkning av glattingsparameteren er en beskrivelse av hvordan nivået på

Tabell 3 Målekriterier og hvordan disse blir beregnet

Målekriterium	Formel
ME	$1 / 12 * \sum_{t=1}^{12} (y_t - \hat{y}_t)$
MPE	$100 / 12 * \sum_{t=1}^{12} [(y_t - \hat{y}_t) / y_t]$
MSE	$1 / 12 * \sum_{t=1}^{12} (y_t - \hat{y}_t)^2$
RMSE	$SQRT[1 / 12 * \sum_{t=1}^{12} (y_t - \hat{y}_t)^2]$
MAE	$1 / 12 * \sum_{t=1}^{12} [ABS(y_t - \hat{y}_t)]$
MAPE	$100 / 12 * \sum_{t=1}^{12} [ABS(y_t - \hat{y}_t) / y_t]$

Tabell 3a Prognose fra modell basert på eksponentiell glatting og faktiske observasjoner

Periode	Abonnement	Virkedager	Volum ts/ab/vdag	Volum tellerskritt	Volum tellerskritt
1992			Prognose	Prognose	Observasjon
august	51322	24	15,73	19375	19279
september	51571	24,4	15,73	19794	21025
oktober	51743	24,7	15,73	20104	21457
november	51845	23,7	15,73	19328	21278
desember	51943	22,2	15,73	18139	19540
1993	52052	23,4	15,73	19160	21225
februar	52359	22,4	15,73	18449	20404
mars	52565	25,4	15,73	21002	23376
april	52648	20,95	15,73	17350	19482
mai	52714	21,45	15,73	17786	20375
juni	52720	24,4	15,73	20235	22630
juli	52721	24,7	15,73	20484	17690

Tabell 3b Testparametere for modell basert på eksponentiell glatting

ME	MPE	MSE	RMSE	MAE	MAPE
1380	6,33	3970055	1992	1862	9,04

Tabell 4a Prognose fra modell basert på Holts metode og faktiske observasjoner

Periode	Abonnement	Virkedager	Volum ts/ab/vdag	Volum tellerskritt	Volum tellerskritt
1992			Prognose	Prognose	Observasjon
august	51322	24	16,63	20484	19279
september	51571	24,4	16,66	20964	21025
oktober	51743	24,7	16,69	21331	21457
november	51845	23,7	16,72	20544	21278
desember	51943	22,2	16,75	19315	19540
1993	52052	23,4	16,79	20451	21225
februar	52359	22,4	16,82	19727	20404
mars	52565	25,4	16,85	22497	23376
april	52648	20,95	16,88	18618	19482
mai	52714	21,45	16,91	19120	20375
juni	52720	24,4	16,94	21791	22630
juli	52721	24,7	16,97	22098	17690

Tabell 4b Testparametere for modell basert på Holts metode

ME	MPE	MSE	RMSE	MAE	MAPE
-68	-0,06	2195517	1482	1004	5,14

dataene endres over tid. Verdien 0,9 er satt skjønnsmessig.

Tabell 3a viser prognose og observasjonsdata som testingen skal foretas på.

Ser vi på modellen, gitt ved likningene (1) og (2), ser vi raskt at eksponentiell glatting er en enkel metode som kun forklarer nivået i tidsrekken, mens volum tellerskritt pr abonnement pr virkedag, som er den variabelen som modelleres her, er en tidsrekke som også består av vekst og sesongsvingninger. Ved å

benytte en glattingsparameter med verdi nær 1,0 vil modellen gi en brukbar beskrivelse av den historiske utvikling (dog ligge noe under trenden), men kun angi prognosene som det siste estimat for nivået. Dette kan vi se på prognosen for volum tellerskritt pr abonnement pr virkedag som er 15,73 i alle perioder, se tabell 3a. Imidlertid blir trend og sesongsvingninger indirekte trukket inn i prognosen for volum tellerskritt ved at 15,73 er multiplisert med antall abonne-

ment og antall virkedager i de ulike periodene.

Tabell 3b viser testparametrene.

Den høye verdien på ME viser klart at prognosene ligger for lavt i forhold til observerte verdier, se formelen i kapittel 2. Den er m a o ikke forventningsrett. I snitt ligger den 6,33 % under observerte verdier (MPE). Dette forholdet blir forstørret i MSE.

4.2 Holts metode

Prognosemodellen er gitt ved to oppdateringslikninger, (1) og (2), og en prognoselikning (3):

$$\hat{\mu}_t = 0,9(\hat{\mu}_{t-1} + \hat{\beta}_{t-1}) + 0,1y_t \quad (1)$$

$$\hat{\beta}_t = 0,8\hat{\beta}_{t-1} + 0,2(\hat{\mu}_t - \hat{\mu}_{t-1}) \quad (2)$$

$$F_{t+k} = \hat{\mu}_t + k\hat{\beta}_t \quad (3)$$

Her er:

y_t = observert verdi av volum tellerskritt pr abonnement pr virkedag, tidspunkt t

$\hat{\mu}_t$ = estimat for tidsrekkens nivå, tidspunkt t

$\hat{\mu}_{t-1}$ = estimat for tidsrekkens nivå, tidspunkt $t-1$

$\hat{\beta}_t$ = estimat for tidsrekkens stigning, tidspunkt t

$\hat{\beta}_{t-1}$ = estimat for tidsrekkens stigning, tidspunkt $t-1$

F_{t+k} = prognosen, k perioder frem ($k = 1, 2, \dots, 12$)

Konstantene 0,9 i (1) og 0,8 (2) kalles glattingsparametere. I (1), som estimerer tidsrekkens nivå, uttrykker konstanten den vekt som legges på historiske estimater for nivået, mens konstanten 0,1 da uttrykker den vekt som legges på ny informasjon om nivået. I (2), som estimerer tidsrekkens stigning, uttrykker konstanten den vekt som legges på historiske estimater for stigningen, mens konstanten 0,2 da uttrykker den vekt som legges på ny informasjon om stigningen.

En annen tolkning av glattingsparametrene er en beskrivelse av hvordan nivået på dataene endres over tid (1) og hvordan stigningen i dataene endres over tid (2). Verdierne 0,9 og 0,8 er satt skjønnsmessig, både i likning (1) og i likning (2).

Tabell 4a viser prognose og observasjonsdata som testingen skal foretas på.

I denne modellen forklares også trend, gitt i likning (2). Vi ser da også at prognosen for volum tellerskritt pr abonnemement pr virkedag er voksende i alle perioder. Når denne prognosen multipliseres med abonnemement og virkedager bidrar dette til at prognosen for volum tellerskritt ligger nærmere de observerte verdier enn tilfellet var i modellen basert på eksponentiell glatting.

Tabell 4b viser testparametrene.

Vi ser at samtlige testparametere er forbedret i forhold til modellen basert på eksponentiell glatting. ME viser at modellen i gjennomsnitt lager (litt) for høye prognoser, siden verdien er negativ. Selv om MAE og MAPE nesten er halvert, er likevel ikke MSE og RMSE tilsvarende forbedret. Dette skyldes modellens manglende evne til å beskrive sesongsvingningene for volum tellerskritt pr abonnemement pr virkedag.

4.3 Holt-Winters metode

Prognosemodellen er gitt ved tre oppdateringslikninger, (1), (2) og (3), og en prognoselikning (4):

$$\hat{\mu}_t = 0,9(\hat{\mu}_{t-1} + \hat{\beta}_{t-1}) + 0,1(y_t - \hat{S}_{t-m}) \quad (1)$$

$$\hat{\beta}_t = 0,8\hat{\beta}_{t-1} + 0,2(\hat{\mu}_t - \hat{\mu}_{t-1}) \quad (2)$$

$$\hat{S}_t = 0,7\hat{S}_{t-m} + 0,3(y_t - \hat{\mu}_t) \quad (3)$$

$$F_{t+k} = \hat{\mu}_t + k\hat{\beta}_t + \hat{S}_{t+k-m} \quad (4)$$

Her er:

$y_t, \hat{\mu}_t, \hat{\mu}_{t-1}, \hat{\beta}_t, \hat{\beta}_{t-1}, F_{t+k}$ og k de samme som i Holts modell, se kapittel

3.2. Dessuten er:

\hat{S}_t = estimat for tidsrekken sesongsvingning, tidspunkt t

\hat{S}_{t-m} = estimat for tidsrekken sesongsvingning, tidspunkt $t-m$

\hat{S}_{t+k-m} = prognose for tidsrekken sesongsvingning k perioder frem

m = antall sesongsvingninger i tidsrekken, her 12.

Konstantene 0,9 og 0,1 i (1) og 0,8 og 0,2 i (2) kalles glattingsparametere og har tilsvarende tolkning som konstantene i hhv likning (1) og (2) i modellen basert på Holts metode. I likning (3), som estimerer tidsrekken sesongsvingninger, uttrykker konstanten 0,7 den vekt som legges på historiske estimater for sesongsvingningene, mens konstanten 0,3

Tabell 5a Prognose fra modell basert på Holt-Winters metode og faktiske observasjoner

Periode	Abonnement	Virkedager	Volum ts/ab/vdag	Volum tellerskritt	Volum tellerskritt
1992			Prognose	Prognose	Observasjon
august	51322	24	15,59	19203	19279
september	51571	24,4	16,65	20951	21025
oktober	51743	24,7	17,29	22097	21457
november	51845	23,7	17,39	21368	21278
desember	51943	22,2	17,48	20157	19540
1993	52052	23,4	17,57	21401	21225
februar	52359	22,4	17,40	20407	20404
mars	52565	25,4	17,85	23832	23376
april	52648	20,95	18,21	20085	19482
mai	52714	21,45	17,83	20161	20375
juni	52720	24,4	17,28	22228	22630
juli	52721	24,7	13,47	17541	17690

ME	MPE	MSE	RMSE	MAE	MAPE
-139	-0,67	136829	370	292	1,41

Tabell 5b Testparametere for modell basert på Holt-Winters metode

uttrykker den vekt som legges på ny informasjon om sesongsvingningene.

Verdiene for de tre glattingsparametrene er satt skjønnsmessig.

Tabell 5a viser prognose og observasjonsdata som testingen skal foretas på.

Av prognosene for volum tellerskritt pr abonnemement pr virkedag ser vi at også sesongsvingningene er ivarettatt: Selv om prognosene stiger over hele prognoseperioden under ett, er prognosen for enkelte perioder lavere enn foregående periode i typiske "lav-sesong-måneder" som f eks juni og juli. Dette gir forventning om en prognose som treffer de ulike periodene bedre enn de foregående modellene.

Tabell 5b viser testparametrene.

Som forventet ser vi en radikal forbedring i MSE og RMSE. I tillegg er MAE og MAPE forbedret med en reduksjon på godt over 2/3. Modellen overestimerer imidlertid utviklingen, idet ME og MPE er negative. ME og MPE angir en noe skjev prognose enn Holts metode.

4.4 Box-Jenkins metode

Prognosemodeller bygget med Box-Jenkins metode er basert på at prognosene kan uttrykkes ved tidligere verdier av tidsrekken og støy-ledd (den uforklarte del av modellen). Vektleggingen av tidligere verdier av tidsrekken og støy-ledd skjer ved estimering av parametere, slik at vi til enhver tid opererer med optimale verdier

for vektene. Hvert element i modellen blir valgt ut fra identifikasjon av avhengigheten mellom observasjonene i tidsserien.

Prognosemodellen som vi har valgt er gitt ved:

$$y_{t+k} = y_{t+k-1} + y_{t+k-12} - y_{t+k-13} - 0,78a_{t+k-1} - 0,80a_{t+k-12}$$

Her er:

y_{t+k} = prognose for volum tellerskritt pr abonnemement pr virkedag, tidspunkt $t+k$

a_i = et stokastisk støy-ledd, med visse egenskaper, som representerer den uforklarte del av tidsrekken, på tidspunkt $i, i = t+k-1, t+k-12$.

Konstantene er estimerte parametere.

Prognosene genereres ved at alle fremtidige støy-ledd settes lik 0.

Modellens form med en parameter for støy-leddet på tidspunkt $t-1$ og en parameter for støy-leddet på tidspunkt $t-12$ (samt en differensiering og en sesong-differensiering), har navnet Airline-modellen fordi den opprinnelig ble utviklet for å lage prognoser for flypassasjertrafikk. Den har senere vist seg å passe godt som prognosemodell for bl a teletrafikk.

Tabell 6a viser prognose og observasjonsdata som testingen skal foretas på.

Modeller bygget med denne fleksible metoden kan i utgangspunktet virke vanskelig å tolke. Imidlertid er det slik at

Tabell 6a Prognose fra modell basert på Box-Jenkins metode og faktiske observasjoner

Periode	Abonnement	Virke-dager	Volum ts/ab/vdag	Volum tellerskritt	Volum tellerskritt
1992			Prognose	Prognose	Observasjon
august	51322	24	15,64	19264	19279
september	51571	24,4	16,57	20851	21025
oktober	51743	24,7	16,92	21625	21457
november	51845	23,7	17,11	21024	21278
desember	51943	22,2	17,18	19811	19540
1993	52052	23,4	17,56	21389	21225
februar	52359	22,4	17,33	20325	20404
mars	52565	25,4	17,70	23632	23376
april	52648	20,95	17,84	19677	19482
mai	52714	21,45	17,75	20070	20375
juni	52720	24,4	17,31	22267	22630
juli	52721	24,7	13,68	17814	17690

Tabell 6b Testparametere for modell basert på Box-Jenkins metode

	ME	MPE	MSE	RMSE	MAE	MAPE
1		-0,01	47788	219	197	0,94

Tabell 7a Prognose fra modell basert på Kalmanfilter og faktiske observasjoner

Periode	Abonnement	Virke-dager	Volum ts/ab/vdag	Volum tellerskritt	Volum tellerskritt
1992			Prognose	Prognose	Observasjon
august	51322	24	15,48	19067	19279
september	51571	24,4	16,45	20700	21025
oktober	51743	24,7	16,86	21548	21457
november	51845	23,7	16,97	20851	21278
desember	51943	22,2	16,95	19546	19540
1993	52052	23,4	17,40	21194	21225
februar	52359	22,4	17,21	20185	20404
mars	52565	25,4	17,56	23445	23376
april	52648	20,95	17,73	19556	19482
mai	52714	21,45	17,54	19833	20375
juni	52720	24,4	17,09	21984	22630
juli	52721	24,7	13,53	17619	17690

Tabell 7b Testparametere for modell basert på Kalmanfilter

	ME	MPE	MSE	RMSE	MAE	MAPE
186	186	0,89	93077	305	226	1,08

både nivå, trend og sesong forklares av modellen hvis identifikasjonsprosessen viser at tidsserien har slike (signifikante) komponenter. I prosessen frem til identifikasjon av disse komponentene er det imidlertid muligheter for å "gå seg bort", i motsetning til f.eks Holt-Winters metode, hvor komponentene er fast definert (bortsett fra valg av glattingsparametere). Hvis modellen gir en god beskrivelse av tidsrekken vil vi imidlertid normalt forvente at denne metoden vil

gi bedre prognoser enn modeller basert på Holt-Winters metode.

Tabell 6b viser testparametrene.

Vi ser at samtlige testparametere er radikalt forbedret i forhold til modellen basert på Holt-Winters metode. ME og MPE indikerer en meget nøyaktig prognose, dvs den er forventningsrett. MSE og RMSE er redusert med hhv ca 65 % og 40 %, mens MAE og MAPE er redusert med hhv ca 32 % og 35 % i forhold til modellen basert på Holt-Winters metode.

4.5 Kalmanfilter

Prognosemodeller basert på Kalmanfilter estimeres ut fra en eller flere systemlikninger som angir sammenhengen mellom ulike parametere, samt en eller flere målelikninger som angir usikkerheten i data.

I vår modell estimeres tidsrekken nivå, trend og sesongkomponenter. Det er de estimerte verdiene for nivå og trend i tidsrekken siste observasjon, samt sesongkomponentene i de siste 12 observasjoner, som angir prognosen.

Prognosemodellen er på generell form gitt ved:

$$F_{t+k} = \hat{\mu}_t + k\hat{\beta}_t + \hat{S}_{t+k-m}$$

der F_{t+k} , $\hat{\mu}_t$, $\hat{\beta}_t$, k og \hat{S}_{t+k-m} har samme tolkning som i modellen basert på Holt-Winters metode, bortsett fra at oppdateringslikningene (som ikke er vist her) er andre.

Den estimerte modellen gav følgende verdier (t er siste tidspunkt i tidsrekken, dvs juli 1993):

$$\hat{\mu}_t = 16,35$$

$$\hat{\beta}_t = 0,058$$

$$\hat{S}_t = -3,52$$

$$\hat{S}_{t-1} = 0,10$$

$$\hat{S}_{t-2} = 0,60$$

$$\hat{S}_{t-3} = 0,85$$

$$\hat{S}_{t-4} = 0,74$$

$$\hat{S}_{t-5} = 0,45$$

$$\hat{S}_{t-6} = 0,70$$

$$\hat{S}_{t-7} = 0,31$$

$$\hat{S}_{t-8} = 0,39$$

$$\hat{S}_{t-9} = 0,33$$

$$\hat{S}_{t-10} = -0,02$$

$$\hat{S}_{t-11} = -0,93$$

Estimat for S_{t-12} er utelatt fordi alle de andre sesongestimatene er definert som endringer i forhold til denne.

Tabell 7a viser prognose og observasjonsdata som testingen skal foretas på.

Kalmanfilter utgjør en modellklasse som er meget fleksibel, idet det kan vises at glattingsmodellene, Box-Jenkins metode og regresjonsmodeller (se 3.6) er spesielt tilfeller av denne modelltypen. Vi forventer derfor at denne modellen skal være en skarp konkurrent til f eks modellen basert på Box-Jenkins metode (og regresjonsmodellen, se 3.6).

Tabell 7b viser testparametrene.

Vi ser at ME og MPE indikerer en prognose som er noe for lav i forhold til observerte data. MSE er nesten dobbelt så stor, mens RMSE er nesten 40 % større enn for modellen basert på Box-Jenkins metode. MAE og MAPE ligger begge ca 15 % over modellen basert på Box-Jenkins metode.

4.6 Regresjon

Prognosemodeller bygget med regresjonsanalyse er basert på at prognosene kan forklares med forklaringsvariable. Dette betyr at vi nå skifter fra envariabelmodeller til flervariabelmodeller, samtidig som prognosemodellene direkte gir prognoser for volum tellerskritt.

Valg av forklaringsvariable er i vårt tilfelle gitt ved datasettet gitt i tabell 1, dvs abonnement og virkedager. I tillegg innlemmer vi såkalte dummy-variable for å forklare tidsrekkens sesong-svingninger. Dette er variable som er 1 i den sesongen vi vil forklare og 0 ellers.

Vektleggingen av de ulike forklaringsvariablene estimeres, slik at vi til enhver tid opererer med optimale verdier. Forklaringsvariable som ikke tilfredsstiller krav til statistisk signifikans er utelatt (ref noen av sesong-dummiene).

Prognosemodellen er gitt ved:

$$y_{t+k} = -41755,41 + 0,90Ab_{t+k} + 674,46Vdag_{t+k} - 532,03S_6 - 4706,92S_7 - 1628,95S_8 - 606,87S_9 - 956,36S_{12}$$

Her er:

y_{t+k} = prognosen for volum tellerskritt i periode k ($k = 1, 2, \dots, 12$)

Ab_{t+k} = prognosen for abonnement i periode k ($k = 1, 2, \dots, 12$)

$Vdag_{t+k}$ = prognosen for virkedager i periode k ($k = 1, 2, \dots, 12$)

S_i = sesongdummy for periode i , $i = 6, 7, 8, 9, 12$

Konstantene er estimerte parametre.

Tabell 8a Prognose fra modell basert på regresjon og faktiske observasjoner

Periode	Abonnement	Virke-dager	Volum tellerskritt	Volum tellerskritt
1992			Prognose	Observasjon
august	51322	24	18937	19279
september	51571	24,4	20452	21025
oktober	51743	24,7	21416	21457
november	51845	23,7	20833	21278
desember	51943	22,2	18955	19540
1993	52052	23,4	20818	21225
februar	52359	22,4	20419	20404
mars	52565	25,4	22626	23376
april	52648	20,95	19702	19482
mai	52714	21,45	20098	20375
juni	52720	24,4	21560	22630
juli	52721	24,7	17588	17690

ME	MPE	MSE	RMSE	MAE	MAPE
363	1,70	249704	500	402	1,90

Tabell 8b Testparametere for modell basert på regresjon

Tabell 8a viser prognose og observasjonsdata som testingen skal foretas på.

Tabell 8b viser testparametrene.

Vi ser at ME og MPE er ca 100 % høyere enn for modellen basert på Kalmanfilter, dvs en noe forventningsskjev prognose i negativ retning (modellen gir for lave prognoser). MSE og RMSE er hhv over 160 % og ca 65 % høyere enn for prognosene basert på Kalmanfilter, mens både MAE og MAPE tilsvarende er ca 80 % høyere.

4.7 Transfer-modell

Transfer-modeller er egentlig en form for videreutvikling av Box-Jenkins metode for envariabelmodeller, ved at forklaringsvariable trekkes inn. Den delen av tidsrekken som forklaringsvariablene ikke forklarer, modelleres på samme måte som i Box-Jenkins metode for envariabelmodeller.

Modellen er gitt ved:

$$y_t = 0,90Ab_t + 669,0Vdag_t + (a_t - 0,80a_{t-1})/\Delta_{12}$$

Her er Ab_t og $Vdag_t$ som i regresjonsmodellen i kapittel 3.6, mens a_t og a_{t-1} er som i Box-Jenkins modell i kapittel 3.4. At støyleddene er dividert med Δ_{12} betyr her at den delen av tidsrekken som forklaringsvariablene ikke forklarer er sesong-differensiert en gang.

Tabell 9a viser prognose og observasjonsdata som testingen skal foretas på.

Normalt forventes at prognosemodellene blir mer nøyaktige når vi går fra Box-Jenkins envariabelmodeller til transfermodeller. Her blir imidlertid ikke modellen i kapittel 3.4 sammenliknbar med denne modellen, idet denne modellen direkte er modellert med volum tellerskritt som prognosevariabel, mens modellen i kapittel 3.4 er modellert med volum tellerskritt pr abonnement pr virkedag som prognosevariabel.

Tabell 9b viser testparametrene.

Vi ser at ME og MPE er hhv ca 12 % og 7 % høyere enn regresjonsmodellen, dvs denne modellen genererer prognoser som er enda litt lavere enn den faktiske utvikling. MSE og RMSE er hhv ca 245 % og 86 % høyere/lavere enn for modellen basert på regresjonsanalyse, mens MAE og MAPE tilsvarende er ca 70 % høyere.

5 Sammenlikning av modellene

I kapittel 3 sammenliknet vi hver modell med den forutgående. Vi summerer her opp testparametrene for alle modellene.

Tabell 10 viser en oppsummering av modellenes treffsikkerhet i form av testparametere.

Den desidert beste prognosemodellen er modellen basert på Box-Jenkins metode:

Tabell 9a Prognose fra modell basert på transfer-modell og faktiske observasjoner

Periode	Abonnement	Virkedager	Volum tellerskritt	Volum tellerskritt
1992			Prognose	Observasjon
august	51322	24	19322	19279
september	51571	24,4	19830	21025
oktober	51743	24,7	21633	21457
november	51845	23,7	20859	21278
desember	51943	22,2	18782	19540
1993	52052	23,4	21954	21225
februar	52359	22,4	20054	20404
mars	52565	25,4	21415	23376
april	52648	20,95	20017	19482
mai	52714	21,45	20564	20375
juni	52720	24,4	20812	22630
juli	52721	24,7	17643	17690

Tabell 9b Testparametere for modell basert på transfer-modell

ME	MPE	MSE	RMSE	MAE	MAPE
406	1,82	861806	928	685	3,17

inkludert i regresjonsmodellen viser da også betraktelig bedre resultater for regresjonsmodellen: 0,59 %-poeng lavere MPE og 0,46 %-poeng lavere MAPE.

At transfer-modellen kommer dårligere ut enn modellen basert på regresjonsanalyse var imidlertid overraskende. Sammenlikner vi modellene basert på de to metodene ser vi at de estimerte parametrene for hhv abonnement og virkedagers påvirkning på volum tellerskritt er svært like. Ulikheten ligger dermed i transfermodellens modellering av den delen av tidsrekken som ikke forklares av forklaringsvariablene, dvs støyen. Det er mulig at datamengden er noe for liten til å identifisere strukturen i støyen. Anbefalt datamengde for identifikasjon er minimum 24 pluss to ganger sesong, dvs 48 observasjoner. Vi har 43 observasjoner til rådighet. Forsøk på modellering ved bruk av alle observasjonene frem til 30. juli 1993, dvs 55 observasjoner, gav imidlertid bare små forbedringer. Som nevnt under kapittel 3.4 er det muligheter for å "gå seg bort" under modelleringen, – det kan være at en grundigere analyse av støystrukturen ville gitt bedre resultater.

At modellen basert på eksponentiell glatting fallt dårligst ut, både mhp forventningsskjevhet (ME og MPE) og treffsikkerhet (MSE, RMSE, MAE og MAPE), var ventet. Denne modelltypen forklarer ikke trend og sesongsvingninger.

Selv om modellen basert på Holts metode kommer godt ut mhp forventningsretthet (ME og MPE), kommer denne modellen heller dårlig ut på treffsikkerhets-målene (MSE, RMSE, MAE og MAPE). Dette var da også som ventet, siden denne modellen ikke forklarer sesongsvingningene.

6 Avsluttende merknader

I innledningen satte vi opp tre hovedkriterier for valg av prognosemodell og trakk ut ett delkriterium som vi har sett på: Hvordan kan vi måle ulike prognosemodellers treffsikkerhet når de er utviklet ved hjelp av ulike metoder og vi ikke har noen erfaring fra hvor gode prognoser de ulike modellene genererer. Dette målte vi ved å simulere treffsikkerheten og så måle denne ved hjelp av kriterier for hvorvidt hver enkelt modell i gjennomsnitt treffer for lavt eller for høyt. Vi målte dette ved hjelp gjennomsnittlig

Tabell 10 Testparametere for de ulike modellene

	ME	MPE	MSE	RMSE	MAE	MAPE
Eksponteniell glatting	1380	6,33	3970055	1992	1862	9,04
Holts metode	68	-0,06	2195517	1482	1004	5,14
Holt-Winters metode	-139	-0,67	136829	370	292	1,41
Box-Jenkins metode	1	-0,01	47788	219	197	0,94
Kalmanfilter	186	0,89	93077	305	226	1,08
Regresjon	363	1,70	249704	500	402	1,90
Transfer-modell	406	1,82	861806	928	685	3,17

Samtlige testparametere er bedre enn for noen annen modell. Modellen har overraskende lave verdier for ME og MPE, dvs den er forventningsrett. I tillegg er den relativt treffsikker, idet den gir lave verdier for MSE, RMSE, MAE og MAPE.

På en god nummer to følger modellen basert på Kalmanfilter. For ME og MPE er den imidlertid slått både av modellen basert på Holts metode og av modellen basert på Holt-Winters metode: Prognosemodellen basert på Kalmanfilter lager for lave prognoser i forhold til den observerte utvikling. Målt ved MSE, RMSE, MAE og MAPE har likevel modellen basert på Kalmanfilter større treffsikkerhet enn modellen basert på Holt-Winters metode og langt større treffsikk-

erhet enn modellen basert på Holts metode.

At modellene basert på Box-Jenkins metode og Kalmanfilter kommer bedre ut enn modellen basert på regresjonsanalyse er ikke overraskende: Det skjer ofte (men slett ikke alltid) når korttids-prognosemodeller sammenliknes. Noe av forklaringen ligger i viktigheten av å inkludere alle relevante forklaringsvariable i en god regresjonsmodell, mens vi her bare har anvendt to. Dessuten er bare signifikante sesongvariable inkludert i vår regresjonsmodell, mens vår versjon av Kalmanfilter-modellen inkluderer alle sesongvariablene uansett om de er signifikante eller ikke. Dette bidrar til at modellen har større treffsikkerhet på sesongsvingningene. Test med alle sesongvariablene

avvik (ME) og gjennomsnittlig prosentvis avvik (MPE).

I tillegg beregnet vi det gjennomsnittlige avviket uavhengig av fortegn, dvs gjennomsnittlig absolutt feil (MAE) og den gjennomsnittlige prosentvis absolutte feil (MAPE). Et alternativ til MAE er gjennomsnittlig kvadrat-feil (MSE) og kvadratrotten av gjennomsnittlig kvadratfeil (RMSE).

Målt ved disse testparametrene var modellen basert på Box-Jenkins metode suverent best: Forventningsrett og svært god treffsikkerhet.

Som innledningen hentyder, er dette bare ett av de kravene som bør stilles til en god prognosemodell. Når de andre kravene testes kan det være at disse går på akkord med hverandre, idet modellen viser seg å ha høy karakter på noen, men ikke på alle kravene. Vi skal kort gi noen eksempler der vi holder vår test opp mot de andre kravene.

Hovedkriterium nr 1 krever at prognosemodellen skal inneholde et teoretisk og statistisk grunnlag. Det forekommer likevel fra tid til annen at prognosemodeller som tilfredsstillt anerkjente teorier for den utvikling som modellene skal beskrive og i tillegg de krav til statistisk signifikans som prognosemetoden er basert på genererer prognoser som treffer svært dårlig. Grunnen til dette er ofte uklart, men forklaringen søkes i momenter som selve teorien for tidsrekkens utvikling, for få data, datakvalitet, m m. I prognosesammenheng står en da overfor det klassiske dilemma om hvorvidt en skal godta de dårlige prognosene eller endre modellen slik at den genererer troverdige prognoser. I praksis vil vi på kort sikt trolig være tvunget til å godta en modell som ikke tilfredsstillt alle krav til teori og statistisk signifikans. Samtidig bør vi likevel på sikt forsøke å rette opp dette ved at alle sider ved prognosemodellen forbedres.

Hovedkriterium nr 2 har vi delvis behandlet ved at vi har simulert prognosemodellenes historiske treffsikkerhet. Vi har imidlertid ikke diskutert hvor godt modellene beskriver den historiske utvikling. Kravet om god beskrivelse av den historiske utvikling er ofte nært knyttet til kriterium nr 1 ovenfor, idet den riktige vei til god beskrivelse ligger i teori for den utvikling som modellene skal forklare og statistisk signifikans. Resonnementet blir derfor svært likt det som er gjort i avsnittet ovenfor. En

detaljert gjennomgang av modellens evne til å beskrive den historiske utvikling vil også kunne avdekke om treffsikkerheten på tilbakeholdte data skyldes tilfeldigheter som vi ikke kan regne med i fremtiden.

Det tryggeste er å skaffe seg erfaring med modellenes treffsikkerhet, ved at de får anledning til å generere prognoser som senere blir kontrollert. Selv om en modell er den beste målt ved våre testparametre, vil det over tid være fornuftig å ha 2–3 av de andre modellene til evaluering parallelt, fordi det kan være at en eller flere av disse da gjør det bedre.

Hovedkriterium nr 3 sier at prognosemodellene må være tilpasset brukerne av prognosene, dvs de må gi svar på det vi ber om prognoser for og være lett å anvende i forhold til tolkning, antall prognoser som skal lages, tid til disposisjon, brukerens kvalifikasjoner m m. Her skiller metodene seg klart ut. Glattingmodellene er svært enkle å håndtere siden modell-likningene er faste. Det samme gjelder vår versjon av modell basert på Kalmanfilter. Dette gjør at tidsaspektet taler til fordel for slike metoder, spesielt hvis mange ulike prognoser skal lages. I tillegg er modellene enkle å tolke.

Selv om det ligger et tungt fundament i form av teoretisk statistikk bak Kalmanfiltermodellene har suksessen med denne metoden likevel vært stor innen f eks trafikkprognoser, spesielt når automatisk deteksjon av nivåskift og outliere er en del av modellen.

Modeller basert på Box-Jenkins metode, regresjonsanalyse og transfer-modeller er imidlertid relativt kompliserte, idet de krever god innsikt i metodene og erfaring fra modellering. På den annen side gir gode modeller basert på regresjon og transfer-modeller større innsikt i hva som påvirker utviklingen i prognosevariabelen i form av forklaringsvariable. Slike modeller gir derfor raskere signaler om endringer i utviklingen.

Et moment som vi ikke har nevnt er prognosehorisonten. I vår test valgte vi et år som horisont. Hvis horisonten er flere år inn i fremtiden er det generelt antatt at modeller med forklaringsvariable vil gi bedre prognoser enn envariabelmodellene. Hvis det skal lages prognoser for kortere periode enn ett år bør en se på hvilke modeller som er mest treffsikre på de ulike avstander fra siste datapunkt. Dette kan gjøres ved å benytte våre test-

parametere på prognoser laget for hhv 1, 2, 3, ..., 12 perioder frem.

Prognoser for abonnementssetterspørsel

AV CARLO HJELKREM OG JOHANNES BØE

001.18:621.39

1 Behandling av grunnlagsdata, data på etterspørselsform

Arbeidet med forslag til etterspørselsprognosene begynner med et betydelig forarbeid i form av behandling av grunnlagsdata på sentralområdenivå. Vi skal her anta at dette arbeidet er utført, dvs at vi har data på etterspørselsform, og at eventuelle korrigeringer for nummertilnæringer og justeringer for grensejusteringer er foretatt. Se [1], hvor dette er diskutert.

Når vi nå har data på (riktig) etterspørselsform kan vi lage prognoser, enten ved at tidsrekken først analyseres med tanke på å bygge en prognosemodell eller ved at en allerede innarbeidet prognosemodell benyttes. Vi skal her presentere enkelte resultater fra dette arbeidet i Region Oslo.

Siden prognosearbeidet stadig er under utvikling, blir beskrivelsene nedenfor eksempler på hvordan prognosene kan lages med data fra Region Oslo: Noen av framgangsmåtene er allerede fraveket idet du leser dette, fordi nye metoder er utprøvd og ny informasjon er tatt i bruk. Likevel vil dette være eksempler på anvendelse av noen av de prognosemetodene som er presentert andre steder i denne utgaven av Telelektronikk.

2 Regionsprognosen

Forslaget til regionsprognose utarbeides alltid først i Region Oslo. Dette er et viktig steg fordi alle de (geografisk) underliggende prognoser knyttes opp mot denne: Sum sentralprognoser = sum kommuneprognoser = regionsprognosen.

Kommunene i regionen er et hjelpenivå som det ikke offentliggjøres prognoser for.

Prognosen for regionen gir oss retningslinjen for hvordan vi tror regionen totalt vil utvikle seg og blir dermed et "lokk" eller en beskrænkning for de underliggende prognoser. Ved utarbeiding av prognoser på lavere geografisk nivå kan det likevel være aktuelt å gå tilbake og endre regionsprognosen, hvis resultatene fra de underliggende nivåene skulle tilsi dette.

Grunnen til at vi begynner på toppen og så senere bearbeider de geografisk underliggende prognosene, finner vi i de erfaringer som er gjort gjennom prognosearbeidet i Region Oslo: Det er lettere å lage gode prognoser på regionsnivå enn på lavere geografiske nivå. Som regel vil utviklingen på regionsnivå være langt mer stabil enn på f.eks. sentralområdenivå. Begynner vi prognosearbeidet på sentralområdenivå, summerer vi opp den usikkerheten som ligger i hver av disse prognosene når regionsprognosene skal beregnes.

Utgangspunktet for regionsprognosene ligger i kjøring av to metningsmodeller; en for boligabonnement og en for forretningsabonnement. Disse gir prognoser for årlig total etterspørsel. I tillegg skal vi her se på resultatene av en enkel regresjonsmodell med tiden som forklaringsfaktor.

Det er også utviklet modeller for månedlig etterspørsel. Her benyttes for det meste ARIMA-modeller, mens vi her skal begrense oss til ett eksempel på Holt-Winters metode.

Som støtte til modell-framskrivingene benyttes informasjon som vi vet er relevant, men som modellene i mindre/ingen grad tar hensyn til. Slik informasjon skal vi også nevne, om enn i en noe kortere form enn resultatene fra modellkjøringene.

2.1 En enkel regresjonsmodell

En enkel regresjonsmodell av formen

$$y_t = a + b \cdot t + u_t \quad (1)$$

hvor

y = total etterspørsel etter abonnement

t = tiden

u = et stokastisk støyledd (med visse egenskaper, som vi ikke går inn på her)

vil kunne gi en antydning av den lineære trend i de historiske data. Tabell 1 viser total etterspørsel etter HA i perioden 1982 – 1990.

Etterspørselstallet for 1990 er justert noe i forhold til offisiell statistikk på grunn av ISDN. Denne justeringen går vi ikke inn på her.

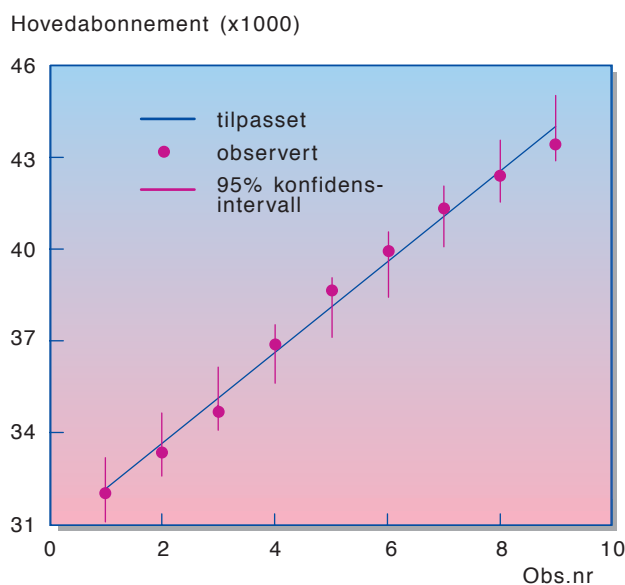
Tabell 2 viser resultatet av en kjøring av regresjonen i formel (1). Programpakken Statgraphics er benyttet. Tabellene og figurene nedenfor gjengir hovedpunktene i resultatene.

Vi ser at estimatene for både a (CONSTANT) og b (AARHA.aar) har høye t -verdier, dvs at variansen er så liten at det ikke er vanskelig å akseptere at de er signifikant forskjellig fra null. For øvrig er usikkerheten i estimatene gitt ved standardavvikene (std. error).

Den multiple korrelasjonskoeffisient (R-SQ) er over 99 %, hvilket indikerer en modell med meget god tilpasning til historiske data.

Går vi til Durbin-Watson indikatoren (DurbWat), er det litt vanskelig å tolke denne fordi vi bruker så få observasjoner: I den tabellen som vi bruker (den er ikke vist her) er 15 observasjoner det laveste som er angitt. Ved ekstrapolasjon av tabellen havner imidlertid verdien på 0,71 i grenseområdet mellom "ingen konklusjon" og "positiv autokorrelasjon". Det er altså mulig at støyleddene er autokorrelerte. Andre tester kan utføres for om mulig å få avklart dette, men det refereres ikke her.

Rent intuitivt kan det likevel sies at det er stor sannsynlighet for positiv autokorrelasjon siden vi benytter akkumulerte data: y vokser i alle perioder. I så fall bidrar positiv autokorrelasjon til overoptimistiske t -verdier (den estimerte variansen på parametrene underestimeres), og likeså overoptimistisk verdi på den multiple korrelasjonskoeffisienten. Alt dette gir prognoser med høyere usikker-



Figur 1 Regresjonslinje, observasjoner og 95 % konfidensintervall

Tabell 1 Total etterspørsel etter HA i Region Oslo, 1982–1990

Obs.nr	År	Total etterspørsel
(1)	1982	320 592
(2)	1983	334 201
(3)	1984	348 882
(4)	1985	368 011
(5)	1986	386 530
(6)	1987	399 706
(7)	1988	413 746
(8)	1989	423 934
(9)	1990	434 590

het enn om autokorrelasjonen ikke hadde vært til stede.

Videre gir tabell 2 også 95 % konfidensintervall for a og b . Vi ser at den årlige veksten antas å ligge mellom 13589 og 15965. Dette er viktig å merke seg når den årlige veksten skal anslås.

I figur 1 er resultatene vist grafisk: Regresjonslinjen, dvs modellen (Fitted) og observasjonene (Observed). I tillegg er også modellens 95 % konfidensintervall inntegnet.

Tabell 3 gir et sammendrag av residualene (støyleddene): Gjennomsnittet av residualene (Residual average) er svært nær null, og standardavviket (Residual standard error) er på 3891 (lavt i forhold til verdiene i data). Begge er tilfredsstillende.

I Figur 3 er residualene plottet rundt verdien null.

Vi ser her enda tydeligere det som figur 1 viser: Modellen overestimerer utviklingen de første 3 årene (dette gir negative støyledd). Deretter underestimeres utviklingen i de 4 neste årene (positive støyledd). Til slutt følger to år med negative støyledd, dvs en overestimering.

Det plott av typen figur 3 kan bidra med, er for det første å gi et visuelt bilde av eventuell avhengighet (autokorrelasjon) i støyleddene. Dette er i første omgang ikke lett å påvise i dette tilfellet. Durbin-Watson observatoren kan derfor være god å støtte seg til.

For det andre kan “formen” på støyleddutviklingen gi indikasjoner på hvilke andre forklaringsfaktorer som bør mod-

Tabell 2 Hovedpunkter fra regresjonsanalyse med programpakken STATGRAPHICS

Model fitting results for: AARHA.totettersp.					
Independent variable	Coefficient	Std. error	t-value	Lower Limit	Upper Limit
CONSTANT	-2.896579E7	9.977226E5	-29.0319	-3.13257E7	-2.66059
AARHA.aar	14776.9	502.377508	29.4139	13588.6	15965

R-SQ. (ADJ.)=0.9908 SE=3891.399448 MAE=3090.982716 DurbWat=0.71
 9 observations fitted, forecast(s) computed for 0 missing val. of dep.var.
 95 per cent confidence intervals for coefficient estimates.

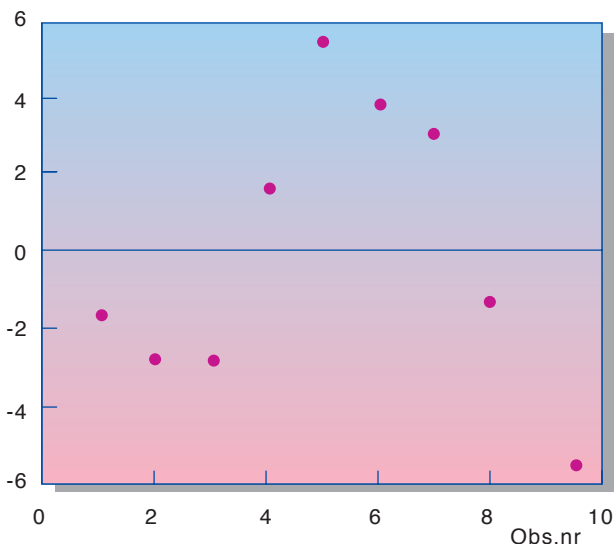
elleres inn. Eksempel: Vi har en mistanke (eller teori) om at aktivitetsnivået hos regionens forretningsabbonnenter er en viktig forklaringsfaktor, og vi har data som representerer denne utviklingen (som vi kan kalle X). En grafisk presentasjon av X-rekkens utvikling kan dermed holdes opp mot støyleddene i figur 3. Hvis de to tidsrekkene viser tilnærmet lik utvikling, kan det være en indikasjon på at X er en forklaringsfaktor som vil kunne bidra til ytterligere forklaring av abonnementsutviklingen og redusere størrelsen på støyleddene.

Figur 3 gir indikasjoner på at det ikke er galt å anta at støyleddene er Normalfordelt (den rette streken indikerer Normalfordelingen), selv om noen av verdiene avviker noe.

I figur 4 er observasjoner, modell og prognose plottet. For å gi et eksempel på hvilken effekt dette gir, er forklaringsfaktoren, dvs årstallene, endret fra 1982, 1983, 1984, ..., 1990 til 1, 2, 3, ..., 9. Dette gir altså samme resultat for den estimerte verdi av b (= 14776,9, se resultat av estimering i figurens overskrift) som i forrige kjøring, dvs hvor de rene årstall ble brukt som forklaringsfaktor. Imidlertid er den estimerte verdien for a endret fra -28965790 i den første kjøringen, til 307248 i denne kjøringen. Denne endringen skyldes den endringen i verdiene på forklaringsfaktoren som er gjort. Resultatene er, tross denne forskjellen, akkurat like gode: Forklaringsfaktoren vokser med en pr observasjon i begge tilfeller – vi har bare endret utgangsverdien (nivået).

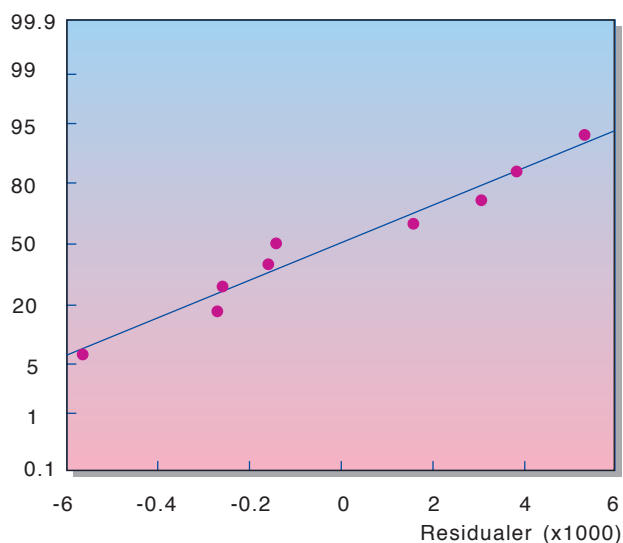
Figur 4 gir et inntrykk av bruken av modellen: Modellen følger “linjalen” (den er lineær), dvs den forsøker å

Residualer (x 1000)



Figur 2 Støyleddene i regresjonsmodellen

Kumulativ %

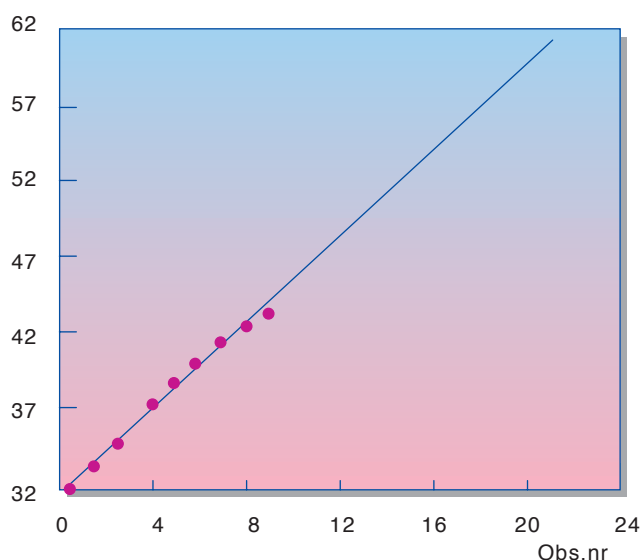


Figur 3 Normalfordelings-sannsynlighetsplot

Tabell 3 Analyse av støyleddene

Residual Summary
Number of observations = 9 (0 missing values excluded)
Residual average = 1.22883E-10
Residual variance = 1.5143E7
Residual standard error = 3891.4
Durbin-Watson statistic = 0.711909

Hovedabonnement (x10 000)



Figur 4 Observasjoner, modell og prognose

beskrive trenden i data og sender prognosene rett til himmels: Selv om dette er den beste kurvetilpasning (dvs modellering av historikken) vi kom fram til med tiden som eneste forklaringsfaktor, tror vi ikke at modellen lager gode prognoser for alle årene fram til år 2000. Den estimerte vekstparameteren, dvs b vil likevel kunne gi oss en indikasjon på utviklingen på kort sikt, dvs for 1991. For perioden 1992 til 2000 søker vi støtte fra andre modelltyper.

2.2 Vekstmodeller

Denne metoden brukes på to separate modeller for årlig utvikling i totalatterspørsel etter henholdsvis bolig- og forretningsabonnement i regionen. Tidsserien for abonnementsutviklingen fordelt på bolig og forretning er konstruert, da det for denne prioden ikke finnes offisielle data for denne fordelingen. Vi skal ikke i detalj gå inn på hvordan denne fordel-

ingen konstrueres, men henviser til [1] som angir ulike metoder for hvordan dette kan gjøres. Modellene er her brukt på årsdata for total etterspørsel etter Totalt antall tilknyttinger (TAT).

Modellen for boligabonnement benytter både tiden og befolkningmengden som forklaringsfaktorer¹, mens modellen for forretningsabonnement kun benytter tiden som forklaringsfaktor.

Modellen for boligabonnements- etterspørsel er gitt ved:

$$y_t = \frac{M}{(1 + e^{a+bt+c \cdot Bef})^g} \quad (2)$$

Her er:

y = total etterspørsel etter boligabonnement

t = tiden

Bef = befolkningen i regionen som er over 18 år (det må en abonnent være for å bestille abonnement)

M = metningsnivået

g = en konstant (som påvirker modellens veksttakt)

a , b og c er parametre som skal estimeres.

For M og g må det settes verdier før estimeringen kan skje.

En kan her tenke seg flere mulige forklaringsfaktorer, men data for flere aktuelle forklaringsfaktorer er ikke å oppdrive på dette geografiske nivå. Et unntak er altså befolkningsstatistikk, som finnes på kommunenivå.

Anslag for metningspunkt (M) i modell (2) var, da disse prognosene ble utarbeidet, basert på Statistisk Sentralbyrås (SSB) befolkningsprognoser, Folke- og boligtellingerne 1980 og Norges Byggeforskningsinstituttets "Boforhold i Oslo og Akershus", 1989. Verdien er satt til 400 000.

Verdien for g er anslått til 20 000. Denne verdien er hentet fra forskningsrapporten [3]. Rapporten undersøker bl a ulike verdier på g i samme modelltype som (2). Konklusjonen viser at selv om modellen brukes på data for så ulike tjenester som Datel, mobiltelefon og teleks, så vil den optimale verdien på g være ca 20 000. Dessuten vil ikke modellprognosene (under gitte betingelser som

vi her ikke går inn på) påvirkes nevneverdig om den "sanne" verdien er forskjellig fra 20 000.

Modellen for forretningsabonnements- etterspørsel er gitt ved:

$$y_t = \frac{M}{(1 + e^{a+bt})^g} \quad (3)$$

Her er y , t , M , a , b og g som for (2), bortsett fra at de nå gjelder for forretningsabonnements- etterspørselen.

Anslag for metningspunkt (M) i modell (3) er basert på antall foretak i regionen og et skjønnsmessig anslag for framtidig utvikling. Dette anslaget er betraktelig mer usikkert enn i tilfellet med metningspunkt for boligabonnements- etterspørselen; vi har lite konkret underlag for anslag på M i dette tilfellet. Verdien er satt til 250 000.

Anslag for g er gjort på samme måte som i tilfellet med boligabonnementsmodellen ovenfor; verdien er satt til 20 000, ref det som er sagt ovenfor.

Tabell 4 viser data for befolkningsutviklingen (18 år og over) og den beregnede boligabonnements- og forretningsabonnementsutvikling i perioden 1982–1990.

Estimeringsprogrammet som benyttes til estimering av a , b og c i (2) og a og b i (3), er internt utarbeidet i Region Oslo. Det er spesialkonstruert med tanke på de transformasjoner som må gjøres for parameterestimering i de logistiske metningsmodellene (2) og (3): Hvis vi skal benytte lineær regresjon, må modellen være lineær i de parametrene som skal estimeres. Det er ikke tilfellet for logistiske metningsmodeller.

Vi beskriver ikke hvordan selve transformasjonene foregår, men viser hvordan den lineær-transformerte funksjonen ser ut for (2):

¹ Det kan diskuteres om det er fornuftig å benytte både tiden (trendvariabel) og befolkningmengden samtidig i samme modell, fordi dette er variable som begge vokser i hele dataperioden: I slike tilfeller oppstår ofte problemer med kolinearitet (multikolinearitetsproblemer) mellom disse variablene og da vil variablene "konkurrere" om å forklare den samme veksten i abonnementsutviklingen. Resultatet blir dårlige estimater for koeffisientene. Se f.eks [2] som beskriver dette problemet.

$$\ln \left(\left(\frac{M}{y_t} \right)^{\frac{1}{g}} - 1 \right) = a + b \cdot t + c \cdot Bef \quad (4)$$

Som vi ser av (4) er nå høyresiden i denne likningen lineær i a , b og c . Vi har med andre ord oppnådd å transformere den ikke-lineære modellen (2) til en lineær modell, og da kan vi benytte lineær regresjon for å estimere a , b , og c .

Ser vi på venstre side av (4), er dette et relativt komplisert uttrykk. Vi kjenner imidlertid verdiene for y , M og g . M og g er konstanter og for hver verdi av y får vi en verdi for "venstre side". Dermed kan venstre side oppfattes som en egen variabel som vi kan kalle y^* . (4) kan da skrives som:

$$y^* = a + b \cdot t + c \cdot Bef \quad (5)$$

der

$$y^* = \ln \left(\left(\frac{M}{y_t} \right)^{\frac{1}{g}} - 1 \right)$$

Likning (5) blir altså regresjonslikningen (med tillegg av et støy-ledd) som benyttes til estimering av a , b og c .

Tilsvarende transformasjon gjøres også før a og b estimeres i modellen for forretningsabonnements-eterspørselen, (3). På tilsvarende måte får vi da:

$$y^* = a + b \cdot t \quad (6)$$

der y^* er det samme som for (5) ovenfor (men med de verdier for y , M og g som tilhører (3)). Dette blir altså regresjonslikningen som benyttes for estimering av a og b i (3) (med tillegg av et støy-ledd).

Det lar seg også gjøre å estimere direkte på (2) og (3). Da må vi imidlertid benytte ikke-lineære estimeringsmetoder, som er kompliserte og i praksis må spesialkonstruere til denne modellformen. I Region Oslo benytter vi oss ikke av slike metoder, siden vi i dette tilfellet kan "lineær-transformere" modellene.

Etter at parametrene a , b , og c i (5) og a og b i (6) er estimert, settes disse verdiene inn i (2) og (3). Dermed kan vi sjekke hvor godt modellen beskriver den historiske utvikling. Med framtidige anslag for t (den er sikker) og Bef (dette er en prognose med en viss usikkerhet) kan vi så lage modellprognoser for de to abonnements-eterspørselene.

Det programmet vi benytter, gir en utskrift som er vist i tabell 5.

Tabell 5 viser, foruten historiske data og prognoser for abonnementsutviklingen fordelt på bolig- og forretningsabonnement, også noen modelltekniske indikatorer:

- Parameterestimat for ALFA (konstanten a i nevnerens eksponent, se modellene (2) og (3))
- BETA1 (angir modellenes vektlegging av tiden som forklaringsfaktor, ref parameter b i nevnerens eksponent)
- BETA2 (angir boligabonnementsmodellens vektlegging av befolkningsutviklingen som forklaringsfaktor, ref konstanten c i (2))
- Den multiple korrelasjonskoeffisient (R^2) og
- Standardavviket rundt regresjonslinjen (3) og regresjonsplanet (2), (S).

Som vi ser, viser begge modellene meget god tilpasning til historiske data, med R^2 godt over 0,99. Dessuten viser standardavvikene, S , meget tilfredsstillende verdier i forhold til verdiene i data. Statistiske indikatorer utover dette må vi søke i andre statistikkpakker. Vi går imidlertid ikke dypere inn på dette her.

Modellen for boligabonnements-eterspørsel ser nå slik ut, når vi setter inn verdiene for konstantene M og g og de estimerte parametre:

$$y_t = \frac{400000}{(1 + e^{13,94508 - 0,00002 \cdot Bef})^{20000}} \quad (7)$$

Modellen for forretningsabonnements-eterspørsel ser slik ut, når vi setter inn verdiene for konstantene M og g og de estimerte parametre:

$$y_t = \frac{250000}{(1 + e^{130,445 - 0,07075t})^{20000}} \quad (8)$$

Ser vi på verdiene for de estimerte parametrene blir vi fort slått av de lave og negative verdiene for BETA-ene. Spesielt er boligabonnementsmodellens BETA2 svært lav (-0,00002).

Fortegnet på BETA-ene er i begge modellene "riktig" bestemt. At dette er tilfellet kan vi se direkte av f eks (2): Med positive verdier på b og c , vil eksponenten i nevneren stadig vokse når t og Bef vokser. Dermed vokser også hele nevneren med økende verdier på t og Bef . At nevneren i en brøk stadig vokser, betyr at hele brøken, når telleren er en konstant, går mot null! Er imidlertid b og c negative, betyr det at eksponenten i

Tabell 4 Befolkningsmengden (18+) og beregnet fordeling av TAT på bolig og forretning. Region Oslo

Region Oslo			
År	Befolkning	Boligabonn.	Forr. abonn.
1982	513820	245517	82324
1983	516714	255298	86692
1984	521602	266186	90744
1985	527573	278119	98744
1986	533344	289505	106494
1987	539041	300222	112850
1988	543481	306219	119421
1989	546049	311314	125079
1990	550376	317111	131347

nevneren avtar med voksende verdier på t og Bef . Da går e opphøyd i eksponenten mot null og dermed hele nevneren mot 1. Den totale effekten på (2) blir m a o at funksjonen går mot M (og det er jo meningen!).

At verdiene er så små (i absoluttverdi) er også "riktig" fordi nivået på verdiene til forklaringsfaktorene er høye: Den estimerte parametre skal multipliseres med den tilhørende forklaringsfaktor når modelltilpasning og prognoser skal regnes ut.

I tillegg gir utskriften en oversikt over regionens historiske og prognostiserte utvikling for total etterspørsel etter TAT fordelt på bolig og forretning, SSBs historiske og prognostiserte befolkningsutvikling, boligabonnementsstettheten (Bol/Bef) og totaltilknytningstettheten (Tot/Bef).

2.3 Holt-Winters metode

Som nevnt er det også utviklet modeller for den månedlige etterspørsel etter abonnement. Generelt er dette viktig fordi vi da utnytter informasjon utover det som årsdata gir oss. Spesielt viktig er dette ved revisjonen av prognosene etter halvgått år. Da bidrar ikke årsdata med ytterligere informasjon: Det er i månedsutviklingen vi må søke hjelp.

Vi har her valgt å vise et eksempel med Holt-Winters additive modell². På generell form kan denne prognosemodellen skrives som:

Tabell 5 Utskrift av prognoseprogram for Region Oslo

28- 1-1991 12:00 Prognose for Region Oslo										
År	Befolkning		Boligabonn.		Forr.abonn.		Tot.ant.tilkn.		Bol/Bef	Tot/Bef
	Total	Vekst	Total	Vekst	Total	Vekst	Total	Vekst		
1982	513820		245517		82324		327841		0,478	0,638
1983	516714	2894	255298	9781	86692	4368	341990	14149	0,494	0,662
1984	521602	4888	266186	10888	90744	4052	356930	14940	0,51	0,684
1985	527573	5971	278119	11933	98744	8000	376863	19933	0,527	0,714
1986	533344	5771	289505	11386	106494	7750	395999	19136	0,543	0,742
1987	539041	5697	300222	10717	112850	6356	413072	17073	0,557	0,766
1988	543481	4440	306219	5997	119421	6571	425640	12568	0,563	0,783
1989	546049	2568	311314	5095	125079	5658	436393	10753	0,57	0,799
1990	550376	4327	317111	5797	131347	6268	448458	12065	0,576	0,815
1991	554207	3831	322668	5557	137255	5908	459924	11466	0,582	0,83
1992	558012	3805	327862	5194	143000	5744	470862	10938	0,588	0,844
1993	561145	3133	331985	4122	148568	5568	480553	9691	0,592	0,856
1994	563895	2750	335481	3496	153951	5383	489432	8879	0,595	0,868
1995	566260	2365	338415	2935	159143	5191	497558	8126	0,598	0,879
1996	568718	2458	341319	2904	164137	4994	505456	7898	0,6	0,889
1997	571158	2440	344079	2760	168931	4794	513010	7554	0,602	0,898
1998	573276	2118	346426	2347	173524	4593	519950	6940	0,604	0,907
1999	575477	2201	348753	2327	177916	4392	526669	6719	0,606	0,915
2000	577614	2137	350931	2178	182108	4192	533039	6370	0,608	0,923
BOLIG:					FORRETNING:					
POTENSIAL:		400000			POTENSIAL:		250000			
R^2:		0,99889			R^2:		0,99528			
ALFA:		13,94508			ALFA:		130,44500			
BETA1:		-0,00762			BETA:		-0,07075			
BETA2:		-0,00002								
S:		975,39270			S:		1306,04059			

$$y_{t+k} = \mu_t + \beta_t k + S_{t+k} - Im \quad (9)$$

der

$$\mu_t = a(\mu_{t-1} + \beta_{t-1}) + (1-a)(y_t - S_{t-m}) \quad (10)$$

$$\beta_t = b\beta_{t-1} + (1-b)(\mu_t - \mu_{t-1}) \quad (11)$$

$$S_t = cS_{t-m} + (1-c)(y_t - \mu_t) \quad (12)$$

hvor

t = tidspunkt

y = abonnementssetterspørselen

μ = tidsrekkens nivå

β = tidsrekkens stigning

S = tidsrekkens sesongutslag

k = antall perioder det lages prognoser for

m = antall sesonger i data (f eks 12 for månedsdata eller 4 for kvartalsdata)

I = det minste positive heltall slik at $k - Im = 0$

a, b og c konstanter (glattingsparametre).

² Se artikkel om glattingsmodeller i denne utgaven av *Elektronikk*.

Tabell 6 Vekter for beregning av antall virkedager

Mandag – fredag:	1
Lørdag – søndag:	0
Off. høytidsdager:	0
Jule-/Nyttårsaften:	0,5
Inneklemte virkedager:	0,5

En av de viktige momenter ved en modell som baserer seg på data for kortere perioder enn år, er sesongsvingningene og estimering av disse. Ettersom det er ulike antall virkedager i hver måned og disse spiller en viktig rolle som forklaring på antall etterspurte abonnemang i hver måned, har vi valgt å bruke etterspørsel pr virkedag som avhengig variabel. I månedsmodellene har vi valgt å lage prognoser for HA- etterspørselen. Dermed blir variabelen som skal prognostiseres (månedlig etterspørsel etter) HA/virkedag. Antall virkedager er her beregnet etter vektleggingen i tabell 6.

Eksempler på “inneklemte virkedager” er fredag etter Kristi Himmelfartsdag, mandag–onsdag før Skjærtorsdag, virkedager i romjulen, osv.

Tabell 7 viser data for den månedlige etterspørsel etter HA pr virkedag fra og med januar 1984 (observasjon nr 1) til og med desember 1990 (observasjon nr 84).

Et vanskelig punkt ved alle typer glattingsmodeller er hvilke verdier vi skal velge for glattingsparametrene. Her benytter den ferske prognosemaker ofte prøve- og feilemetoden, mens den erfarne har en lettere jobb. En måte å unngå dette problemet på, er å bruke en optimeringsrutine til å estimere de optimale verdier for disse parametrene. De fleste tilgjengelige statistikk-/prognosepakker har ikke slike funksjoner, men den programpakken vi benytter til å kjøre Holt-Winter modeller har denne optimeringsfunksjonen. Den heter FORECAST, og er en avansert prognosepakke som er utviklet av Televerket og Norsk Regnesentral.

De estimerte (optimale) verdier av glattingsparametrene ble av FORECAST beregnet til: Nivå (a) = 0,76, stigning (b) = 0,98 og sesong (c) = 1,0. Setter vi disse verdiene inn i (9), aksepterer at siste observasjon er nr 84 (desember 1990), at vi har 12 sesonger og skal lage en prognose for mars måned (3 perioder fram), blir prognosemodellen:

$$y_{87} = \mu_{84} + \beta_{84}^3 + S_{75} \quad (13)$$

der

$$\mu_{84} = 0,76(y_{84} - S_{76}) + 0,34(\mu_{83} + \beta_{83}) \quad (14)$$

$$\beta_{84} = 0,98(\mu_{84} - \mu_{83}) + 0,02\beta_{83} \quad (15)$$

$$S_{75} = 1,00S_{63} + 0,0(y_{75} - \mu_{75}) \quad (16)$$

y -ene er her gitt i tabell 6 ovenfor, hvor fotskriftene referer seg til observasjonsnummeret. Verdiene for S_{63} , S_{76} , μ_{75} , y_{75} , μ_{83} , μ_{84} og β_{83} avhenger av tidligere verdier på disse, som oppdateres rekursivt. Disse refereres ikke her.

Slik som Holt-Winters modell er formulert ved (9) – (12) (og i FORECAST), uttrykker glattingsparameteren hvor stor vekt som skal legges på de gamle estimatene for nivå, stigning og sesong. I minus glattingsparameteren, er dermed et uttrykk for hvor mye vekt som skal tillegges ny informasjon om nivå, stigning og sesong.

Modellen legger altså all vekt på historikkens sesongsvingninger og ingen på de nye som måtte komme inn. Modellen legger også meget stor vekt på den historiske stigningen og lite på ny informasjon om denne. Grunnen til dette er muligens å finne i selve det at vi lar en optimerings-prosedyre bestemme hvilke glattingsparametre som gir den beste til-

pasning til data: Vi legger vekt på å beskrive den historiske utvikling og å tilpasse modellens glattingsparametre til denne. Hvis det er stor usikkerhet i data i form av kraftige og tilfeldige vekst-/sesongsvingninger kan det være at usikkerheten blir stor. I så fall vil optimeringsrutinen medføre at historikken tillegges stor vekt, mens nye (og usikre) data tillegges minimal vekt. En annen mulighet er at optimeringen ikke har funnet de (globale) optimale verdier for parametrene, men har stoppet ved et “lokalt optimum”.

Om dette virkelig gir den beste prognosemodell og dermed de beste prognosene skal

Tabell 7 Månedlig etterspørsel etter HA pr virkedag fra og med januar 1984 til og med desember 1990

(1)	66,6363	(19)	83,4347	(37)	92,2381	(55)	-9,3809	(73)	70,045
(2)	71,6666	(20)	87,5454	(38)	65,8	(56)	25,2608	(74)	58,35
(3)	66,0909	(21)	92,1904	(39)	57,2727	(57)	13,0454	(75)	54,5909
(4)	60,7272	(22)	52,5217	(40)	81,3142	(58)	19,9047	(76)	-13,5757
(5)	43,6923	(23)	65,8571	(41)	52,2162	(59)	34,6363	(77)	44
(6)	49,8461	(24)	55,6111	(42)	71,6666	(60)	25,421	(78)	10,9
(7)	35,4545	(25)	86,909	(43)	52,7826	(61)	47,5909	(79)	2,8636
(8)	58,7391	(26)	89,15	(44)	81,4761	(62)	51,05	(80)	40
(9)	59,9	(27)	69,6363	(45)	78,6818	(63)	16,7567	(81)	66,45
(10)	73,1304	(28)	86,2727	(46)	54,2727	(64)	42,35	(82)	34,7826
(11)	74,2727	(29)	62,3529	(47)	56,3809	(65)	21,1666	(83)	96,6363
(12)	46,1764	(30)	67	(48)	51,8974	(66)	10,2272	(84)	49,7714
(13)	102,455	(31)	26,7391	(49)	83,25	(67)	-10,619		
(14)	83,75	(32)	85,2381	(50)	103,809	(68)	64,7826		
(15)	59,9523	(33)	96,2272	(51)	99,317	(69)	54		
(16)	78,6285	(34)	71,6521	(52)	59,5789	(70)	50,1818		
(17)	72,8947	(35)	84,25	(53)	53,2777	(71)	75,6363		
(18)	90,35	(36)	76,421	(54)	76,6818	(72)	64,4864		

vi ikke diskutere her. Poenget er å vise hvordan en slik optimeringsrutine virker på et datasett.

Siden vår prognosevariabel er HA/virke- dag, blir prognosene også uttrykt i HA/virkedag. Kalenderen er “determin- istisk”, og bringer derfor ikke noen ytter- ligere usikkerhet inn i prognosene når vi multipliserer opp med antall virkedager i

hver måned for å få månedsprognosene for HA. I tabell 8 er denne utregningen foretatt.

Av andre analyser og vurderinger som gjøres når prognosene skal utarbeides, nevner vi kort noen nedenfor.

2.4 Andre metoder for kort- siktige prognoser

Ovenfor viste vi eksempel på Holt-Wint- ers metode. Det finnes imidlertid flere andre metoder som er utviklet med hen- blikk på kortsiktige prognoser. Av de mest anerkjente og brukte kan nevnes Box-Jenkins metode. En annen metode som er kommet sterkt de siste årene er Kalmanfilter-modeller, se [7]. Andre metoder finnes også, men vi går ikke inn på disse her.

2.5 Vurdering av den under- liggende trend – glidende gjennomsnitt

At det kan være vanskelig å se/identi- fisere hvordan utviklingen virkelig har vært ved bare å betrakte et plott av data, kan figur 5 være et eksempel på. Der er månedsdata for etterspørsel etter HA i Region Oslo for perioden januar 1984 til desember 1990 plottet.

For å eliminere støy og sesongsvingninger i data har vi i tillegg plottet et 12-måned- ers glidende gjennomsnitt. Når støy og se- songsvingninger er eliminert, står vi til- bake med trenden i data.

Det glidende 12 måneders gjenn- omsnittet er summen av 12 observasjoner, delt på 12. Når en ny observasjon bringes inn, “kastes” den eldste ut, slik at vi hele tiden operer med snittet av 12 observasjoner.

Vi ser av figuren at trenden over den siste perioden har vært svakt negativ.

Bortsett fra dette gir figuren et godt overblikk over utviklingen i dataperioden: Først en relativt lang periode med svak negativ trend, så en periode med sterk negativ trend (som skyldes oppsigelsesopprydding i Region Oslo), som så snus til en periode med sterk positiv trend (oppsigelsesoppryddingen er over, og vi returnerer til et mer “norm- alt” nivå). Til slutt får vi en peri- ode som likner mer på den første, men trenden er nå mindre negativ enn i den første.

“Motstykket” til 12-måneders glidende gjennomsnitt er årsdata: Vi forsøker jo å se hvordan utviklingen (kontinuerlig) har vært over 12 måneder. Årstrenden estimerte vi i regresjonsmodellen i kapittel 2.1.

Innen et år kan trenden godt ha skiftet. Et eksempel på dette er 1989: Den negative trend i 1. halvår ble snudd til positiv for 2. halvår. Det er da naturlig å spørre om det er normalt med en svak utvikling i første halvår og en sterk utvikling i andre. På det tidspunkt prognosene lages, er det derfor viktig å merke seg hva slags utvikling en er inne i og hvor lenge denne har vart. Dette er til støtte når en gjør antakelser om de kommende perioder og vil dessuten forklare noe av månedsmodellenes prog- noseresultat mht hvor godt har modellen fanget opp trendskiftet.

2.6 Periodisert månedsandel

Både på det tidspunkt som prognosene utarbeides og i den månedlige oppfølging gir “etterspørsel hittil i år, som peri- odisert prosentvis andel av den totale årsetterspørsel” en pekepinn på hva års- prognosen kan bli. Eksempel: Vi har data for etterspørselen de to første månedene i året. Tidligere periodisering av januar og februar har vist at disse månedene i gjennomsnitt utgjør henholdsvis 11,3 % og 10,4 % av den totale etterspørselen over året. Til sammen utgjør disse månedene 21,7 % av årsetterspørselen. Det er da ikke vanskelig å finne hvor mye 100 % (etterspørselen for hele året) vil være. Etter hvert som månedene går, sier dette også hvor stor den resterende månedlige etterspørsel må være for at vi skal nå prognosen.

2.7 Vurdering av samfunns- økonomisk utvikling

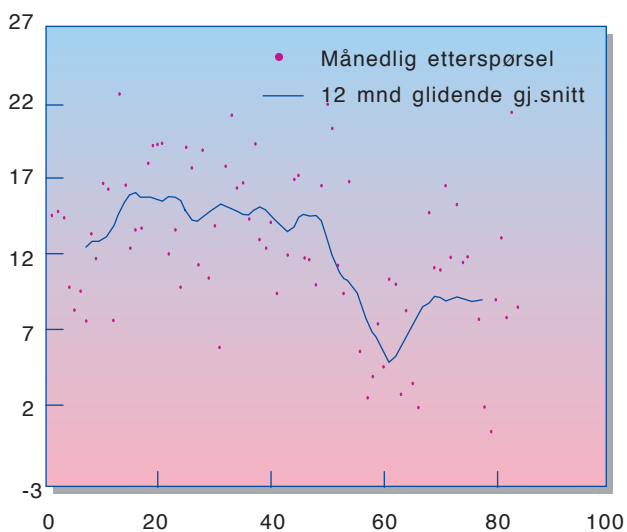
Konjunkturanalyser og prognoser for viktige samfunnsøkonomiske nøkkelvari- abler som f eks privat konsum, real- investeringer, boligbygging m fl følges opp, og modellframskrivningene holdes opp mot disse. De viktigste kildene er: Nasjonalbudsjettet (St. meld. nr 1), Lang- tidsprogrammet (St. meld. nr 4), Penger og Kreditt (Norges Bank) og Økonom- iske Analyser (SSB). I tillegg kommer prognoser/ konjunkturanalyser fra de største bankene, Byggearealstatistikk, Foretaks-/konkursstatistikk, m m.

Et problem med de aller fleste slike dokumenter er at de analyserer og gir prognoser for hvordan landet totalt

Tabell 8 Prognose for etterspørsel etter HA, 1991

Prognose for 1991			
Måned	HA/ virkedag	Virke- dager	HA
Januar	70,047	22	1541
Februar	66,37	20	1327
Mars	52,652	17, 5	921
April	48,35	21	1009
Mai	41,52	18,5	768
Juni	51,369	20	1027
Juli	17,504	23	403
August	70,267	22	1546
September	54,675	21	1148
Oktober	46,316	23	1065
November	61,306	21	1287
Desember	44,518	17,5	779
Sum 1991	624,579	246,5	12822

Hovedab0nnement (x 100)



Figur 5 Månedlig netto etterspørsel etter HA og 12 måneders glidende gjennomsnitt for Region Oslo

utvikler seg, mens det enkelte telegion godt kan oppleve en annen utvikling.

2.8 Vurdering av annen foreliggende intern/ekstern informasjon

Alle modellkjøringene og all annen tilleggsinformasjon som kan bidra til å korrigerer disse, må selvfølgelig holdes opp mot hverandre før de endelige prognosene settes: Ulike modelltyper og ulik tilleggsinformasjon kan ofte gi motstridende resultater.

Slik tilleggsinformasjon kan være alt fra planer til uttalelser om teknikkutvikling og markedsforhold fra egen avdeling, markedsanalyser, internasjonale trender, osv.

Det er ovenfor gitt spredte kommentarer om de ulike modellens styrke og svakheter. En detaljert diskusjon som viser hvordan Region Oslo bedømte de ulike modellresultatene og tilleggsinformasjon opp mot hverandre, vil føre for langt her. Likevel er det en "symbiose" av alle slike resultater og informasjon som til slutt gir de endelige prognosene.

3 Kommuneprognoser

Prognosene som utarbeides for kommunene i Region Oslo er ikke offisielle og distribueres derfor ikke. Kommunene utgjør likevel et nyttig hjelpenivå i arbeidet med å knytte de ulike geografiske aggregeringsnivå sammen. Dessuten er viktig offentlig statistikk, bl a befolkningsstatistikk, fordelt på kommuner.

Vi skal her se på et eksempel hvor det benyttes to logistiske modeller for hver kommune i regionen: En modell for boligabonnemement og en modell for forretningsabonnemement. Disse tilsvarer modellene for regionen totalt, se modellene (2) og (3).

Hver kommunemodell har et metningsnivå som er satt på tilsvarende måte som for regionsmodellene.

Konstanten g er satt til 20 000 i alle modeller, se det som er sagt om dette i avsnittet om vekstmodellene i kapittel 2.2.

På lik linje med modellene for regionen bruker vi også her, i tillegg til forklaringsfaktoren tid, SSBs tall for befolkningsstatistikk/framskrivning som forklaringsfaktor i modellene for boligabonnemements-etterspørsel. Selv om befolkningsframskrivningene til tider har

vist seg å være i utakt med den faktiske utvikling, er det antatt at forklaringsvariabelen bidrar til å forklare de relative forskjeller mellom kommunene.

For ordens skyld repeterer vi modellen for boligabonnemements-etterspørsel:

$$y_t = \frac{M}{(1 + e^{a+bt+cBef})^g} \quad (2)$$

Her er y total etterspørsel etter boligabonnemement i den gjeldende kommune, t tiden og Bef befolkningen som er over 18 år i den gjeldende kommune. a , b og c er parametrene som skal estimeres, mens M er metningsnivået og g er en konstant (som påvirker modellens veksttakt). For de to sistnevnte må det settes verdier før estimeringen kan skje.

Tilsvarende har modellene for forretningsabonnemements-etterspørselen kun tiden som forklaringsfaktor:

$$y_t = \frac{M}{(1 + e^{a+bt})^g} \quad (3)$$

Her er y , t , M , a , b og g som for (2), bortsett fra at de nå gjelder for forretningsabonnemements-etterspørselen i den gjeldende kommune.

Modellene blir også i dette tilfellet brukt på årsdata for total etterspørsel etter TAT³. Splittingen i markedssegmentene bolig og forretning gjøres på tilsvarende måte som for regionsdata, se ovenfor.

Vi har valgt å vise data, modell og prognoser fra en av kommunene i Region Oslo. I tabell 9 er det presentert data for befolkningsutviklingen (18 år og over), boligabonnemement (konstruert) og forretningsabonnemement (konstruert) for Oslo kommune.

Selve transformasjonen av data, før minste kvadraters metode kan benyttes til estimering av parametrene a , b og c i boligabonnemementsmodellene og a og b i forretningsabonnemementsmodellene, utføres på samme måte som tilfellet var for de tilsvarende regionsmodellene.

³ Hvis det er HA en ønsker prognoser for, kan selvfølgelig tilsvarende modeller bygges for denne variabelen. Alternativt kan forholdet mellom HA og TAT trekkes ut i de prognosene som er presentert her.

Utskriftene fra prognoseprogrammet er helt tilsvarende som for regionen totalt, se tabell 5. Vi viser ikke denne her, men setter inn verdiene for M og g , samt de estimerte verdiene for a , b , og c .

Modell for boligabonnemements-etterspørsel for Oslo kommune:

$$y_t = \frac{25900}{(1 + e^{159,8106 - 0,07862t - 0,00004Bef})^{20000}} \quad (17)$$

Modellen for forretningsabonnemements-etterspørsel for Oslo kommune:

$$y_t = \frac{259000}{(1 + e^{118,67177 - 0,06483t})^{20000}} \quad (18)$$

Innledningsvis nevnte vi at alle (geografisk) underliggende prognoser knyttes opp mot regionsprognosen: Sum sentralprognoser = sum kommuneprognoser = regionsprognosen. Dette betyr at det som regel ikke er de direkte resultatene fra kommune-modellene som vurderes: Summen av de individuelle kommuneprognosene er sjeldent lik distriktsprognosen⁴. Derfor må det skje en avstemming av kommuneprognosene slik at summen blir konsistent med distriktsprognosen.

Det tidligere omtalte prognoseprogram for de logistiske modellene utfører denne avstemmingen automatisk. Prinsippet for avstemmingen kan forklares ved følgende eksempel: Anta at sum av alle kommuneprognoser innen regionen i første omgang gir høyere tall enn regions-

Tabell 9 Befolkningsmengden (18+) og beregnet fordeling av TAT på bolig og forretning. Oslo kommune

Oslo kommune			
År	Befolkning	Boligab.	Forr. ab.
1982	369218	185596	66761
1983	369326	191922	70021
1984	370338	199266	73160
1985	372404	206829	79336
1986	374332	214354	85248
1987	376480	221374	88959
1988	378047	224981	93267
1989	378702	227783	97363
1990	380187	231116	101707

prognosen. Da “beskjæres” de enkelte kommuneprognosene etter hvilke kommuner som har den høyeste årlige etterspørsel: Den kommunen som har den høyeste årlige etterspørselen blir justert mest ned, den kommunen som har nest høyest etterspørsel blir justert nest mest ned, osv. Det omvendte skjer om sum av alle kommuneprognosene skulle bli lavere enn regionsprognosen: Da “løftes” de enkelte kommuneprognosene også etter hvilken kommune som har den høyeste årlige etterspørsel.

Det foretas en proratisk fordeling av differansen mellom regionsprognosen og summen av de underliggende kommuners prognoser. Dette slik at etterspørselen i ett gitt år, for en vilkårlig kommune, framkommer ved:

$$KP = KP_i^* + \frac{\Delta KP_i^*}{\sum_i \Delta KP_i^*} \cdot \left(RP - \sum_i KP_i^* \right) \quad (19)$$

hvor

KP_i = Konsistent kommuneprognose, kommune i

KP_i^* = Modellprognose, kommune i

RP = Regionsprognose

Δ = Årlig vekst

Σ = Sum

En slik enkel regel er selvfølgelig ikke “riktig” i alle tilfeller, men har f eks vist seg å fungere bedre enn å rangere justeringen etter total etterspørsel: Da vil store kommuner med liten årlig etterspørsel risikere å bli justert ned til negativ netto etterspørsel. Se for øvrig [5] som forklarer nedbryting fra teledistrikt til teleområder i Televerkets gamle organisasjon.

En annen mulig konsistens-regel kunne være å la den kommunen som har modell med størst usikkerhet i prognosene bære den største andelen av differansen mellom sum kommuneprognoser og regionsprognosen, osv. Dette er en mye brukt metode når det skal lages prognoser for store trafikkmatiser, se f eks [4].

I tillegg til bruk av vekstmodeller sammenholdes statistikker og annen relevant informasjon på kommunenivå for eventuelle korreksjoner av modellframskrivningene. Dette går vi ikke inn på her.

4 Sentralprognosene

Når sentralprognosene skal utarbeides, tar vi i utgangspunktet fatt i dette (laveste) geografiske nivå “uavhengig” av de ovenfor nevnte prognoser. Når så prognosene er utarbeidet for hvert område individuelt, konsistentgjøres disse med kommuneprognosene, dvs sum sentralprognoser innen en kommune = kommuneprognosen. Dette foregår på samme måte som når kommuneprognosene konsistentgjøres med regionsprognosen:

I det tilfellet at sum av sentralområdeprognosene innen en kommune er høyere enn kommuneprognosen (dvs den konsistente kommuneprognosen, ref det som er sagt om dette i kapittel 3) “beskjæres” sentralområdeprognosene etter hvilket sentralområde som har den høyeste årlige vekst, dvs det sentralområdet som har den høyeste (årlige) etterspørselen blir justert mest ned, det sentralområde som har nest høyest (årlig) etterspørsel blir justert nest mest ned, osv.

I vårt eksempel skal vi se på utarbeiding av sentralprognoser på årsdata for TAT.

Månedssdata, eventuelt kvartalsdata, kan i prinsippet også benyttes som grunnlag for prognoser på dette nivå. Imidlertid har disse vært ansett som usikre, og månedlig nummerlansstatistikk (se [1]) har ikke vært tilgjengelig før i 1991. Dessuten blir det svært mye data å holde styr på.

I prinsippet kan en tenke seg en tilsvarende oppsplitting på bolig- og forretningsabonnement som på regions- og kommunenivå. Slike data har imidlertid ikke vært tilgjengelig før inntil for kort tid siden. I Region Oslo var det heller ikke foretatt noen “teoretisk” oppsplitting på brukskategori på sentralområdenivå da disse prognosene ble laget. Dette er imidlertid en mulig vei til forbedret underlag for prognosene.

Før framskrivningene finner sted er selvfølgelig de historiske årsdata bearbeidet, slik at hver sentrals historikk består av “rene” sentralgrenser og er korrigeret for grensejusteringer. Dette er omtalt i [1].

Selve prognosene lages i dag på bakgrunn av Kalmanfiltermodeller, se [7]. Det vil her føre for langt å referere fra disse modellene. Vi skal imidlertid se på hvordan de på en enklere måte ble utarbeidet før slike modeller ble tatt i bruk.

Da ble prognosene laget delvis på bakgrunn av gjennomsnittsbetraktninger over historisk vekst og delvis på bakgrunn av

lokal kjennskap og antakelser om utviklingen i hvert enkelt sentralområde. Markedskunnskap på sentralområdenivå er her spesielt viktig i prosessen fram til de endelige sentralområdeprognosene, fordi utviklingen innen hvert enkelt sentralområde ofte foregår ujevnt: Den generelle trendutvikling blir ofte forstyrret av utsatte/omgjorte byggeplaner, store kunder med mange abonnement som flytter eller “rydder opp” i sine abonnement, etc.

For å klargjøre prinsippene i denne måten å utarbeide prognosene på, skal vi ta utgangspunkt i situasjonen før prognosene for perioden 1990–2000 skulle utarbeides. Vi går altså ett år tilbake i tid i forhold til de eksemplene som er gitt på regions- og kommunenivå ovenfor. Dette fordi vi på dette tidspunkt faktisk utarbeidet prognosene på denne måten i Region Oslo.

Det vil føre for langt å skulle gjennomgå alle sentralområder i regionen. Vi tar derfor utgangspunkt i ett sentralområde, Holmlia sentralområde. Behandlingen av alle andre sentralområder følger samme mønster.

Tabell 10 viser de historiske data for total og årlig etterspørsel etter TAT og for Holmlia sentralområde i perioden 1982–1989.

Som vi ser har utviklingen i Holmlia sentralområde vært noe ujevn, med variasjon i årlig vekst fra høye 2880 i 1984 til lave 558 i 1989. Tar vi snittet av den årlige vekst over dataperioden gir det følgende resultat:

$$\frac{628 + 2880 + 1454 + 1185 + 1637 + 812 + 558}{7} = 1307,7 \quad (20)$$

Hvis prognosen utelukkende skulle baseres på gjennomsnittet av historikken for den årlige vekst, ble altså prognosen 1308 TAT pr år. Studerer vi den årlige etterspørsel år for år ser vi imidlertid at dette må være altfor høyt: Den årlige etterspørselen etter TAT i Holmlia sentralområde har (med unntak av 1987) avtatt hvert år siden toppen i 1984. Hvis

⁴ Dette skyldes at hver kommuneprognose har en usikkerhet. Når alle kommuneprognosene summeres opp, summeres derfor også denne usikkerheten.

denne utviklingen fortsetter, skal prognosen for 1990 i hvert fall ikke være høyere enn 558, og de derpå følgende år skal muligens ha en lavere årlig etterspørsel⁵.

Med bakgrunn i den informasjon som kom fram i diskusjonen i avsnittet ovenfor og lokal kunnskap om sentralområdet, ble derfor prognosene justert ned slik at etterspørselen for 1990 ble satt lavere enn det historiske snitt tilsier. Den årlige etterspørselen for resten av prognoseperioden ble også manuelt justert til å være avtakende i forhold til 1990-prognosen.

I den videre prosessen ble sentralområdet avstemt mot prognosen for Oslo kommune, slik at sum teleområdeprognoser = kommuneprognosen.

Prinsippene for utarbeidingen av prognosen for Holmlia sentralområde er vist i figur 6.

Figur 6 viser historikk for total etterspørsel etter TAT i perioden 1982–1989. For perioden 1990–2000 viser figuren tre mulige prognoser:

- 1 “Snitt-prognose” er en framskrivning på bakgrunn av snitt-veksten i historiske data, slik som nevnt ovenfor. Dette gir den høyeste prognosen.
- 2 “Manuell prognose” er således en manuell korrigering av “Snitt-prognose”, gjort på bakgrunn av kjennskap til/antakelser om markedsforholdene i sentralområdet, ref det som er sagt ovenfor.
- 3 Den siste prognosen er benevnt “Konsistent prognose”, og dette er resultatet etter at “Manuell prognose” er avstemt mot kommuneprognosen for Oslo kommune.

Som figur 6 viser, er det et klart avvik mellom snitt-framskrivningen og den manuelle framskrivning. På grunn av den sterke historiske veksten i dette sentralområdet, som for øvrig er et typisk boligområde som er relativt nytt, sender et

snitt av årlig vekst prognosene “rett inn i himmelen”. Dette tror vi selvsagt ikke på, og den manuelle prognosen anslår utviklingen slik vi tror den vil bli. Deretter lar vi kommuneprognosen være bestemmende for den endelige utforming av områdets utvikling, dvs det er den konsistente prognosen som gir det endelige prognoseforlag.

Eksemplet fra Holmlia sentralområde viser klart at metoder basert på ren snitt-vekst⁶ vil være lite egnet i mange tilfeller. Et sentralområde som Holmlia, som har vært under sterk utbygging, men hvor veksten har avtatt, får en svært uheldig framskrivning. En metode som legger mer vekt på de siste observasjoner enn på de tidligere, vil derfor være ønskelig.

5 Andre geografiske inndelinger

Vi har ovenfor gitt eksempler fra utarbeiding av regions-, kommune- og sentralområdeprognoser. I Televerkets nye organisasjon er det også opprettet såkalte teledistrikter innen hver region⁷. Siden vi har sikret oss konsistens mellom region, kommune og sentralområder kan derfor alle sentralprognosene innen et teledistrikt summeres til et valgt teledistrikts prognose. Sum av prognosene for teledistriktene i Region Oslo blir dermed lik prognosen for regionen totalt.

Det er selvfølgelig fullt mulig å utarbeide teledistriktsprognosene før sentralområdeprognosene og deretter sørge for konsistens mellom disse. Siden Oslo kommune er delt mellom flere av regionens teledistrikter, har vi imidlertid valgt å gå den andre veien; kommunene utgjør med sin befolkningsstatistikk et viktig hjelpenivå. For regioner og teledistrikt hvor kommuner ikke er delt mellom flere teleområder, kan altså forslaget være å gå veien om teledistriktsprognosene før sentralområdeprognosene utarbeides.

Referanser

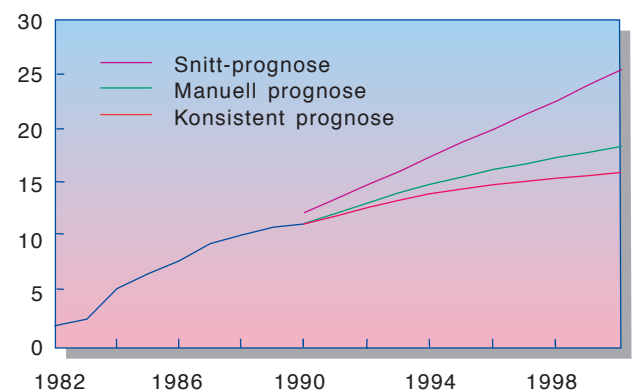
- 1 Hjelkrem, C, Bøe, J. Underlag for prognoser for telefonabonnement: data og analyse. *Teletronikk* 90(1), 65–75, 1994 (denne utgave).
- 2 Stewart, M B, Wallis, K F. *Introductory econometrics, 2nd edition*. Oxford, Blackwell, 1981.

- 3 Blindheim, E, Rygg, M. *En modell for å utarbeide prognoser for nye teletjenester*. Kjeller, Televerkets Forskningsinstitutt, 1982. (TF-Rapport 13/82.)
- 4 Stordahl, K, Holden, L. *Traffic forecasting models based on top down and bottom up models*. Kyoto, 1985. (ITC 11.)
- 5 ØSP. *Prognoser for etterspørsel etter hovedabonnement for telefon*. Oslo, Prognosekontoret Teledirektoratet, 1987.
- 6 Bøe, J, Hjelkrem, C. Utarbeiding av prognoser for etterspørsel etter telefonabonnement i Oslo teledistrikt. *Teletronikk* 87(1), 93–109, 1991.

Tabell 10 Total og årlig etterspørsel etter TAT for Holmlia sentralområde

År	Total etterspørsel	Årlig etterspørsel
1982	1548	
1983	2176	628
1984	5056	2880
1985	6510	1454
1986	7695	1185
1987	9332	1637
1988	10144	812
1989	10702	558

Totalt antall tilknytninger (x 1000)



Figur 6 Total etterspørsel etter TAT, Holmlia sentralområde

⁵ Dette er for øvrig i samsvar med teorien om at abonnementssetterspørselen kan beskrives ved en logistisk modell.

⁶ En prognosemetode som ligger nært opp til dette er enkel regresjon med tiden som eneste forklaringsfaktor.

⁷ Det er bare Divisjon Personmarked som har denne inndelingen.

- 7 Stordahl, K. Prognosemetoder: en oversikt. *Telektronikk* 90(1), 5–52, 1994 (denne utgave).
- 8 Makridakis, S, Wheelwright, S C, McGee, V E. *Forecasting: methods and applications, 2nd edition*. New York, John Wiley, 1983.
- 9 Log, S. Diffusjonsmodeller for nye teletjenester. Kjeller, Televerkets Forskningsinstitutt, 1982. (TF-Rapport 14/82.)
- 10 Winter, P R. Forecasting sales by exponentially wheighted moving averages. *Management Science*, 6, 324–342, 1960.

Tellerskritt – grunnlagsdata og prognoser

AV JOHANNES BØE OG CARLO HJELKREM

Resymé

Denne artikkelen vil beskrive hvordan tallmaterialet som blir hentet fra datasystemer som framkommer i interne statistikker blir justert avhengig av hvilken analyse som skal utføres. Med justering menes her justering for endring i tellerskrittsslengde, flytting av tjenester fra en takst til en annen, osv. Slike justeringer vil aldri bli like, selv om vi, som arbeider med dette på en profesjonell måte, i stor grad lykkes med å arbeide fram metodikk som til enhver tid gir oss en justering som kan forklares best mulig og som er lik år for år. Vi justerer historien slik at historien blir korrekt i

forhold til nåtiden og som dermed gir oss kontroll med framtiden. Jo bedre vi kjenner fortiden/historien, desto bedre kan vi fortelle om framtiden.

Målet med denne artikkelen er å beskrive en arbeidsmetodikk som benyttes i Region Oslo til forberedende analyser for videre å kunne lage prognoser. Tallmaterialet som blir presentert i artikkelen inneholder ingen realitet i forhold til virkeligheten, ettersom det er metodene og metodikken som er det viktigste.

1 Datagrunnlag for tellerskritt og analyse av disse

Før vi i det hele tatt kan begynne å analysere data må det skaffes til veie en oversikt over de data som er tilgjengelige og informasjon om hvilke fordeler og ulemper disse har. Vi vil her ta for oss de

datasystemer og hjelpemidler som det i dag er mulig å hente ut informasjon om tellerskritt fra. Videre vil bearbeidingen av tellerskrittdata bli utført og vise at dette kan gi nyttig informasjon både med hensyn på prognoser og markeds- og ledelsesinformasjon. I utgangspunktet vil en slik analyse være utført med hensyn på et prognosearbeid, slik at man kan si at

først utføres en analyse på data, så lages en prognose, men "biproduktene" vil likevel være til stede.

1.1 Krav til data

Hovedsakelig vil månedlige data for etterspørselen etter tellerskritt benyttes. For å kunne analysere slike data er det nødvendig å ha sammenhengende data eller observasjoner fra flere år tilbake (minst to års historikk av metode-tekniske årsaker, ideelt bør vi ha minst fire år historikk). Årsaken til dette er at analysen av dataene skal gi informasjon om sesongbetoningen og veksten i etterspørselen. Ut fra analysene lages det så modeller som beskriver hvorledes dataene varierer. Modellene brukes så til å predikere eller prognostisere den videre utviklingen.

1.2 Debiterte/påløpte tellerskritt

Etttersom Tellerarkivet i Region Oslo blir oppdatert i slutten av hver måned er det naturlig å ta utgangspunkt i månedsavlesninger. Det er derfor blitt laget programmer så tidlig som 1988 for å "plukke" data fra Tellerarkivet slik at muligheten for å få statistikk over debiterte tellerskritt er til stede. Ulempen med disse dataene er at det ikke finnes oppsplitting på takstklasser. Dette medfører at detaljeringsgraden er liten.

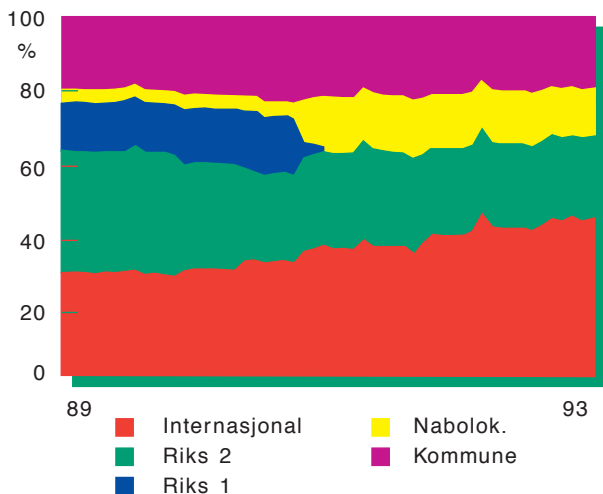
Tellerarkivet er grunnlaget for regningsutsendelsene og vil derfor inneha meget høy kvalitet. Da dataene er spesifisert ned på hver sentral vil dette gi oss mulighet til en riktig geografisk oppsplitting, hvis disse er definert på forhånd. I artikkelen tar vi kun for oss tallmateriale hentet fra et geografisk område i Region Oslo.

1.3 Tellerskritt fordelt på takstklasser

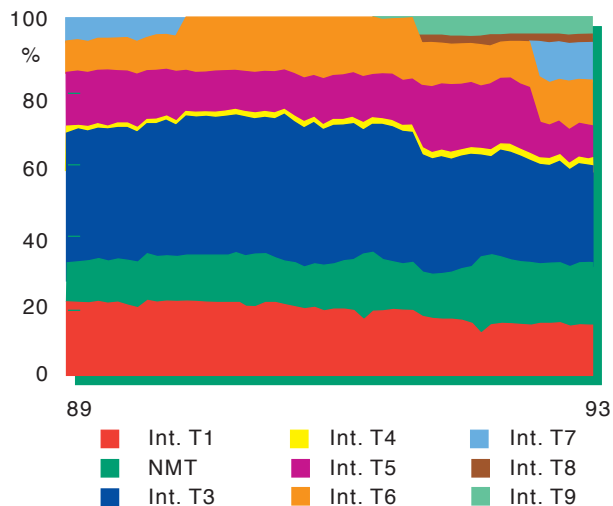
Det kan hentes tellerskritt ut fra de fleste sentraler som viser antall tellerskritt som er påløpt de respektive sentralene. Avlesningene er inndelt i takstklasser (internasjonale og nasjonale takstklasser). Disse tellerskrittene er direkte avlest fra tellere på sentralene, der hver takstklasse har sin egen teller. Dette er tellere som teller antall tellerskritt som genereres av abonnentene tilkoblet sentralene, fordelt på takstklasser.

Vi har nå totalt antall tellerskritt som genereres i vårt geografiske område og i kombinasjon med tellerskrittfordelingen mellom takstklassene er det mulig å beregne en total fordeling av tellerskrittene hentet fra Tellerarkivet.

Figur 1a og 1b viser prosentfordeling av tellerskritt for perioden 1989 til juni 1993, slik de er avlest. Det må gjøres oppmerksom på at de internasjonale takstklassene inneholder i tillegg til internasjonal tra-



Figur 1a



Figur 1b

001.18:621.39

Tabell 1 Oversikt over takstklassenes dekningsområder

Nasjonale takstklasser	Dekningsområde
Lokal takst	Innen egen kommune
Nærtakst	Nabokommunene
Fjertakst	Resten av landet/ åpen linje/televoting
Internasjonale takstklasser	Dekningsområde
Internasjonal takst 1	Norden
Internasjonal takst 2	NMT
Internasjonal takst 3	Europa
Internasjonal takst 4	Inmarsat
Internasjonal takst 5	USA Canada
Internasjonal takst 6	Nære østen med flere
Internasjonal takst 7	Opplysningstjenesten
Internasjonal takst 8	Australia ++
Internasjonal takst 9	Resten av verden

Tabell 2

Tidspkt.	Takstklasse	Fra	Til
09.89	Int. takst 5 full takst	4,5	5,3
	Int. takst 5 red. takst	6,0	6,4
	Int. takst 7	2,5	2,8
01.90	Rikstakst 2 full takst	22,5	27,7
	Rikstakst 2 red.takst	28,1	34,6
	Int. takst 1 full takst	20,6	21,4
	Int. takst 1 red. takst	25,7	26,8
	Int. takst 5 full takst	5,3	6,4
	Int. takst 5 red. takst	6,4	8,1
	Int. takst 7	2,8	3,3
07.90	Rikstakst 2 full takst	27,7	36,0
	Rikstakst 2 red.takst	34,6	45,0
01.91	Rikstakst 1 full takst	36,0	90,0
	Rikstakst 1 red.takst	45,0	180,0
	Nabolokal red.takst	270,0	180,0
	Int. takst 3 red.takst	10,7	13,4
	Int. takst 4	1,2	1,1
	Int. takst 6	3,3	3,1
	Int. takst 7	3,3	2,3

fikk også trafikk til blant annet Teletorg, Opplysningstjenesten og NMT.

2 Bearbeiding av underlagsdata

2.1 Mål for bearbeiding av data

Det aller første som må gjøres er å rydde data slik at de er sammenliknbare, ellers kan vi risikere å lage prognoser eller analyser på bakgrunn av et heller tvilsomt materiale. (Det er ingen god ide å lage prognoser for antall lastebiler i Norge på bakgrunn av totalbilsaget.)

Vårt mål med behandlingen av grunnlaget er å få direkte sammenliknbare data i de ulike avlesningsperiodene samtidig som et eventuelt sesongmønster tas vare på. Dette er viktig for at den statistiske modellen vi arbeider med skal kunne gi oss best mulig resultat uten å miste informasjon eller ødelegge kvaliteten på data. Vi antar at følgende variable påvirker våre data direkte:

- Tidligere utførte takstendringer
- Samtaletidsfordeling på de ulike takstklassene
- Antall virkedager i måleperiodene
- Abonnementsvariasjonen på de respektive målingene.

De ovennevnte variablene påvirker våre data slik at tidsserien vi har ikke er sammenliknbar, og det må derfor justeres for dette. Justering er nærmere beskrevet videre.

2.2 Takstendringer. En beskrivelse av takstklassene

Takstklassebenevninger som er brukt her er vist i tabell 1 sammen med dekningsområdene. Det må her nevnes at tidligere rikstakst 1 og nabolokal takst fra 01.01.91 er satt sammen til én takstklasse og har skiftet navn til *nærtakst*. Rikstakst 2 har skiftet navn til *fjertakst*.

Det må nevnes at tjenester som Teletorg takseres som internasjonal trafikk. Opplysningstjenesten er en tjeneste som har hoppet fra én takstklasse til en annen, men som pr i dag (1993) har fått egen takstklasse. Det har vært flyttinger av områder og tjenester innen de internasjonale takstklassene i perioden.

Avklaring med hensyn på hovedbegrepene

Tellerskritt:

Tellerskritt som til enhver tid er avlest. Disse er ikke justert for takstendringer. Enheten er tellerskritt.

“Volum” tellerskritt:

Tellerskritt som er avlest, justert for takstendringer (tellerskrittsslengde) som har vært utført i avlesningsperioden. Enhet er tellerskritt.

“Volum” tellerskritt er altså en konstruert tidsserie som er sammenliknbar over tid og uavhengig av endringer i tellerskrittsslengden.

Vi ønsker her å kunne benytte samme enhet “volum” tellerskritt, denne enheten er sammenliknbar over tid og uavhengig av endringer i tellerskrittsslengden. Dette er viktig ettersom tellerskrittsslengden har endret seg for hver takstklasse i tiden vi har avlesninger over. Dataene inneholder enkelte hopp/nivåskift og årsaken til dette er at et tellerskritt i 1989 ikke er like langt som et tellerskritt i 1992. Det må her presiseres at vi ønsker å justere for endringer i tellerskrittsslengden og ikke i pris pr tellerskritt.

Foretatte takstendringer (endring i tellerskrittsslengden)

Det er viktig å kunne justere data slik at de ikke inneholder nivåskift/hopp i data. Nivåskift i disse dataene er eventuelle takstendringer som har vært i avlesningsperioden. Det er om å gjøre å tilstrebe at historiske data er konstruert slik at de inneholder samme takststruktur som siste observasjon. De takstendringene våre tellerskrittsslengde data er justert for vises i tabell 2.

Muligheten til å kunne justere for endringer i tellerskrittsslengden er til stede ettersom dataene fra Tellerarkivet er beregnet oppsplittet på takstklasser. Dermed det ikke foretas en justering for takstendringer vil vi, som tidligere nevnt, ved hver takstendring få et hopp/nivåskift i dataene. Ved for mange slike nivåskift vil prognosemodellene våre begynne å bli ustabile og gi dårlige resultater. Etter denne justeringen vil dataene nå beskrive en kontinuerlig og sammenliknbar utvikling i tellerskritt basert på siste status i tellerskrittsslengde (her 01.01.93). Vi slipper dermed å bruke ekstra parametere i prognosemodellen for å forklare endringer i tellerskrittintervall, noe som vil medføre større usikkerhet i selve prognosen.

Tabell 3 Antall samtaler i prosent som har varighet på x antall sekunder for fjerntakst

Intervall		Prosent	Intervall		Prosent	Intervall		Prosent
fra	til		fra	til		fra	til	
0	5	2,58	50	55	2,47	140	145	1,83
5	10	5,80	55	60	2,27	150	155	1,70
10	15	5,75	60	65	2,11	160	165	1,58
15	20	5,16	70	75	3,79	170	175	1,47
20	25	4,57	80	85	3,32	180	185	1,38
25	30	4,05	90	95	2,95	190	195	1,29
30	35	3,62	100	105	2,64	200	205	1,22
35	40	3,26	110	115	2,39	210	215	1,15
40	45	2,95	120	125	2,17	220	225	1,09
45	50	2,69	130	135	1,99	230	235	1,03

Tabell 2 viser som eksempel de takstendringer som har vært utført i perioden 1989 – 1991. *Tidspkt.* angir måned og år takstendringen er foretatt. *Fra* og *Til* angir gammel og ny tellerskritt lengde i sekunder.

3 Samtaletidsfordeling og justering for tidligere takstendringer

For å kunne beregne en slags indeks for endring i tellerskritt lengden er det også nødvendig med informasjon om samtale-tidsmønsteret til abonnentene. Det er nødvendig å foreta slike målinger ved jevne mellomrom for å kontrollere at resultatene blir mest mulig riktige. Vi har her funnet målinger som er utført, og vi antar/forutsetter at dette kan anses som et snitt for abonnentene våre.

Tabell 3 er et utdrag av de målingene som er brukt i beregningene. Målinger

finnes opp til 600 sekunder (10 min) med 5 sekunders intervaller. Vi antar i utgangspunktet at den akkumulerte samtale-tidsfordelingen $F(t)$ følger følgende funksjon:

akkumulert samtale-tidsfordeling

$$F(t) = 1/(1 + e^{(a + bt^n)})^m$$

Ved å transformere dataene i tabellen over til lineær form lages det et nytt tallsett som vi benevner med t' . Altså:

$$t' = t^{-n}$$

og

$$F'(t) = \text{Ln}[(1/F(t'))^{1/m} - 1]$$

Dermed er

$$F'(t) = \alpha + \beta t'$$

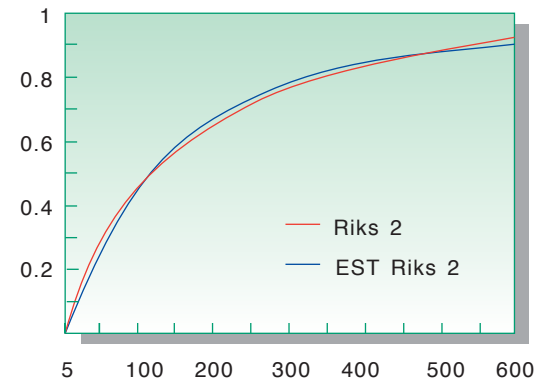
Konstantene m og n estimeres:

$$m = 20$$

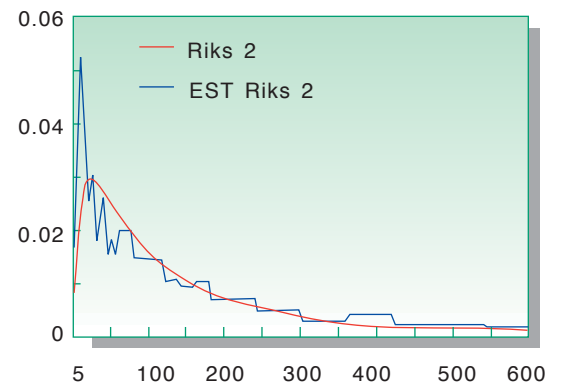
$$n = 0,25$$

Tabell 4 Resultater av regresjonsberegningene på data fra tabell 3

Takstklasse	Estimert a	Estimert b	Gitt m	Gitt n	Beregning av gjennomsnittlig samtale-tid i sek
Kommune	0,213964	-1,2386	20	0,25	167
Nabolok	0,500012	-1,22698	20	0,25	214
Riks 1	0,550997	-1,1936	20	0,25	247
Riks 2	0,403256	-1,14449	20	0,25	261
Utland	0,088702	-1,13096	20	0,25	216
Sum	0,24459	-1,2035	20	0,25	191



Figur 2



Figur 3

m og n er estimert til de verdiene som gir best tilpasning til de data som benyttes.

Det er nå mulig å bruke lineær regresjon på det transformerte tallsettet og dermed er beregningen til α og β enkel. Tilsvarende transformasjoner og regresjonsanalyse er beskrevet i spesialartikkel i denne utgave av *Elektronikk* og lærebøker om statistikk.

Resultatet av disse beregningene framkommer i figur 2, som viser målt og estimert akkumulert samtale-tidsfordeling for fjerntakst.

Vi har nå konstruert en estimert samtale-tidsfunksjon for samtale-tidsfordeling. Figur 3 viser avleste verdier og estimerte verdier for samtale-tiden fordelt på hver takstklasse.

Figur 3 viser målt og estimert samtale-tidsfordeling for fjern-takst.

For å begrense antall tabeller og regneeksempler vises nå beregningene for fjern-takst, full takst med de data som er tilgjengelig.

Tabell 5 Antall samtalesekunder 100 samtaler genererer, fordelt på takstklasser

Takstklasser	100 samtaler samtaletid i sek
Nasjonale takstklasser	
Lokal takst	14697
Nærtakst	18537
Fjerntakst	21550
Internasjonale takstklasser	
Internasjonal takst 1	17939
Internasjonal takst 2	17939
Internasjonal takst 3	17939
Internasjonal takst 4	17939
Internasjonal takst 5	17939
Internasjonal takst 6	17939
Internasjonal takst 7	17939
Internasjonal takst 8	17939
Internasjonal takst 9	17939

Beregning av gjennomsnittlig samtaletid gjøres ved hjelp av simulering og ikke matematisk. Dette vil gi en feil, men denne er så liten at vi kan se bort fra den.

Regresjonsberegningene er utført ved hjelp av et vanlig regneark, EXCEL, benyttet i Region Oslo. Dette gav resultater som vist i tabell 4.

Et estimat for samtaletiden kan nå sies å være på plass.

Når ovennevnte beregninger er utført er det en enkel sak å beregne hvor mange samtalesekund hvert tellerskritt utgjør i gjennomsnitt på hver enkelt takstklasse. Tabell 5 viser takstklassene med beregnet antall samtalesekunder 100 samtaler genererer, beregnet fra samtalestidsfordelingen, uavhengig av tellerskrittets lengden.

Videre må man finne ut hva de samme samtalene koster i antall tellerskritt avhengig av tellerskrittintervallet. Vi velger her kun å presentere Fjerntakst, men prinsippet er likt for alle takstklassene.

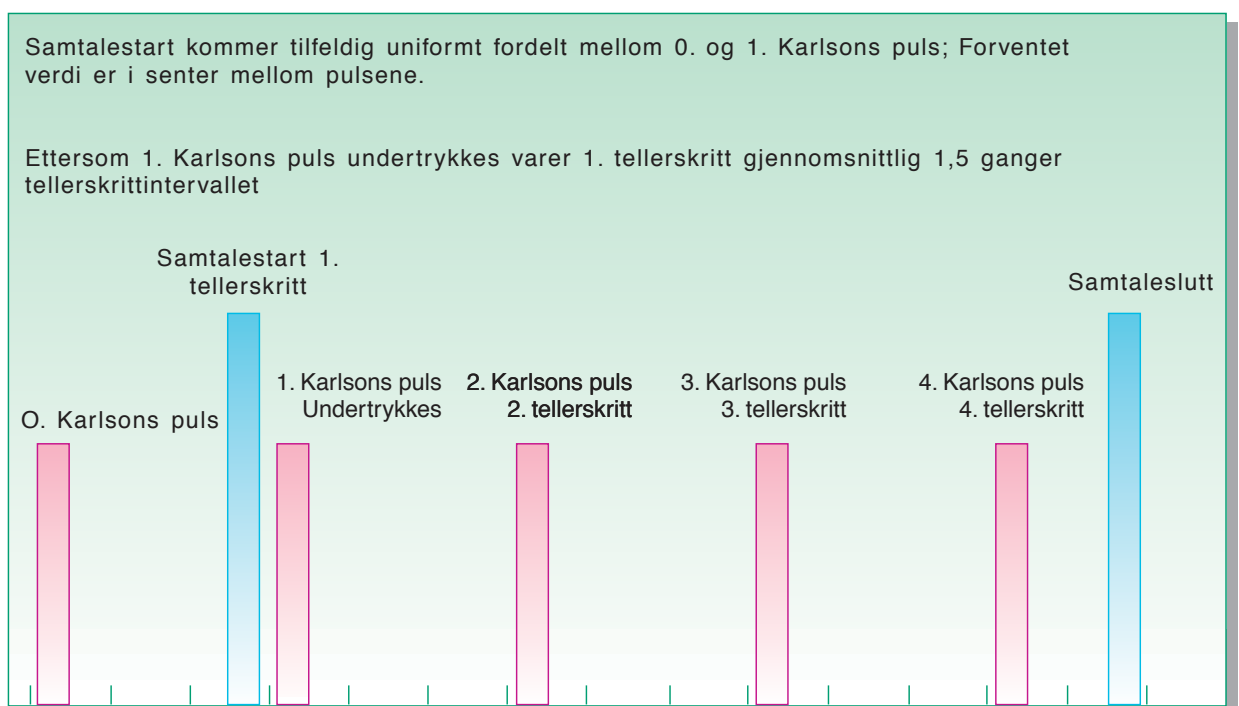
De forskjellige kolonnene viser antall tellerskritt som påløpte ved forskjellige priser/takstintervaller (som nevnt tidligere har det vært en del endringer). Det må her gjøres oppmerksom på at slik tell-

erskrittssystemet fungerer teknisk blir 1. Karlsons takstpuls undertrykket, noe som medfører at første tellerskritt vil være gjennomsnittlig 1,5 tellerskritt. Når en samtale blir gjennomkoplest blir det satt på ett tellerskritt. Denne samtalen kommer inn i et pulstog hvor hver puls har samme avstand i tid som takstintervallet. Når første puls kommer blir denne undertrykket, det vil si at den ikke blir talt. Samtalen har til nå kostet ett tellerskritt. Når neste takstpuls (2. Karlsons takstpuls) kommer blir denne talt, og samtalen har kostet to tellerskritt. Fra nå av teller hver takstpuls ett tellerskritt.

Figur 4 viser skjematisk hvordan pulstog med takstmarkeringer blir omgjort til tellerskritt avhengig av lengden på en samtale.

Ved å benytte ovennevnte tabell i kombinasjon med samtalestidsfordelingen er det mulig å beregne hva 100 samtaler på Fjerntakst koster (samme prosedyre utføres for hver takstklasse).

Når disse beregningene er utført for alle takstklasser vil man kunne lage en "indeks" for hver takstklasse. Og etter som våre tellerskrittdata er splittet opp på takstklasser kan en takstclasses



1
Figur 4

beregnete indeks benyttes til å si hvor mange tellerskritt som påløpte innen enhver tidsperiode.

Det er nå laget en indeks på hva 100 samtaler koster i forhold til siste takstnivå (dagens) for hver takstklasse. Denne indeksen vil fra nå av bli kalt *justeringsfaktor*. Som en forenkling har vi nå jobbet med tellerskritt lengder for såkalt fulltakst periode. For å gjøre resultatet virkelig nøyaktig må dette også gjøres for reduserttakst periode og de to indeksene veies. Veiingen av de to indeksene må gjøres på bakgrunn av trafikkmålinger/tellerskritt målinger i fulltakst- og reduserttakst perioden. Målinger utført på sentraler i Region Oslo viser at ca 20 – 25 % av inntektene kommer i redusert tid. Det må også utføres en veiing av takstklassenes justeringsfaktorer slik at det framkommer en *total justeringsfaktor*

for summen av alle takstklasser. Dette gjøres for å forenkle beregningene videre.

Regneøvelsen som er beskrevet over bør utføres både for full og redusert takst og veies sammen med 0,25 for redusert takst og 0,75 for full takst. Vektene er framkommet fra målinger utført på timebasis.

Vi har nå klargjort den informasjon som er nødvendig for å kunne justere de avleste tellerskrittene slik at alle avlesningene er sammenliknbare og enheten er etter beregningene *volum tellerskritt*.

4 Justering for variasjon i antall abonnement

Et problem med dataene er antall abonnement som genererer disse tellerskrittene. Det har vært variasjoner i

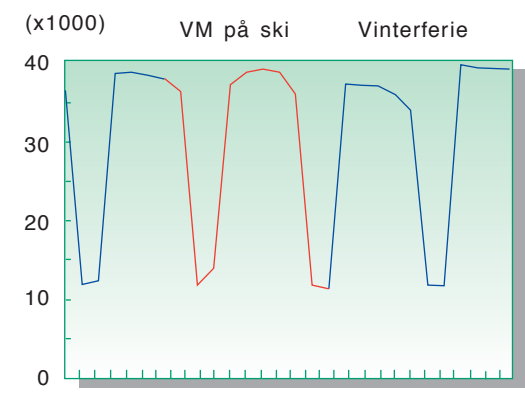
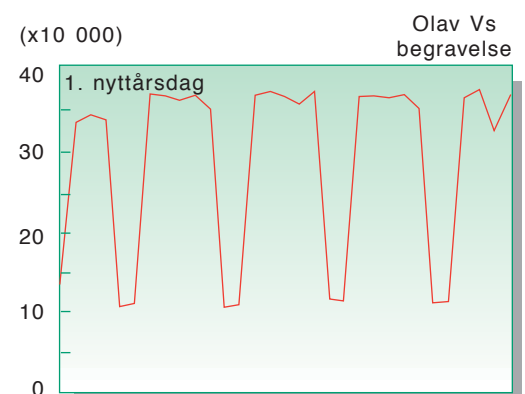
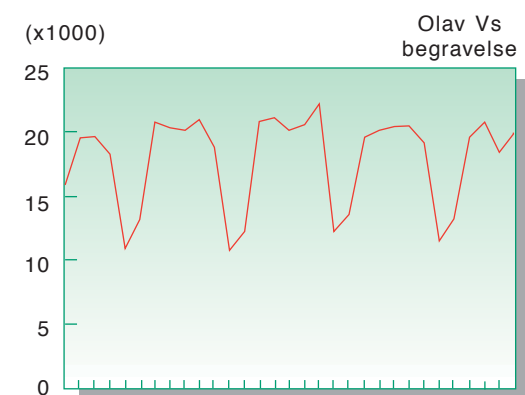
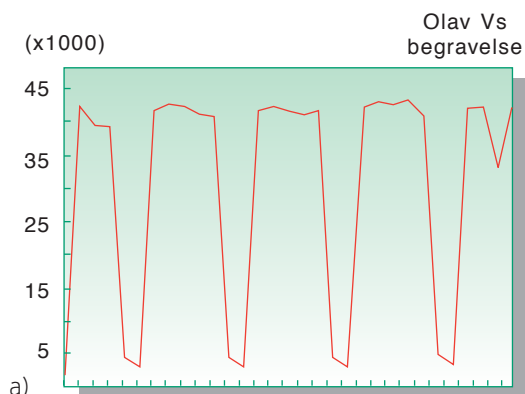
abonnementsmassen, som igjen medfører variasjon i avleste tellerskritt. Ved å konstruere et nytt datasett hvor enheten er volum tellerskritt pr abonnement pr måned, er variasjonen abonnementsmassen ivarettatt. Dataene har etter justering med hensyn på antall abonnement fått enhet *volum tellerskritt pr abonnement*.

4.1 Justering for variasjon i antall virkedager

Det må også justeres for alle variable helligdager, lørdager og søndager samt “inneklemt” dager. Når dette er gjort er periodene sammenliknbare fra år til år, vi har altså konstruert en tidsserie med sammenfallende sesong på 12 perioder. Denne tidsserien har enheten *volum tellerskritt pr abonnement pr virkedag*.

Tabell 6 Hvor mye en samtale med “gjennomsnittsfjerntakst” full takst koster avhengig av lengden på samtalen, lengden på tellerskrittet og hvilke tidsrom det er

		22,5	22,5	27,7	36	36	40	40
		01.01.89 – 31.08.89	01.09.89 – 31.12.89	01.01.90 – 30.06.90	01.07.90 – 31.12.90	01.01.91 – 31.12.91	01.01.92 – 31.12.92	01.01.93 – 30.06.93
fra	til							
0	5	1	1	1	1	1	1	1
5	10	1	1	1	1	1	1	1
10	15	1	1	1	1	1	1	1
15	20	1	1	1	1	1	1	1
20	25	1	1	1	1	1	1	1
25	30	1	1	1	1	1	1	1
30	35	2	2	1	1	1	1	1
35	40	2	2	1	1	1	1	1
40	45	2	2	2	1	1	1	1
45	50	2	2	2	1	1	1	1
50	55	2	2	2	2	2	1	1
55	60	3	3	2	2	2	2	2
60	65	3	3	2	2	2	2	2
65	75	3	3	3	2	2	2	2
75	85	4	4	3	2	2	2	2
85	95	4	4	3	3	3	2	2
95	105	5	5	4	3	3	3	3
105	115	5	5	4	3	3	3	3
115	125	6	6	5	3	3	3	3
125	135	6	6	5	4	4	3	3
135	145	6	6	5	4	4	4	4
145	155	7	7	6	4	4	4	4
155	165	7	7	6	5	5	4	4
165	175	8	8	6	5	5	4	4
175	185	8	8	7	5	5	5	5
185	195	9	9	7	5	5	5	5
195	205	9	9	7	6	6	5	5
205	215	10	10	8	6	6	5	5
215	225	10	10	8	6	6	6	6
225	235	10	10	8	7	7	6	6



Figur 5 a–d

Det er mulig å kunne avlese antall på-løpte tellerskritt for hver dag som går. Dette gir oss en mulighet for å kunne finne ut hvor mye helligdager, lørdag, søndager og inneklemt dager skal telle/veie i forhold til en vanlig “arbeidsdag”.

Generelt kan man si at vektene på de forskjellige dager i året er basert på faktiske målinger. Disse målingene viser at:

Vekt

0,3 Lørdag/søndag har vekt 0,3 (det er kun disse dagene som har redusert takst)

0,2 – ,07 Helligdager, offentlige fridager og inneklemt dager måles spesielt

1 Resten

Dette er en betraktning som kun vil gjelde Region Oslo, men den kan antakelig brukes av andre også. Å bruke vektene her vil tross alt være bedre enn om vi ikke tar hensyn til slike variasjoner. Slike vekter er forskjellige fra sentral til sentral, fra kommune til kommune og dermed fra region til region ettersom vektene er avhengige av hvilken fordeling det er mellom bolig- og forretningsabonnement og takststruktur. Figurene 5a, b, c og d viser en typisk forretnings-sentral og en typisk boligsentral. Felles-trekkene er at det påløper vesentlig flere tellerskritt mandag til fredag enn lørdag og søndag, samt at tallserien er svært stabil. Forskjellen er at forretnings-sentralen ikke har nevneverdig aktivitet lørdag og søndag, mens boligsentralen derimot har betydelig aktivitet. Det kan nesten se ut til at boligsentralen er mer aktiv på lørdag og søndag enn ellers, da tatt i betraktning reduserte priser.

Figur 5a, b, c og d viser dagsmålinger for januar og februar 1991 for en bolig-sentral og en forretnings-sentral. Som en kuriositet kan vi se effekten av kong Olavs begravelse i slutten av januar samt effekten av VM på ski 1991 og vinterferien for skolene.

For å understreke hvor viktig slike justeringer er, er det bare å trekke fram påske og pinse som veksler mellom henholdsvis mars/april og mai/juni. Ved å korrigere for antall virkedager vil slike effekter nesten elimineres og de rene sesongvariasjonene kommer bedre fram. Figur 6 viser dette tydelig.

Figur 6 viser registrerte data før eventuelle justeringer.

Figur 7a og b viser volum tellerskritt/abonnement/dag slik at sesongutslagene avtegnes som et tydelig mønster som gjentar seg år for år.

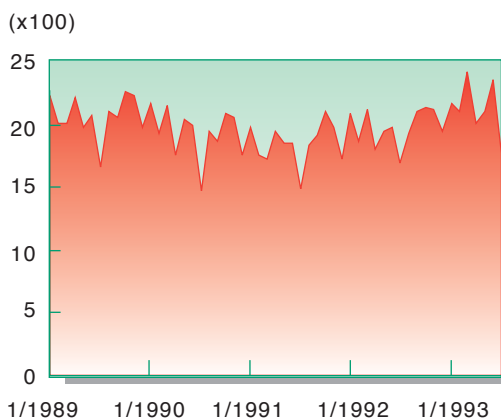
4.2 Resultat av alle justeringer

Det ses ut fra de grafiske kurvene i figur 7a og 7b at alle data har falt på plass i et riktig mønster i forhold til hverandre. Sesongmønsteret er relativt likt for hvert år og kommer tydelig fram. Det som nå gjenstår er å lage ett/flere prognosealternativ på bakgrunn av denne tidsserien.

Det må her nevnes at volum tellerskritt ikke må oppfattes som endelige (de er derfor ikke uforanderlige). Det vil som regel være en kontinuerlig oppfølging av volumberegningene (justeringene). En slik oppfølging vil bli foretatt hver gang det er endringer i tellerskrittintervallet for en takstklasse.

5 Tellerskrittprognoser, valg av metode

Det må gjøres oppmerksom på at tallmaterialet i artikkelen ikke er offisielle tall og representerer ikke Region Oslo. Hensikten med kapittelet har vært å kunne skissere en framgangsmåte og ikke beskrive Region Oslos prognoser og grunnlagsmateriale for produksjon av prognoser for tellerskritt. Det må også nevnes at slike prognoser er benyttet som et hjelpemiddel i inntektsprognosene for regionen på totalnivå samt som grunnlag for oppsplitting på divisjoner.



Figur 6

Tabell 7 Hva 100 samtaler på Fjerntakst koster avhengig av hvilket tellerskrittintervall det til enhver tid er

Periode	01.01.89 – 31.08.89	01.09.89 – 31.12.89	01.01.90 – 30.06.90	01.07.90 – 31.12.90	01.01.91 – 31.12.91	01.01.92 – 31.12.92	01.01.93 – 30.06.93
Tellerskritt- intervall	22,5	22,5	27,7	36	36	40	40
Tellerskritt	962	962	782	605	605	549	549
	549/962	549/962	549/782	549/605	549/605	549/549	549/549
Indeks	0,57069	0,57069	0,70205	0,90744	0,90744	1,0	1,0

5.1 Aktuelle prognoser

Målsettingen med dette arbeidet har vært å få bedre oversikt over inntektsutviklingen fra tellerskrittene. Det er derfor utviklet prognosemodeller som produserer prognoser på månedsbasis. Disse gir grunnlag for

- å lage prognoser for neste års budsjett
- å kunne monitorere månedlig utvikling av tellerskritt og inntekter
- å lage månedlige prognoser for resten av budsjettperioden
- å lage periodisert inntektsbudsjett
- å lage prognoser for langtidsbudsjettet (neste fem-års periode).

Når det gjelder å estimere effekten av nedsatte takster, er dette svært vanskelig. Dette skyldes blant annet abonnentenes oppfatning av endringene i takststrukturen.

5.2 Modell og grunnlagsdata benyttet i arbeid med korttidsprognoser for tellerskritt

De prognosemodeller/metoder som det er aktuelt å bruke, er:

- Regresjonsmodell med sesongkomponenter
- ARIMA-modeller
- Transferfunksjonsmodeller
- Holt-Winters metode
- Kalmanfilter.

Vi velger her å benytte Kalmanfiltermodellen, fordi

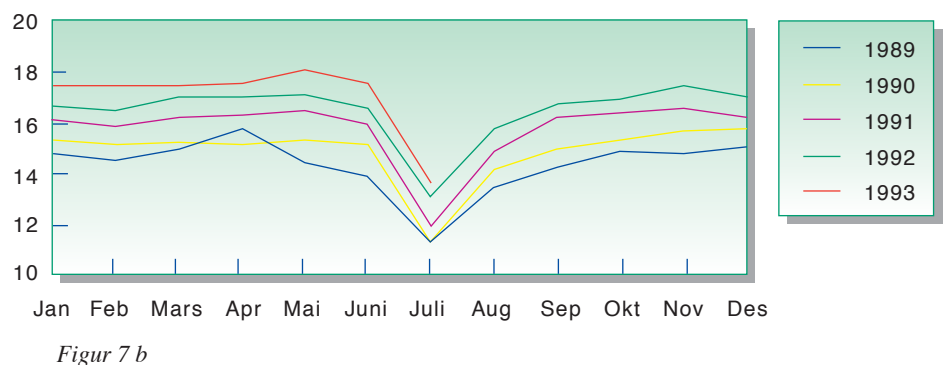
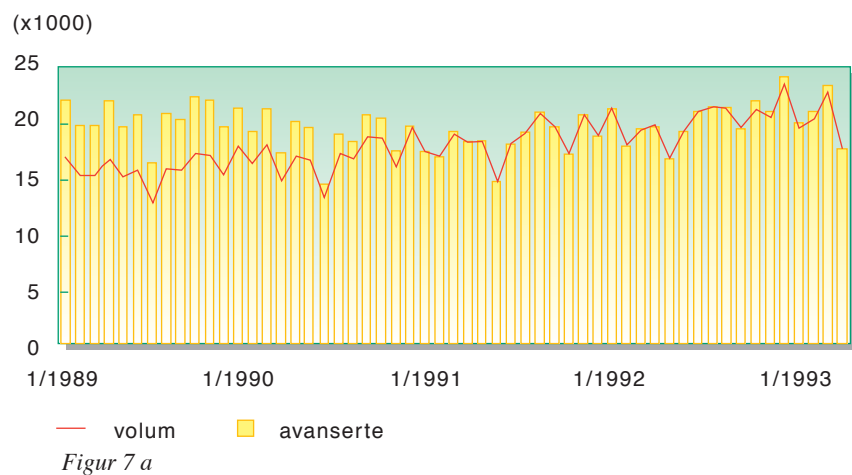
- modellen er enkel å bruke
- den gir stabil god kvalitet på prognosene
- den takler såkalt "outlier" observasjoner.

Grunnlagsdata for korttidsprognoser

Tidligere i artikkelen ble det etter en nøysommelig prosess definert en tidsserie som vi skulle bruke som grunnlagsdata i forbindelse med å utarbeide prognoser for tellerskritt. Vi skal nå benytte oss av det grunnlaget som er presentert tidligere i tabellform til å lage prognoser for det område (de sentraler) som er valgt.

5.3 Prognoser basert på Kalmanfiltermodellen

Kalmanfiltermodellen er meget avansert, som tidligere beskrevet i innledningsartikkel. Modellen har en del funksjoner som kan/bør utnyttes, dog ikke ukritisk. Dette er bl a nivåskift, som vi kan spesifisere tidspunkt på eller modellen kan søke etter og estimere nivåskiftet automatisk. I tillegg estimerer modellen såkalte "outliers" ved å angi disse på samme måte som ved nivåskift. "Outliers" er observasjoner som ligger utenfor



Tabell 8 viser de avleste tellerskrittene som finnes og omregnet fra direkte avleste tellerskritt til volum tellerskritt (justerte) pr abonnement pr virkedag

Periode	Abonnement	Avleste tellerskritt	Volum tellerskritt	Virkedager	Avleste ts/ab /v.dag	Volum ts. /ab./v.dag	Total takstjustering
1989	46508	22071	16949	24,65	19,25	14,78	0,767946
Februar	46673	19927	15303	22,4	19,06	14,64	0,767946
Mars	46815	19968	15334	21,95	19,43	14,92	0,767946
April	46893	22021	16911	23	20,42	15,68	0,767946
Mai	46990	19700	15129	22,3	18,80	14,44	0,767946
Juni	46970	20717	15909	24,4	18,08	13,88	0,767946
Juli	46877	16508	12677	24	14,67	11,27	0,767946
August	47051	20835	16000	25,4	17,43	13,39	0,767946
September	47233	20359	15803	23,7	18,19	14,12	0,776190
Oktober	47347	22413	17397	24,7	19,17	14,88	0,776190
November	47544	22013	17086	24,4	18,98	14,73	0,776190
Desember	47658	19743	15324	21,6	19,18	14,89	0,776190
1990	47846	21345	18069	24,8	17,99	15,23	0,846510
Februar	48045	19208	16260	22,4	17,85	15,11	0,846510
Mars	48206	21322	18049	24,7	17,91	15,16	0,846510
April	48296	17429	14754	20,25	17,82	15,09	0,846510
Mai	48402	20223	17119	23,15	18,05	15,28	0,846510
Juni	48432	19635	16621	22,85	17,74	15,02	0,846510
Juli	48437	14639	13317	24,7	12,24	11,13	0,909696
August	48510	19095	17370	25,4	15,50	14,10	0,909696
September	48622	18422	16759	23	16,47	14,99	0,909696
Oktober	48763	20713	18842	25,4	16,72	15,21	0,909696
November	49099	20466	18618	24,4	17,08	15,54	0,909696
Desember	49172	17557	15971	20,6	17,33	15,77	0,909696
1991	49289	19784	19702	24,8	16,19	16,12	0,995860
Februar	49461	17516	17443	22,4	15,81	15,74	0,995860
Mars	49604	17075	17004	21,2	16,24	16,17	0,995860
April	49688	19245	19165	23,75	16,31	16,24	0,995860
Mai	49824	18342	18266	22,3	16,51	16,44	0,995860
Juni	49881	18464	18387	23	16,09	16,03	0,995860
Juli	49862	14926	14864	25,4	11,79	11,74	0,995860
August	49952	18308	18232	24,7	14,84	14,78	0,995860
September	50089	19134	19054	23,7	16,12	16,05	0,995860
Oktober	50274	20915	20828	25,4	16,38	16,31	0,995860
November	50435	19759	19677	23,7	16,53	16,46	0,995860
Desember	50528	17311	17239	21,3	16,08	16,02	0,99586
1992	50697	20692	20754	24,8	16,46	16,51	1,002987
Februar	50839	18843	18899	22,7	16,33	16,38	1,002987
Mars	50913	21334	21398	24,7	16,96	17,02	1,002987
April	50974	18023	18076	20,95	16,88	16,93	1,002987
Mai	51082	19373	19431	22,35	16,97	17,02	1,002987
Juni	51150	19755	19814	23,55	16,40	16,45	1,002987
Juli	51117	16853	16892	25,4	12,98	13,01	1,002329
August	51322	19234	19279	24	15,62	15,65	1,002329
September	51571	20977	21025	24,4	16,67	16,71	1,002329
Oktober	51743	21407	21457	24,7	16,75	16,79	1,002329
November	51845	21228	21278	23,7	17,28	17,32	1,002329
Desember	51943	19495	19540	22,2	16,91	16,95	1,002329
1993	52052	21865	21225	23,4	17,95	17,43	0,970741
Februar	52359	21019	20404	22,4	17,92	17,40	0,970741
Mars	52565	24080	23376	25,4	18,04	17,51	0,970741
April	52648	20069	19482	20,95	18,20	17,66	0,970741
Mai	52714	20989	20375	21,45	18,56	18,02	0,970741
Juni	52720	23312	22630	24,4	18,12	17,59	0,970741
Juli	52721	17690	17690	24,7	13,58	13,58	1,000000

Tabell 9 Volum tellerskritt pr abonnement pr virkedag

Måned	Volum tskritt/ab./virkedag	Volum tskritt/ab./virkedag	Volum tskritt/ab./virkedag	Volum tskritt/ab./virkedag	Volum tskritt/ab./virkedag
	1989	1990	1991	1992	1993
1	14,78	15,23	16,12	16,51	17,43
2	14,64	15,11	15,74	16,38	17,40
3	14,92	15,16	16,17	17,02	17,51
4	15,68	15,09	16,24	16,93	17,66
5	14,44	15,28	16,44	17,02	18,02
6	13,88	15,02	16,03	16,45	17,59
7	11,27	11,13	11,74	13,01	13,58
8	13,39	14,10	14,78	15,65	
9	14,12	14,99	16,05	16,71	
10	14,88	15,21	16,31	16,79	
11	14,73	15,54	16,46	17,32	
12	14,89	15,77	16,02	16,95	

det som man kan kalle "normalen" (2 ganger standardavviket). Slike "outliers" kan forekomme ved feilmålinger, situasjoner i samfunnet som normalt ikke forekommer, osv. I dette eksempelet vil vi ikke bringe inn faktorer som ovennevnte nivåskift eller outlier.

Som resultat viser modellen med vårt grunnlag en trend med 4,3 % årlig vekst og videre et standardavvik avrundet til 0,0. Vi kan da med god sikkerhet, 95 % sannsynlighet, si at trenden som er ved siste observasjon er 4,3 % årlig vekst.

Figur 8 viser en prognose for tellerskrittsvolum pr abonnement pr virkedag.

En slik prognose kan i korthet kalles en trendprognose. Vi estimerer her kun en trend med sesongkomponenter og framskriver disse. For å se det hele i

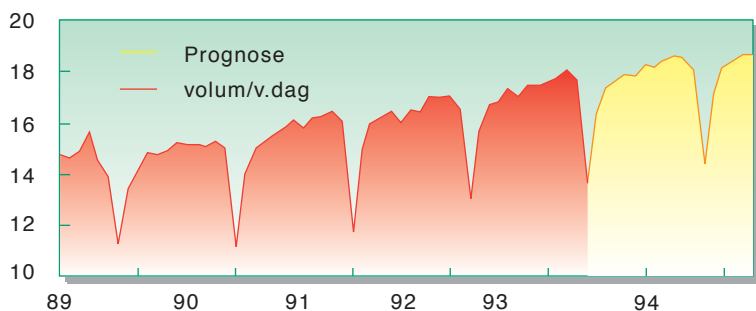
Tabell 10 Volum tellerskritt (justerte) pr abonnement pr virkedag. Uthevede tall viser hvordan prognosen ble

Måned	Volum tskritt/ab./virkedag					
	1989	1990	1991	1992	1993	1994
1	14,78	15,23	16,12	16,51	17,43	18,21
2	14,64	15,11	15,74	16,38	17,40	18,05
3	14,92	15,16	16,17	17,02	17,51	18,35
4	15,68	15,09	16,24	16,93	17,66	18,52
5	14,44	15,28	16,44	17,02	18,02	18,44
6	13,88	15,02	16,03	16,45	17,59	17,99
7	11,27	11,13	11,74	13,01	13,58	14,34
8	13,39	14,10	14,78	15,65	16,31	17,04
9	14,12	14,99	16,05	16,71	17,30	18,03
10	14,88	15,21	16,31	16,79	17,63	18,36
11	14,73	15,54	16,46	17,32	17,84	18,58
12	14,89	15,77	16,02	16,95	17,74	18,47

sammenheng må den prognostiserte trenden settes inn i de totale tellerskrittinntektene og på en slik måte at endringer i abonnementsmassen og endringer i antall virkedager blir tatt hensyn til. Ved å sette opp en tabell og beregne de størrelsene som er aktuelle, får vi følgende resultater.

Prognosen er i Volum tellerskritt/abonnement/virkedag. For å finne totalinntektene må det multipliseres opp med antall abonnement (et estimat for antall innkoplete abonnement) og antall virkedager for de respektive måneder (virkedagene beregnes på samme måte som tidligere beskrevet). Ved å summere sammen for hvert år kan man enkelt se hvilke faktorer som påvirker inntektene og hvilken vekt de forskjellige innehar. Vi ser av tabell 11 at antall virkedager for 1994 bidrar med 0,2 % av årsresultatet, veksten i abonnement (gjennomsnitt over året) utgjør 3,0 %, trenden bidrar med 4,3 %, og totalt gir dette en volumvekst fra 1993 til 1994 på rundt 7,5 %. Det må gjøres oppmerksom på at dette er volumvekst i antall tellerskritt, alle tall er her presentert i lik tellerskrittintervall over hele perioden.

Tabell 11 viser altså et anslag på utviklingen for volum tellerskritt under like forhold i prognoseperioden som i historisk tid. Det må nå vurderes hvilke faktorer som pr i dag er kjent og som vil komme til å påvirke volumutviklingen i tidsperioden vi ser på. Det kan være informasjon om endringer i tjenester som



Figur 8

Tabell 11 Komplette datasett med historikk og prognoser slik det normalt presenteres. Uthevet tall viser prognosen

Periode	Virke- dager	Volum tskritt	Innk. abonn.	Volum ts/ab/vdag	% endring fra året før
1993	23,4	21225	52052	17,43	5,57 %
Februar	22,4	20404	52359	17,40	6,23 %
Mars	25,4	23376	52565	17,51	2,90 %
April	20,95	19482	52648	17,66	4,35 %
Mai	21,45	20375	52714	18,02	5,87 %
Juni	24,4	22630	52720	17,59	6,95 %
Juli	24,7	17690	52721	13,58	4,42 %
August	24,7	21321	52926	16,31	4,21 %
September	24,4	22446	53174	17,30	3,54 %
Oktober	24	22572	53346	17,63	5,01 %
November	24,4	23266	53449	17,84	3,02 %
Desember	22,4	21278	53547	17,74	4,69 %
1994	23,95	23401	53656	18,21	4,50 %
Februar	22,4	21818	53962	18,05	3,75 %
Mars	23,55	23408	54168	18,35	4,81 %
April	22,1	22205	54251	18,52	4,85 %
Mai	22,15	22186	54317	18,44	2,33 %
Juni	24,4	23846	54324	17,99	2,26 %
Juli	24	18696	54324	14,34	5,56 %
August	25,4	23601	54529	17,04	4,48 %
September	24,4	24099	54778	18,03	4,22 %
Oktober	24	24213	54950	18,36	4,14 %
November	24,4	24958	55052	18,58	4,15 %
Desember	22,35	22766	55150	18,47	4,11 %

	1989	1990	1991	1992	1993	1994
Virkedager	282,50	281,65 -0,30 %	281,65 0,00 %	283,45 0,64 %	282,60 -0,30 %	283,10 0,18 %
Abonnenter	47047	48486 3,06 %	49907 2,93 %	51266 2,72 %	52852 3,09 %	54455 3,03 %
Volum pr dag pr ab.	14,28	14,77 3,44 %	15,64 5,87 %	16,37 4,64 %	17,14 4,75 %	17,88 4,31 %
Volum totalt	189822	201749 6,28 %	219863 8,98 %	237844 8,18 %	256065 7,66 %	275197 7,47 %

i perioden vil endre abonnentenes ringe-interesse, f.eks. ved etablering av attraktive tjenester som kan bidra med kraftig vekst på grunn av høy pris, eller ved reduksjon i tilbud av dyre, attraktive tjenester. Prognosen over beskriver volum tellerskritt og må ikke forveksles med trafikkutviklingen til abonnentene. Innbakt i prognosen ligger eventuelle vridninger av trafikkinteressen fra rimelige takserte tjenester til dyrere tjenester, f.eks. overgang fra lokaltakst til fjerntakst (eventuelt Teletorg). For å kunne beskrive slike vridninger kreves en langt større analyse enn den som er beskrevet her.

5.4 Presentasjon av prognosene

Når prognoser skal presenteres må dette gjøres på en mest mulig oversiktlig måte, slik at den/de som mottar og skal eventuelt revurdere eller bruke disse forstår hvordan man er kommet fram til resultatene. Det er her ikke nødvendig å snakke om modeller og andre tekniske/statistiske begrep, men en ren presentasjon av data og resultater.

Vi skal nå vise en enkel presentasjon av både korttidsprognosen og langtidsprognosen. I prinsippet ønsker vi å gjøre presentasjonen på så lite papir som mulig, kombinert med grafiske framstillinger. Presentasjonen er her tenkt som presentasjon overfor ledelse. Overfor ledelse er det best å presentere resultatene (og datagrunnlaget) grafisk samt noe informasjon påskrevet de grafiske framstillingene.

Informasjon til bruk for videre analyse og annet arbeid innen økonomi, plan og marked vil kreve mer detaljerte tabeller – slike vil bli vist senere og krever langt mer papir.

5.5 Oppfølging av prognosene og rapportering av styringsparametere

Det er viktig å kunne følge opp prognosene etter hvert som tiden går både med hensyn til treffsikkerhet, se foregående avsnitt, og hvordan prognosen vil påvirke totaltall for året. Dette vil kunne gi korreksjoner med hensyn på styringsparametere og når det gjelder spesielle hendelser i løpet av året.

Utviklingsparametere som er nyttige i oppfølgingsammenheng er følgende:

Glidende 12 måneder gjennomsnitt. Tallet beregnet for januar 1990 er gjennomsnittet av volum tellerskritt/

ab./v.dag for februar 1989 til og med januar 1990.

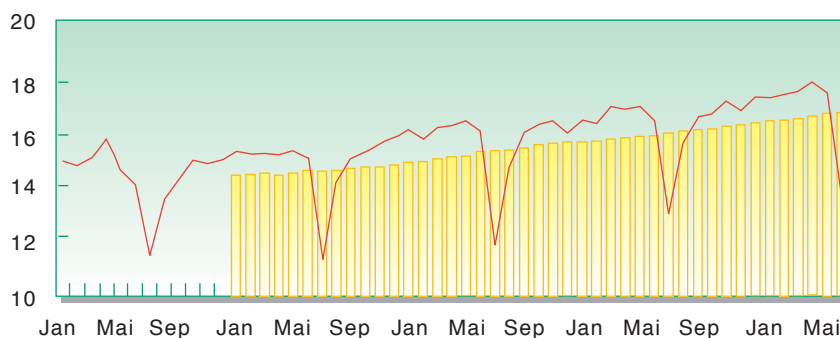
En prosentvis endring i volum tellerskritt pr abonnement pr virkedag mot samme periode året før.

Figur 9a og b viser tidsseriene de to ovennevnte tidsseriene sammen med volum tellerskritt pr abonnement pr virkedag.

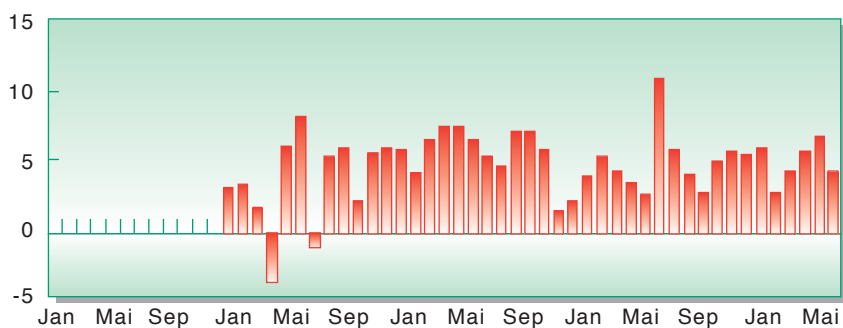
Figur 9a viser for volum tellerskritt/abonnement/virkedag og glidende 12 måneder gjennomsnitt.

Figur 9b viser prosentvis endring i volum tellerskritt pr abonnement pr virkedag mot samme periode året før.

Tabell 12 viser volumutvikling for hele år (for 1993 er januar – juli avlesninger, august – desember og hele 1994 er prognose). I 1993 ser det med denne prognosen ut til å bli en volumvekst på 7,66 %. Denne kan igjen splittes opp i en endring i antall virkedager på -0,3 %, en økning i antall abonnement på 3,09 % og til slutt en økning i volum tellerskritt pr abonnement pr virkedag på 4,75 %. Videre er det satt opp sammenlikning i totalt avleste tellerskritt og prosentvis endring av dette år for år.



Figur 9 a



Figur 9 b

Tabell 12 Årsresultatene splittet opp på enkeltkomponenter

	1989	1990	1991	1992	1993	1994
Virkedager	282,50	281,65	281,65	283,45	282,60	283,10
		-0,30 %	0,00 %	0,64 %	-0,30 %	0,18 %
Abonnenter	47047	48486	49907	51266	52852	54455
		3,06 %	2,93 %	2,72 %	3,09 %	3,03 %
Volum pr dag pr ab.	14,28	14,77	15,64	16,37	17,14	17,88
		3,44 %	5,87 %	4,64 %	4,75 %	4,31 %
Volum totalt	189822	201749	219863	237844	256065	275197
		6,28 %	8,98 %	8,18 %	7,66 %	7,47 %

Prognoser for nye tjenester hvor data er mangelfulle

AV JOHANNES BØE OG CARLO HJELKREM

001.18:621.39

Case: ISDN prognoser, valg av metode, modell og grunnlagsdata

1 Innledning

Formålet med denne artikkelen er å belyse de metoder som kan brukes når det skal lages prognoser for nye tjenester (tjenester vi ikke har erfaring med).

En ny teletjeneste (eller nye tjenester generelt) har i de fleste kjente tilfeller en begynnelse/innføringsfase, en "stabil" vekstfase og en metningsfase der markedet kan sies å være i metning. Vi skal her prøve å belyse, ved hjelp av et eksempel, hvordan det går å kombinere generelle kunnskaper om utvikling av nye tjenester, strategibeslutninger og markedsinformasjon, samt vurderinger av eksperter/eksperttips sammen med statistiske og matematiske funksjoner.

ISDN er en ny tjeneste i dag. Vi har lite erfaringmateriale og det finnes heller ikke historiske tall kanskje bortsett fra prøvetjenesten, hvor ISDN er prøvd av kunder. Derfor er ISDN en meget aktuell

tjeneste å se på som et case på hvordan man kan lage prognoser for nye tjenester. Uttalelser som at "ingen kan lage gode prognoser" passer spesielt godt i ISDN-sammenheng (ettersom det her ikke finnes historiske data eller erfaringstall). Men enkelte kan lage bedre prognoser enn andre, fordi de sitter inne med opplysninger, metodekunnskap og informasjon om strategier andre ikke har eller ikke ser betydningen av. Vi har forsøkt å kombinere metodekunnskap med de opplysningene som finnes, slik at "best mulig" prognoser kan framkomme.

Vi vet at utviklingen av nye tjenester er noenlunde lik på enkelte områder uansett hvilken tjeneste det gjelder. Dette er vist i spesialartikkel om S-kurver. Der ble det fokusert på at etterspørselen etter en ny tjeneste hadde visse likhetstrekk som kan utnyttes. Vi gjentar enkelte figurer her for å repetere likhetstrekkene.

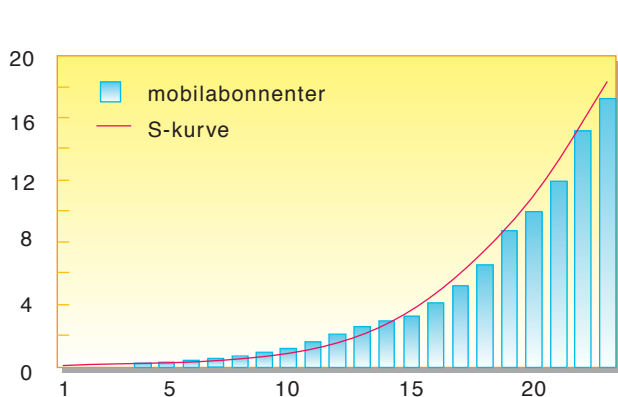
Figur 1 viser utviklingen på mobiltelefon sammenliknet med tilpasset S-kurve.

Figur 2 viser utviklingen på teleks sammenliknet med tilpasset S-kurve.

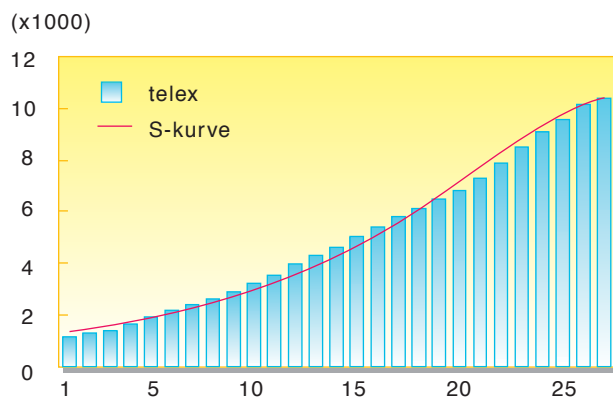
Figur 3a viser utviklingen på mobiltelefon sammenliknet med tilpasset S-kurve og forlenget S-kurve.

Figur 3b viser utviklingen på teleks sammenliknet med tilpasset S-kurve og forlenget S-kurve.

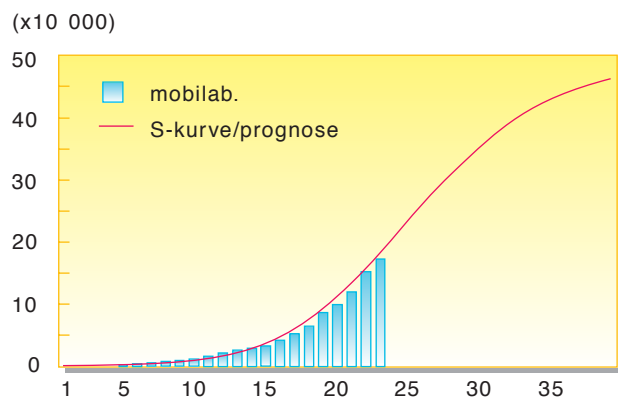
Som artikler i denne utgave av Teletronikk har omtalt tar vanligvis prognosene i Televerket utgangspunkt i historisk utvikling. Disse prognosene baseres på modeller som er testet på tidligere historisk utvikling. Dermed vil ISDN-prognosene, i motsetning til de fleste andre prognosene i Televerket, ikke kunne bygges på historisk utvikling, fordi dette gjelder en ny tjeneste. Et annet problem er mangel på informasjon, som vanligvis brukes under prognosearbeid med andre nett og tjenester. Dette kan være informasjon som kan brukes enten direkte eller indirekte, og som vi kan anta at etterspørselen etter ISDN kan sies å være avhengig av. Ettersom vi har begrenset mengde med



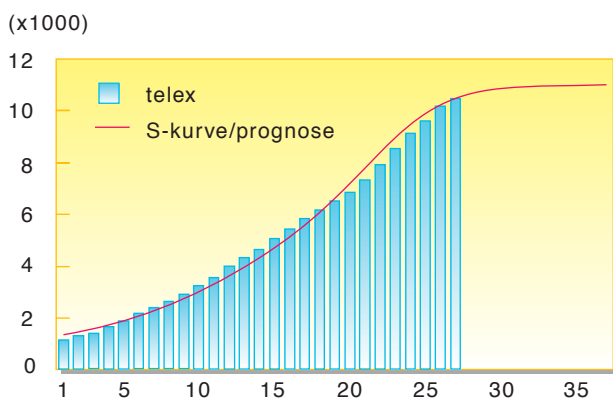
Figur 1 Utviklingen på mobiltelefon sammenliknet med tilpasset S-kurve



Figur 2 Utviklingen på teleks sammenliknet med tilpasset S-kurve



Figur 3a Utviklingen på mobiltelefon sammenliknet med tilpasset S-kurve og forlenget S-kurve.



Figur 3b Utviklingen på teleks sammenliknet med tilpasset S-kurve og forlenget S-kurve

informasjon må vi finne en modell med relativt få parametere. Etterspørselsfunksjonen må også velges slik at den er robust (det vil si stabil, lite påvirkelig) de første årene (det er jo de som er viktigst for en ny tjeneste som ISDN, sett fra plansiden).

ISDN-prognosene kan brukes som underlag for:

1 Strategiarbeid

Prognosene vil danne underlag for en rekke strategiske beslutninger, takst-arbeid, vurderinger i forhold til andre nett og tjenester, etc.

2 Teknisk planlegging

Prognosene er underlag for planlegging av hvilke sentraler som på gitte tidspunkt skal oppgraderes til ISDN-sentraler. Dette for å kunne dekke behovet etter ISDN-aksesser på den teknisk/økonomisk mest hensiktsmessige måte. Hver sentral vil da måtte vurderes i en samlet plan for utbygging, basert på prognosene for hver endesentral. Prognosene vil også danne basis for trafikkdimensjonering av nett (og sentraler).

3 Økonomisk planlegging

På bakgrunn av planene for teknisk utbygging av ISDN får vi tall for investering/drift. Ut fra prognosene og takstene får vi inntektstall for tilknytningsavgift og fastavgiftene. Sammen med framtidige trafikk-prognoser vil vi også få prognoser for trafikkinntektene.

4 Markedsplanlegging/markedsføring

Markedsplanlegging vil blant annet bestå i gjennomføring av markedsundersøkelser som også danner grunnlag for prognoser. Men prognosene vil også danne grunnlag for videre markedsundersøkelser og markedsføring. I tillegg kan prognosene danne grunnlag for fastsettelse av salgsmål.

Innføringen av ISDN har så store konsekvenser for den tekniske utbyggingen og for økonomien at ISDN-prognosene bør ofres tilstrekkelig oppmerksomhet og være så gode som mulig helt fra starten av (muligheten for feilinvesteringer er størst i startfasen). Ved for høye prognoser blir bestillingene for høye og dermed kapitalkostnadene større enn nødvendig. Er derimot prognosene for lave, vil tapte inntekter på grunn av for liten kapasitet ha stor betydning, samt tapte inntekter på grunn av at Televerkets kunder går over til andre private tjeneste-

leverandører fordi vi da ikke klarer å dekke kundenes behov.

Gode ISDN-prognoser vil betinge et godt kjennskap til markedet og kundenes behov.

2 Beskrivelse av tilgjengelig informasjon

Prisinformasjon om ISDN er et vanskelig tema ettersom prisene i skrivende stund ikke er bestemt. Vi må forutsette at pris på ISDN i stor grad vil være lik med telefoni.

Markedsundersøkelser er utført nasjonalt, men gir oss lite nyttig informasjon i prognosearbeidet, bortsett fra at markedet viser stor interesse for ISDN. Det er usikkert om interessen skyldes:

- Tjenestene i ISDN
- Digital linjetilknytning
- Ønske om innvalg
- Rimelig løsning, tatt pris for PABX i betraktning.

Problemet med denne type markedsundersøkelser er at respondenten ofte vet lite om hva ISDN innebærer. Videre har vi en mengde undersøkelser som viser at behovet for kommunikasjon er voksende også i små bedrifter og i husstandsmarkedet. Undersøkelser viser at kundene i første omgang kun fokuserer på pris, rimelig tilknytningsform. Først når de har fått en rimelig tilknytning og brukt denne en periode vil de spørre seg om hva ISDN kan benyttes til utover vanlig telefoni.

Televerkets egne dataregistre inneholder en mengde informasjon som kan brukes i prognosearbeidet. Denne informasjonen er vanskelig tilgjengelig hvis man ikke har kompetanse til å jobbe direkte på databasene for å lage spesialrapporter.

3 Valg av grunnlagsdata

Et av motivene for ISDN er å møte konkurransen om (med) våre strategisk kunder. Dette er kunder som for en stor del etablerer interne nett motivert ut fra bedre og billigere datakommunikasjon. Videre er eksisterende forslag til takstsystem i ISDN utformet slik at det er mest å "spare" for brukere med datakommunikasjon ved en overgang til ISDN. Dette betyr at datakommunikasjon er en av hovednøkklene til utbredelsen av ISDN i første fase.

Vi velger likevel kun å ta utgangspunkt i telefonbrukerne og deres behov når grunnlagsdataene skal vurderes. Dette gjøres av flere årsaker, blant annet fordi telefonbrukerne er den største gruppen av abonnenter som har fordel av ISDN og at alle telefonbrukere vil få gjennomsnittlig overkapasitet i B-kanaler (telefonkanaler) som kan nyttes til dataoverføring (dette vises senere). Videre er datatjenestene et incitament for å etterspørre ISDN, og dette har stor betydning i første fase av innføringen av ISDN. Større bedrifter har gjerne et stort data- og tekstkommunikasjonsbehov, dette vil vi ta hensyn til når prognosen skal vurderes.

Ettersom vi først og fremst har konsentrert oss om telefonabonnement og ISDN, ønsker vi å hente data fra *Abonnementssystemet* og bruke disse direkte i prognosearbeidet. Vi kan da si at dataene vi bruker er generelt tilgjengelig for alle sentralområder. Dette vil være til stor hjelp ettersom flere områder kan behandles likt og arbeidsmengden da blir mindre enn hvis hvert område skulle behandles separat. Det kan allikevel være grunnlag for å behandle spesielle områder annerledes, typiske kan være store sentralområder med en konsentrert forretningsvirksomhet, eller områder med kun boligbebyggelse.

Vi anbefaler gjennom eksemplet at markedet skilles i tre segmenter. Segmentene er oppdelt ut fra antall telefonabonnement kundene har.

- 1 Små og middels forretningsvirksomheter
- 2 Store forretningsvirksomheter
- 3 Husstandsabonnenter.

Grunnlaget fra *Abonnementssystemet* er brukt for å beregne markedspotensieller for aksesstyper innenfor bolig- og forretnings-segmentene (disse segmentene blir nærmere beskrevet senere). Det må gjøres oppmerksom på at slike beregninger aldri vil bli bedre enn kvaliteten på informasjonen hentet fra databasen.

Grunnlaget fra *Abonnementssystemet* viser antall abonnenter med x antall abonnement som har brukskategori forretningsabonnement som er koplet i gruppe. Tabell 1 viser et eksempel på en slik rapport.

3.1 Vurderinger av foreliggende informasjon som har betydning for alle prognosene

Prognosene bygger på at ISDN er generelt tilgjengelig fra 1994, enten som direkte tilknytning til en ISDN-sentral

Tabell 1 Oversikt antall forretningsabonnenter for en gruppe sentraler

Antall bylinjer	Sentral1	Sentral2	Sentral3	Sentral4
1	3432	5876	2819	1349
2	225	368	195	98
3	168	250	117	58
4	127	218	98	42
5	49	71	35	18
6	60	137	48	23
7	17	36	16	7
8	28	62	12	10
9	10	11	8	5
10	21	38	11	9
11	5	12	6	3
12	8	27	8	5
13	6	10	2	3
14	5	16	5	1
15	4	7	3	1
16	10	17	4	1
17	2	4	1	0
18	2	6	2	0
19	4	5	0	2
20	6	18	3	2
21-25	8	20	2	3
26-30	3	13	4	5
31-35	3	10	1	2
36-40	6	2	4	2
41-45	0	5	2	0
46-50	2	5	0	0
51-100	11	11	1	2
101-150	3	4	1	0
151-200	2	1	0	0
201-250	0	1	0	0
251- 999	0	2	0	0

eller som fjernetilknytning. Fram til 1994 vil det være prøvedrift med et begrenset antall kunder.

Tallene for 1993 er basert på kjennskap til prøvetjeneste-abonnement, sentralutstyr og terminalutstyr som skal brukes under prøvetjenesten. Det er også gjort vurderinger ut fra forslag til pris på tilkopling og abonnementsavgift i ISDN i forhold til dagens telefonipriser, samt priser og prisutvikling på terminalutstyr. I tillegg kommer også dagens priser på tilsvarende tjenester i datanettene og telefonnettet. Spesielt kan abonnentenes antatte årlige fastavgifter i ISDN i forhold til data- og telefoninettet nevnes som en av de viktigste faktorene for å bestemme markedspotensialet for ISDN. Det kan her listes en del faktorer som kan være til hjelp i disse vurderingene:

- Priser/antatt prisutvikling på terminalutstyr
- Avskrivningstid på disse og beslektet utstyr
- Andel av bedrifter som benytter datakommunikasjon
- Andel av bedrifter som benytter teleks/telefaks
- Utskiftningstakt på hussentraler
- Avskrivningstid på hussentraler
- Kostnad for utskiftning/oppgradering av hussentraler
- Forskjell mellom små og store bedrifters behov/kostnader
- Hva vet bedriftene om ISDN, og hvor stor er påvirkningen fra utstyrsleverandørene
- Informasjon om planlagte applikasjoner /mulige applikasjoner.

4 Segmentering av markedet

Brukere av ISDN kan grovt deles inn i tre hovedgrupper, ut fra naturlige skiller i grunnlagsdata og tilknytningsform i ISDN:

- Husstander
- Små bedrifter
- Store bedrifter.

Denne grove inndelingen er valgt fordi det er mangel på sikre data når grunnlagsdataene brytes ned på markedssegmenter. Hvis tilstrekkelig informasjon hadde vært til stede, kunne man laget markedssegmenterte prognoser med basis i applikasjoner/bransjeløsninger.

I det følgende skal vi se på hvilke muligheter ISDN gir disse gruppene.

4.1 Husstander

Gruppen kalt husstander kan igjen deles inn i to undergrupper:

- De "vanlige" (stort flertall)
- De "avanserte" (lite mindretall).

Aktuelle fellestrekk for disse er følgende:

- En vanlig telefonlinje
- Ett eller flere biapparater
- Alarm (f eks trykkløst alarm for eldre).

De avanserte vil i tillegg også kunne ha behov for datakommunikasjon til PC.

A Den vanlige husstand i ISDN:

For husstander med lite behov utover telefoni innebærer ISDN

- bedre telefonkvalitet (digital telefonlinje)
- mange nye muligheter (les tilleggstjenester)
- doblet kapasitet på ut/inngående samtaler.

Vi regner ikke disse kundene som aktuelle etterspørere etter ISDN før etter år 2000.

B Den avanserte husstand i ISDN:

For denne gruppen betyr ISDN (grunn-tilknytning, 2B+D)

- tilknytning av opp til 8 terminaler på samme linje
- mulighet for samtidig bruk av flere terminaler
- mange flere muligheter på samme linje
- større fleksibilitet i bruken.

Tre viktige faktorer påvirker disse brukerne til å ta i bruk ISDN:

- 1 Lave inngangskostnader, enten gjennom billige "black box"-er (terminaladaptere), software for å utvide deres eksisterende utstyr (PC-er o l)

eller billige ISDN-apparater (terminaler)

- 2 Behov hos abonnenten som dekkes av tjenester i ISDN (f eks hjemmebank, elektronisk postkasse, kjøp og salg gjennom telenettet, alarm, billedoverføring, osv)
- 3 Bedriften betaler abonnement for at kunden kan ha hjemmekontor.

4.2 Små og mellomstore bedrifter

Små og mellomstore bedrifter er her definert som kunder med behov for $n \times 2B+D$ aksesser.

Generelt vil nok utnyttelsen av ISDN blant annet være proporsjonalt med den bruken bedriften har av EDB og telekommunikasjon i vid forstand i dag.

Blant disse små og mellomstore bedriftene – i telekommunikasjonssammenheng – kan det være små lokalavdelinger av større bedrifter eller selvstendige foretak fra kolonialbutikker med 1–2 ansatte til bedrifter med inntil ca 100 ansatte.

I dag har disse muligens en eller flere av disse alternativene:

- Apparatsystem/liten hussentral
- Telefaks (muligens et eget abonnement til denne)
- Teleks/teletex
- PC (selvstendig eller koplet sammen i LAN)
- Datakommunikasjon mot et av Televerkets datanett
- Leid linje inn mot eventuelt hovedkontor.

Dette krever telefonabonnement (kanskje tilkoppelt en liten hussentral/apparatsystem), pluss abonnement på teleks/teletex og noen datalinjer.

I dag innebærer dette også ulike abonnementsforhold for hver tjeneste (telefon, teleks, data), med hver sine særtjenester, oppkopplingsfunksjoner, priser, regningsrutiner og grensesnitt mot Televerket.

Med $n \times 2B+D$ aksesser vil ISDN kunne gi denne gruppen:

- telefon, teleks, data- og bildetenester på samme “veggkontakt”
- raskere overføring – f eks telefaks gruppe 4
- alle teletjenester tilgjengelig fra starten (terminalavhengig)

- samtrafikk med andre nett
- muligheter for kombinasjon av telefon og data
- ett abonnement med én regningsrutine
- ett grensesnitt mot Televerket.

4.3 Store bedrifter

Denne type bedrifter/organisasjoner er store nok til å ha sin egen telekommunikasjonsavdeling eller telekommunikasjonsansvarlig (gjerne i kombinasjon med EDB-drift) for koordinering av (kommunikasjon) virksomheten både internt og eksternt. Mange av disse bedriftene har allerede bedriftsinterne data- og hussentralnett og klarer på den måten å levere sine ansatte kvasi-ISDN tjenester. Ettersom teknologien er kjent for dem og satsingen på kommunikasjon er stor, vil antakeligvis denne type kunder være meget interessert i ISDN. Ikke bare av økonomiske årsaker, men også for å få et forsprang på andre konkurrenter og for å øke sin egenkompetanse innen ISDN.

5 Beskrivelse av valgt metode og etterspørselsfunksjon

Man kan si at det er et begrenset antall abonnenter med et bestemt behov for ISDN-aksesser i dag. Det vil igjen si at dagens totale marked har et begrenset behov for antall ISDN-aksesser, avhengig av antall abonnenter og abonnentenes telekommunikasjonsbehov. Samtidig tilsier erfaringen at antall abonnenter og deres behov endrer seg relativt langsomt sett over tid.

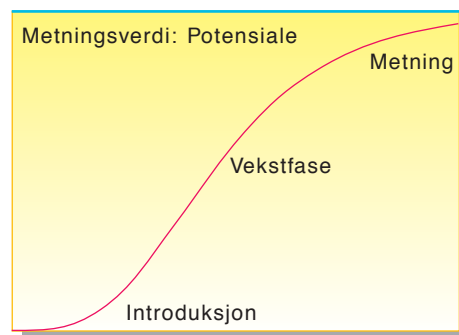
Siden vi ikke kan gjøre analyser på bakgrunn av historisk utvikling, må vi trekke slutninger på bakgrunn av erfaringer fra andre tjenester og innsikt i den nye tjenesten, her ISDN. Vi må derfor velge en etterspørselsmodell som hjelper oss i beskrivelse av framtiden. Etterspørselsmodellen som er valgt er en såkalt S-kurve (også kalt metningsfunksjon, potensialfunksjonen eller diffusjonsmodell). Detaljert informasjon om modellene er gitt i spesialartikkel i denne utgaven av *Elektronikk*. Dette er en modell som tidligere har vist seg å passe godt til prognoser for nye produkter og tjenester (teledata, mobiltelefon, ...). Modellen har en kurveform som likner en S, og det antas at etterspørselen følger denne. (Se figur 4 og 5.) Modellen beskriver i første del av perioden en økende årlig vekst inntil

modellen når et vendepunkt, den årlige veksten når et maksimum, og fra det tidspunkt vil modellen avta år for år og til slutt nærmer seg metning. Endringen i etterspørselen vil aldri være negativ. Det vil derfor med denne modellen ikke være mulig å beskrive en negativ etterspørsel. Kort beskrevet er dette en modell som sier at jo flere som har tjenesten, desto flere ønsker å ta tjenesten i bruk inntil man når et vendepunkt og den totale etterspørselen går mot en metning.

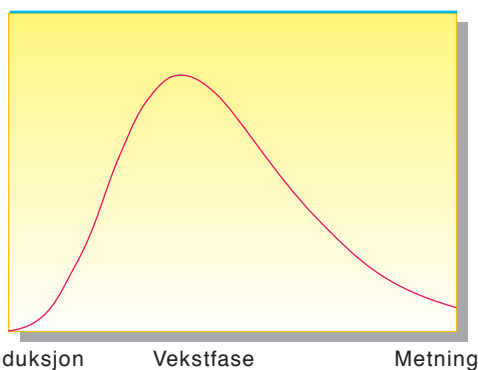
Figur 4 viser hvordan markedet mettes. En tenker seg at markedspotensialet er det totale marked. Dette markedet (markedspotensialet) er bestemt ut fra dagens abonnenter og deres antatte behov for ISDN-aksesser.

Figur 5 viser hvordan årlig vekst endrer seg med tiden. I første periode øker veksten i etterspørselen hvert år, og etter en viss tid når den årlige veksten et maksimum ved “vendepunktet”, og deretter avtar etterspørselen for hvert år. Den totale etterspørselen nærmer seg til slutt totalmarkedet/ metningspunktet.

En utvidet variant av funksjonen kan også med fordel benyttes for langsiktige prognoser for vanlig telefonabonnement. I Region Oslo har det vist seg at den



Figur 4



Figur 5

passer utmerket, med relativt lav feilmargin.

S-kurven er også tidligere testet av Televerkets Forskningsinstitutt og rapporten forelå august 1982 (TF-rapport R13/82). Den viste at S-kurven/ metningsfunksjonen gav et rimelig godt bilde av virkeligheten. Modellen ble da testet på gamle data og resultatet ble sammenliknet med nyere data. Det vil si at modellen ble testet på data fra blant annet Telex tjenestens utvikling de første årene for så å sammenlikne resultatene med de senere målingene. Konklusjonen ble at modellen (S-kurven) kunne brukes til å beskrive utviklingen av nye teletjenester de første årene etter en introduksjon.

Med hensyn på ISDN har vi ennå ikke tilstrekkelige data slik at skal modellen kunne brukes må det gjennomføres markedsundersøkelser/analyser for å fastslå markedspotensialet for tjenesten. I tillegg må det anslås minimum 2 verdier i forskjellig tidspunkt for etterspørselen. Markedspotensialet og de to verdiene kan ansees som konstanter som anslås på bakgrunn av generell informasjon og intelligent gjetting.

5.1 Beskrivelse av parametere i modellen

I tabell 1 var det en oversikt over antall kunder og hvor stort behov disse har i dag. Slike oversikter er hentet fra TEL-SIS og er brutt ned på sentralområder. Disse er som nevnt i kap. 3 segmentert i tre grupper:

- Store forretningskunder med behov for $n \times 30B+D$
- Små forretningskunder med behov for $n \times 2B+D$
- Boligkunder (disse har kun behov for én $2B+D$ aksess).

Det er samtidig foretatt en totalvurdering av resultatene og det har da blitt utført små korreksjoner på hver enkelt gruppes behov/framtidig behov. Store forretningskunder med behov for $n \times 30B+D$ får gjennomsnittlig i ISDN en overkapasitet, og vi har dermed vært noe strengere i beregningene av markedspotensialet for disse. Mindre forretningskunder som har behov for $n \times 2B+D$ har vi derimot gitt noe ekstra på grunn av liten overkapasitet. Boligabonentene har vi forutsatt normalt ikke har behov for mer enn én $2B+D$ aksess.

Videre må det også gjøres anslag av startverdier og verdi nummer to på kurven. Disse to verdiene settes direkte inn i prognosen for de forskjellige markedssegmentene. Startverdiene for forretningsabonentene er med utgangspunkt i innmeldt behov i prøvetjenesten, mens startverdi for boligabonentene er basert på skjønn. Det samme er verdi nummer to for alle tre segmentene. Disse verdiene er viktig for at vi skal kunne beregne etterspørselsfunksjonens prinsipielle forløp, og samtidig ta hensyn til markeds-
potensialet/metningspunktet.

Etterspørselsfunksjonen har følgende matematisk form:

$$F(t) = \frac{POT}{(1 + e^{(\alpha + \beta * t)})^c}$$

hvor:

- $F(t)$ er prognosen for etterspørselen etter tjenesten uttrykt i antall aksesser i tidspunkt t
- POT er et uttrykk for markeds-
potensialet
- t er tidsfaktor med år som måleenhet
- α og β er koeffisienter som estimeres.

For å kunne estimere α og β må det bestemmes fire konstante parametere i modellen:

- Verdi på c : anbefalt i TF-rapport R13/82 satt til 20000
- Markedspotensialet (vil være avhengig av prisforhold mellom $2B+D/30B+D$, og data fra TELSIS)
- Startverdi (avhengig av prøvetjenesten)
- Verdi nummer to (avhengig av markeds-
potensialet).

Disse parametrene må fastsettes på forhånd med grunnlag i informasjon som er tilgjengelig og intelligent gjetting. På bakgrunn av disse beregnes stigning og nivå på kurvens prinsipielle forløp. Disse beregningene er utført ved hjelp av regresjon. Det henvises til artiklene om metningsmodeller og regresjon i denne utgaven av Telelektronikk angående mer informasjon om regresjon.

Metoden går i korthet ut på innen hvert segment å bestemme markedspotensialet og to verdier, samt å bruke regresjon (statistiske beregninger) for å estimere parametrene (α og β). Deretter settes disse inn i modellen og det kan beregnes en prognose innen hvert segment.

- Verdi på c : anbefalt i TF-rapport R13/82 til 20000
- Startverdi: settes på bakgrunn av antall abonnement i prøvetjenesten
- Verdi nr 2: antatt verdi i år 2000 basert på et anslag av dagens behov som har ISDN-abonnement. År 2000 er valgt fordi året ligger utenfor planperioden. Det vises til slutt at verdien ikke påvirker de første 3 prognoseårene (planperioden) vesentlig
- Markeds-
potensialet er basert på erfaringer fra segmenterte prognoser utført for telefoniabonnement.

5.2 Eksempel på beregning av dagens markeds- potensial fordelt på aksesstyper og markedssegmentene bolig og forretning

Ettersom abonnementspris for ISDN i skrivende stund ikke er vedtatt, forutsetter vi at prisforholdet mellom en $2B+D$ aksess og en $30B+D$ aksess er 1 til 7. Samtidig forutsetter vi at to vanlige telefoniabonnement erstattes med en $2B+D$ aksess, fire telefoniabonnement erstattes med to $2B+D$ aksesser og så videre. Det er mulig å kategorisere forretningsabonentene i to kategorier: små forretningskunder som trenger x antall $2B+D$ aksesser og store, som trenger x antall $30B+D$. Ut fra forutsetningen om prisforholdet mellom $2B+D$ og $30B+D$ (1:7) vil de kunder som har behov for x $2B+D$ aksesser heller velge én $30B+D$ aksess. Det må her nevnes at brukerstyret ikke er kalkulert inn i prisforholdet, noe som i praksis vil endre på hvorvidt en kunde velger $2B+D$ eller $30B+D$.

Ut fra forutsetningen kan vi anta følgende:

De abonnenter som har inntil 12 abonnement for telefon i dag vil en gang i framtiden velge et gitt antall $2B+D$ aksesser. De abonnenter med mer enn 12 abonnement i dag vil med ISDN velge $30B+D$. Disse grensene danner grunnlag for å bestemme hvilke markeds-
potensialer som er aktuelt å benytte for de to ulike aksessformene. Grensen på 12 abonnement kan virke usikker fordi vi kun tar hensyn til forutsetning om Televerkets egen pris på aksessstype og ikke hussentralleverandør-
enes utstyrpris, men vi skal vise til slutt at markeds-
potensialet ikke er av avgjørende betydning de tre til fire første årene.

Tabell 2 Små forretningsabonnenters behov for 2B+D pr i dag

Antall abonnement	Behov 2B+D	Abonnenter	2B+D aksesser
1	1	9494	9494
2	1	482	482
3	2	399	798
4	2	296	592
5	3	135	405
6	3	156	468
7	4	54	216
8	4	59	236
9	5	26	130
10	5	48	240
11	6	21	126
12	6	39	<u>234</u>
			13421

Tabell 3 Store forretningsabonnenters behov for 30B+D pr i dag

Antall abonnement	Behov 30B+D	Abonnenter	30B+D aksesser
13	1	13	13
14	1	15	15
15	1	15	15
16	1	20	20
17	1	4	4
18	1	4	4
19	1	6	6
20	1	15	15
21-25	1	22	22
26-30	1	18	18
31-35	2	6	12
36-40	2	9	18
41-45	2	9	18
46-50	2	4	8
51-100	3	18	54
101-150	5	3	15
151-200	6	4	24
201-250	8	1	8
251-999	9	0	<u>0</u>
			289

Tabell 2 gir en oversikt over antall små forretningskunder fordelt på teleområde og summert på distriktet, og antall 2B+D aksesser disse antas å velge. Behovet for 2B+D aksesser er antatt på bakgrunn av at kundens behov for telefoni i dag dekkes opp ved overgang til ISDN.

Tabell 3 gir en oversikt over antall store forretningskunder fordelt på teleområde og summert på distriktet, og antall 30B+D aksesser disse antas å velge. Behovet for 30B+D aksesser er antatt på bakgrunn av at kundens behov for telefoni i dag dekkes opp ved overgang til ISDN.

Årsaken til intervallene på antall hovedabonnement er kort og godt for å begrense antall beregninger.

Vi kan nå regne ut fra tabellene hvilket behov det er pr i dag for ISDN-aksess 2B+D, samt gjennomsnittlig antall analoge aksesser en 2B+D erstatter fordelt på sentralnivå. Dette gjøres også for 30B+D.

Tabell 4 gir en oversikt for teleområdene over gjennomsnittlig antall analoge hovedabonnement som erstattes av henholdsvis en 30B+D aksess eller en 2B+D aksess brukt av forretningskunder.

Det gjenstår nå å finne dagens behov for 2B+D i boligsegmentet. Dette kan gjøres på følgende måte: Vi har registrert antall hovedabonnement på hvert enkelt sentralområde. Samtidig kjenner vi dagens behov for både 2B+D og 30B+D for forretningsabonnentene, gjennomsnittlig forhold mellom 2B+D forretning og analoge HA, og

30B+D forretning og analoge HA. Da sier vi at dagens behov for 2B+D til boligabonnentene er:

- status innkoplede abonnement pr i dag
- antall 2B+D forretning * konverteringsforhold
- antall 30B+D forretning * konverteringsforhold

=dagens behov for 2B+D til boligabonnentene

Disse beregningene er utført i tabell 5 og viser følgende:

- Antall hovedabonnement innkoplede
- Dagens behov for 2B+D aksesser
- Konverteringsforhold 2B+D
- Dagens behov for 30B+D aksesser
- Konverteringsforhold 30B+D
- Dagens behov for antall 2B+D til boligabonnenter.

I disse beregningene er det forutsatt at alle abonnenter/abonnement er potensielle ISDN-abonnenter, samt at alle konverterer til ISDN før eller senere. Det vil si at dagens totalpotensial er de som pr i dag er innkoplede. Når markedspotensialene blir beregnet, vil forutsetninger om abonnementsvekst bli lagt inn.

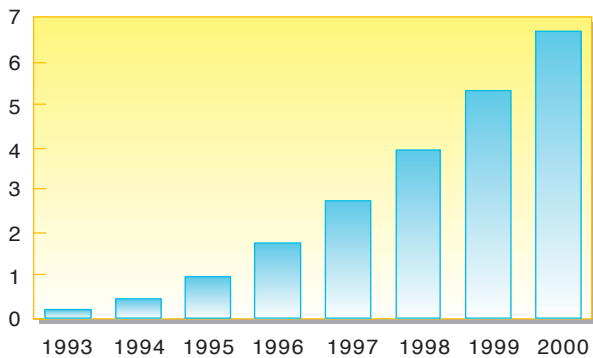
6 Prognoser fordelt på aksesstyper og markedssegmentene bolig og forretning

6.1 2B+D forretningsabonnenter

Forutsetninger:

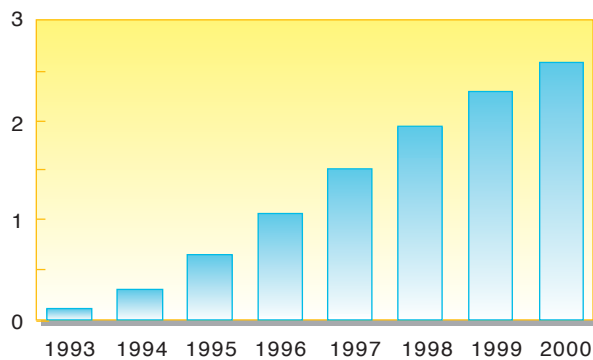
- Abonnementspris på 2B+D aksess ca 3 x abonnementspris for et hovedabonnement i dag
- 2B+D er tilgjengelig i prøvetjenesten 1993
- Anslag for abonnement i 1993 er basert på opplysninger om tilgjengelig utstyr i prøvetjenesten
- 2B+D er kommersielt tilgjengelig fra og med 1994
- 50 % av dagens behov er konvertert innen 2000
- Markedspotensialet/metningspunktet er 30 % høyere enn dagens behov.

2B+D aksesser (x1000)



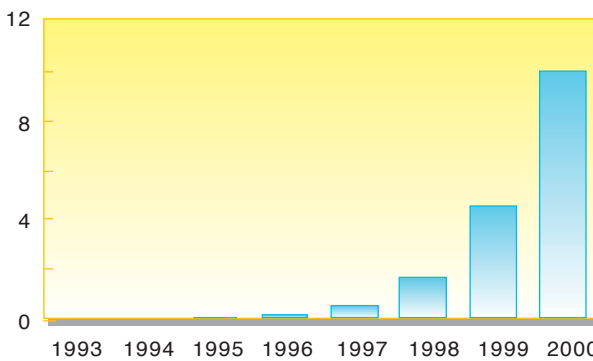
Figur 6

30B+D aksesser (x100)



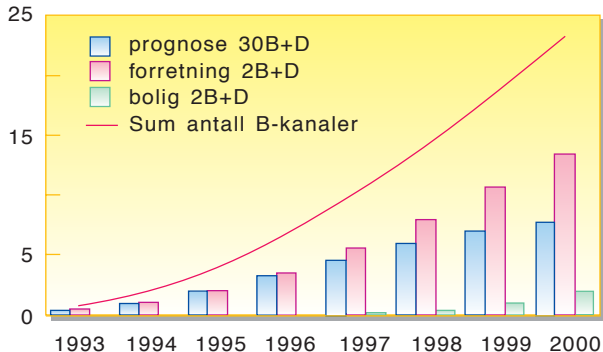
Figur 7

2B+D aksesser (x100)



Figur 8

Antall B-kanaler (x1000)



Figur 9

Tilvekst pga. datatrafikk har ingen betydning for prognosen ettersom det er medregnet en gjennomsnittlig overkapasitet, men derimot meget stor betydning for datanettene ref. konverteringsforhold, tabell 5.

Vi har nå bestemt to verdier og ett markedspotensial. I tillegg er parameteren c bestemt å være 20000 som en konstant for alle aksessformer. Vi er da i stand til å beregne utviklingen for 2B+D forretning fra og med 1993 til og med 2000. Før vi kan estimere parametrene må etterspørselsfunksjonen transformeres slik at parametrene kan bestemmes ved hjelp av regresjon.

Vi skal her tilpasse funksjonen

$$y_t = \frac{M}{(1 + e^{(a+bt)})^c}$$

hvor

M - markedspotensialet

a - parameter som skal estimeres

b - parameter som skal estimeres

c - parameter som er en konstant (= 20000)

t - tiden,

slik at det er mulig å benytte lineær regresjon. Dette gjøres ved å transformere y_t over til lineær form y'_t .

Modellen på transformert form blir følgende:

$$LN\{(M/y_t)^{1/c} - 1\} = a + bt$$

$$y'_t = a + bt;$$

hvor $y'_t = LN\{(M/y_t)^{1/c} - 1\}$

Ved å foreta regresjon på de transformerte verdiene kan parameter a og b beregnes

$$a = 430,813$$

$$b = -0,2204$$

Figur 6 viser prognosen for 2B+D forretningsabonnenter.

6.2 30B+D forretningsabonnenter

Forutsetninger:

- 30B+D er tilgjengelig i prøvetjenesten 1991
- Anslag for abonnement i 1993 er basert på opplysninger om tilgjengelig utstyr i prøvetjenesten
- 30B+D er kommersielt tilgjengelig fra og med 1994
- 90 % av dagens behov er konvertert innen 2000
- Markedspotensialet/metningspunktet er 15 % høyere enn dagens behov.

Tilvekst pga. datatrafikk har ingen betydning for prognosen ettersom det er medregnet en gjennomsnittlig overkapasitet, men derimot meget stor betydning for datanettene, ref. konverteringsforhold, tabell 5.

Beregningene foretas på samme måte som for 2B+D forretning, men med nye verdier.

Figur 7 viser prognosen for 30B+D forretningsabonnenter.

6.3 2B+D boligabonnenter

Forutsetninger:

- 2B+D er kommersielt tilgjengelig fra og med 1994. Testfasen er da over
- Første verdi som anslås er 2 abonnenter i 1995
- 4 % av dagens behov er konvertert innen 2000
- Markedspotensialet/metningspunktet er 10 % høyere enn dagens behov.

Beregningene foretas på samme måte som for 2B+D forretning, men med nye verdier.

Figur 8 viser prognosen for 2B+D boligabonnenter.

Figur 9 viser antall B-kanaler for sum bolig og forretningsabonnement.

	Opprinnelige verdier	Transformerte verdier	
Verdi 1993	y_{1993}	200	y'_{1993} -8,4063
Verdi 2000	y_{2000}	6711	y'_{2000} -9,949
Markedspotensial			13654

Tabell 4 Forhold mellom dagens behov for 2B+D og 30B+D aksesser og analoge abonnement pr i dag. Forretningsabonnenter

2B+D aksesser	13421
Analoge abonnement	16776
Konverteringsforhold	1,25
30B+D aksesser	289
Analoge abonnement	6257
Konverteringsforhold	21,65

Tabell 5 Beregning av antall boligabonnenter

Status hovedabonnement	48131
Behov 2B+D	13421
Konverteringsforhold	1,25
Behov 30B+D	289
Konverteringsforhold	21,65
Antall boligabonnement	25124

6.4 ISDN prognosene/analog telefoniprognose

På bakgrunn av de tre ovennevnte ISDN del-prognoser samt omregningsfaktorene, kan det regnes ut hvilken påvirkning ISDN har på dagens telefoni-prognoser (og hvilke konsekvenser dette får for telefonnettet). Ved å ta antall aksesser og multiplisere med konverteringsforholdet for hver sentral for hver del-prognose vil man få antall analoge ekvivalenter.

Figur 10 viser ISDN-prognosene omregnet til analoge ekvivalenter slik at de er sammenliknbare med dagens analoge telefoniprognoser.

Disse analoge ekvivalenter er nå direkte sammenliknbare med dagens telefoni-prognoser. ISDN-prognosene omregnet til analoge ekvivalenter kan videre brukes til å beregne en ny analog telefoniprognose påvirket av ISDN.

Dagens offisielle analoge telefoniprognose

- ISDN-prognosene omregnet til analoge ekvivalenter

=Dagens offisielle analoge prognose påvirket av ISDN.

Figur 11 viser dagens analoge telefoni-prognose og denne prognosen påvirket av ISDN prognosene.

6.5 Følsomhet til de enkelte prognoser i forhold til inngangsverdiene

Startverdi:

Startverdiene er basert på intelligent gjetting og kjennskap til kundene. Prognosene vil være meget følsomme overfor disse startverdiene. Det kan beskrives slik at startverdiene bestemmer kurvens forløp de første årene. Er verdien lav, blir prognosene lavere de første årene – er verdien høy, blir prognosene høyere de første årene. Feilen blir stor i prosent, men lav i antall. Økonomisk har dette liten betydning i de første årene.

Figur 12 viser hvordan prognosen varierer med forskjellige startverdier. Dette er prognosen for 2B+D forretning hvor den opprinnelige startverdien er 200.

Verdi nr 2:

Verdi nr 2 er basert på intelligent gjetting i et scenario for år 2000. Vi har her prøvd å ta hensyn til utskiftingstakt på hussentraler og apparat-systemer, avskrivningstid o.l. Prognosene er meget følsomme overfor verdi nr 2, men ettersom verdi nr 2 er bestemt til år 2000 blir konsekvensene av feil anslag her ikke vesentlig før 1995 og utover. Feil her har liten betydning for de første prognoseårene.

Figur 13 viser hvordan prognosen varierer med forskjellige verdi nr 2 (verdi år 2000). Dette er prognosen for 2B+D forretning hvor den opprinnelige verdien for år 2000 var 6711, det vil 50 % av dagens behov. Vi ser her hvordan prognosen varierer med andel av dagens behov i år 2000.

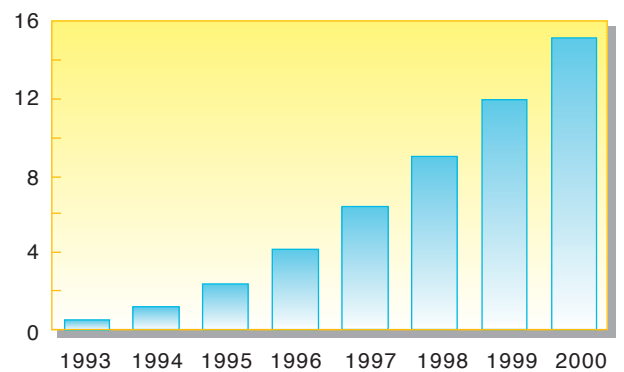
Markedspotensialet:

Denne verdien har liten innvirkning på prognosene de første årene (det vil si før vendepunktet), men jo nærmere prognosen er metningspunktet, desto mer følsom er den.

Figur 14 viser hvordan prognosen varierer med markedspotensialet. Dette er prognosen for 2B+D forretning hvor den opprinnelige verdien for markedspotensialet var 17447. Vi ser her hvordan prognosen varierer med varierende markedspotensial.

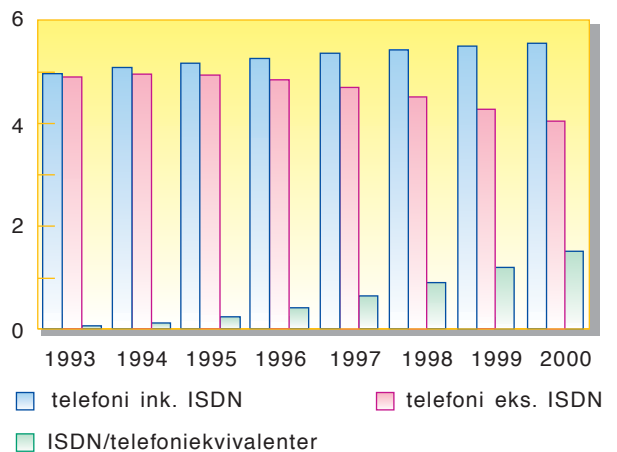
Skille mellom valg av 30B+D og 2B+D: Forutsetningen for valg mellom disse aksesser er gjort på bakgrunn av forutset-

Analoge ekvivalenter (x1000)



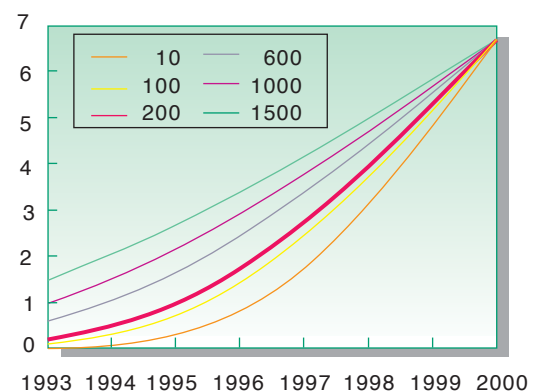
Figur 10

telefoniekvivalenter (x1000)



Figur 11

(x1000)



Figur 12

Tabell 6 Input i beregningsmodellen for forretningsabonnement 2B+D

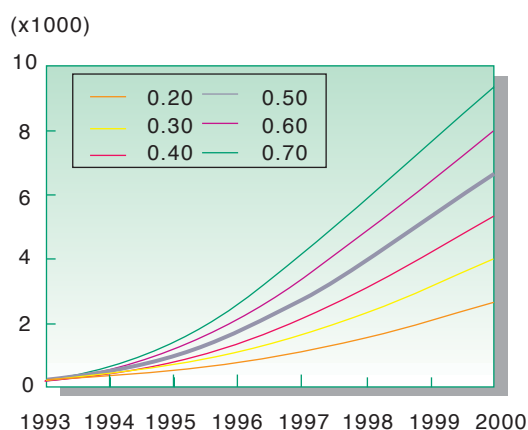
Verdi 1993	200
Verdi år 2000	6711
Markedspotensial	13654

Tabell 7 Input i beregningsmodellen for forretningsabonnement 30B+D

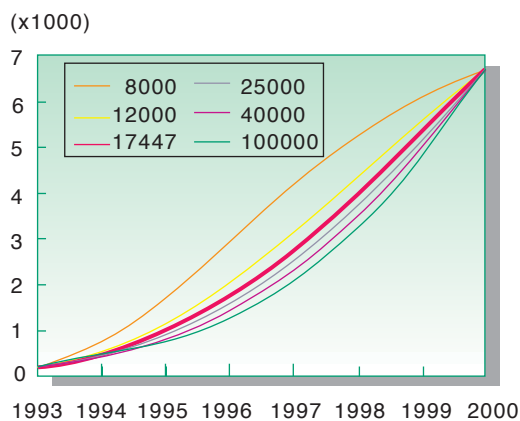
Verdi 1993	10
Verdi 2000	266
Markedspotensial	340

Tabell 8 Input i beregningsmodellen for boligabonnement 2B+D

Verdi 1995	2
Verdi år 2000	1154
Markedspotensial	25361



Figur 13



Figur 14

ninger om pris som er nevnt tidligere. Som man ser av beregningene av markedspotensialet har det innbyrdes prisforhold mellom 30B+D og 2B+D meget stor betydning. Ved endring av prisforholdet mellom 2B+D og 30B+D endres grensen for valg mellom 2B+D og 30B+D, som igjen endrer markedspotensialene for de respektive aksessformer dramatisk i antall aksesser, og konsekvensene for de respektive prognosene blir store. Prognosen for 2B+D for boligsegmentet er uavhengig av denne grensen, fordi antall boligabonnenter hele tiden er konstant.

Derimot vil allikevel prognosen for analoge ekvivalenter være relativt uberørt, fordi endring i markedspotensialet for 2B+D for forretning øker nesten tilsvarende som markedspotensialet for 30B+D synker. Og derfor er med andre ord denne prognosen (analoge ekvivalenter) stort sett upåvirket av det innbyrdes prisforhold mellom ISDN-aksessene.

7 Konklusjon

De grafiske framstillingene viser, under nevnte forutsetninger, at forretningsstanden vil være den viktigste og største kunden for ISDN langt ut i neste århundre. Det viser seg også at den største pågangen etter 30B+D er over før 2000. Men 2B+D for forretning får den største veksten rundt år 2000. Boligabonnentene vil først virkelig etterspørre ISDN etter år 2000. Enten fordi de har et behov eller fordi de ikke har noe valg. Kort sagt er forretningsabonnentene de som vil etterspørre ISDN i størst grad, det er for øvrig de som vil ha det største økonomiske utbyttet av dette.

Generelt er det en viss usikkerhet omkring forutsetningene for prognosene for ISDN. Dette gjelder spesielt ikke avklarte spørsmål, som

- Hvilke tjenester vil være pådrivere for ISDN?
- Hva vil de endelige takstene (tilkopling, fastavgifter og trafikkavgifter) bli?
- Hvor mye vil det koste å konvertere eksisterende utstyr til ISDN grensesnitt (f eks konvertere fra analoge tilknytninger til 30B+D)?
- Hvordan vil tilgjengeligheten til ISDN være i framtiden?
- Hvordan vil tilgjengeligheten av utstyr være i framtiden?
- Hvilket incitament har utstyrsleverandørene for å anbefale en ISDN-løsning framfor vanlig telefoniløsning?

Ser vi bort fra den ovennevnte usikkerhet og antar at alle andre forutsetninger holder, kan vi noe upresist si at prognosene kun avhenger av kvaliteten på grunnlagsdata fra TELSIS. Anslagene på verdi nummer to er imidlertid vage og representerer dermed en betydelig usikkerhet. Inntil vi får bedre erfaringsmateriale er dette likevel det beste vi kan produsere med de opplysningene som er tilgjengelig i dag. Prognosene gir oss således et godt bilde på hvordan utviklingen kan bli med de siste tidsplaner og de alternative prisforhold.

Tabell 10 ISDN prognosene fordelt mellom aksessform og marked

	Omregningsfaktor	Markedspotensial	1993	1994	1995	1996	1997	1998	1999	2000
30B+D	21,65	332	10	30	65	108	154	197	232	260
2B+D forretning	1,25	17447	200	484	983	1737	2741	3953	5302	6711
2B+D bolig		27637	0	0	2	11	51	171	456	1005

Tabell 11 Telefoniprognose inklusiv ISDN, ISDN-prognosene og telefoniprognosen eksklusiv ISDN

	1993	1994	1995	1996	1997	1998	1999	2000
Telefoni inkl ISDN	68331	70015	71493	72841	74067	75220	76341	77405
ISDN	560	1199	2367	4154	6448	9120	12042	15154
Telefoni ekskl ISDN	48863	49375	49217	48351	46896	45012	42857	40473

Prognoser for pris på nettkomponenter basert på utvidelse av lærekurvemodellen

AV BORGAR TØRRE OLSEN OG KJELL STORDAHL

001.18:621.39

1 Sammendrag

Denne artikkelen gir bakgrunn for utvikling av de såkalte lærekurvemodeller som brukes for å anslå fremtidige (prognostiserte) kostnader på produksjon av utstyr som funksjon av produksjonsvolum. Det forutsettes at metoden også kan brukes på pris på utstyr som funksjon av produksjonsvolum.

Ved innføring av Logistisk modell for etterspørsel etter utstyr og sette denne sammen med Wrights modell for lærekurver, er det i denne artikkelen vist hvorledes det er mulig å lage modeller for prognostisert pris på utstyr som funksjon av tiden.

Metoden er benyttet i tilknytning til RACE-prosjektet TITAN som er et prosjekt som utvikler et planverktøy for utbygging av abonnentnett basert på nye nettkomponenter. Prisutviklingen på disse nettkomponentene sammen med etterspørsel for høykapasitetstjenester er avgjørende for hvilke nettarkitekturer og løsninger som vil bli anbefalt.

Artikkelen viser hvorledes parametrene i modellene estimeres (beregnes) på optimal måte ved bruk av statistiske metoder. I tillegg vises det hvorledes parametrene anslås når det er mangel på observasjoner.

2 Lærekurvemodeller

Lærekurvemodeller brukes for å anslå i hvilken grad kostnad pr utstyrsenhet avtar etter hvert som produksjonsvolumet øker. I litteraturen brukes også lærekurvemodeller om individuell læring. I en industriprosess er imidlertid ikke individuell læring det primære. Her vil det være en sammensetning av en rekke faktorer som kan være avgjørende for kostnadsnedgang pr produsert enhet etter hvert som produksjonsvolumet øker. Faktorer som har betydning er:

- Effektivitetsforbedring av arbeidsstokk
- Organisasjonsmessige endringer
- Bedre styring av produksjonsprosessen
- Innføring av nye metoder
- Innføring av ny teknologi
- Raffinering av produktet (redesign) for å redusere kostnader
- Standardisering av produksjonsprosessen
- Automatisering.

Mange av disse faktorene griper inn i hverandre og er med andre ord ikke uavhengige. Eksempelvis kan en programvare-utvikling som understøtter utviklingen av produktet, være sammensatt av bedre styring av produksjonsprosessen, innføring av nye metoder, samt standardisering av produksjonsprosessen.

Wright introduserte i 1936 [1] den klassiske lærekurvemodellen. Den er gitt ved:

$$P = P_0 n^\beta \quad (2.1)$$

der

P er kostnad pr produsert enhet

P_0 er kostnad ved produksjon av første enhet

n er totalt antall produserte enheter

β er parameter i modellen ($\beta < 0$).

Det ses av formelen at kostnadene reduseres eksponentielt omvendt proporsjonalt med produksjonsvolumet og basert på en gitt parameter. Parameteren er selv sagt forskjellig for ulike typer utstyr eller produkter. Den kan estimeres ved bruk av regresjonsanalyse og minste kvadraters metode. Ut fra tidligere erfaringer og konkrete data (observasjoner), finnes da den optimale verdi på parameteren.

Ofte benyttes lærekurveraten K . Den angir hvor mye kostnaden på produktet reduseres med når produksjonsvolumet doubles. Det kan lett vises til at sammenhengen mellom β og K er gitt ved:

$$\beta = \frac{\ln K}{\ln 2} \quad (2.2)$$

Lærekurveraten K er et tall mindre enn 1. Dersom det multipliseres med 100, kan det brukes som prosentmål. Typiske verdier for lærekurveraten er fra 70 % til 95 %. Det avhenger i høy grad av hvilke produktgrupper det er snakk om. Små prosentverdier indikerer en relativt brattere lærekurve. Med andre ord vil det skje større kostnadsreduksjoner som funksjon av økende produksjonsvolum når K har en lav prosentverdi i forhold til en høy prosentverdi.

Det er to svakheter ved Wrights lærekurvemodell. Modellen har bare en parameter og er derfor ikke så fleksibel som ønskelig med hensyn til tilpasning av historiske observasjoner og med hensyn til å beskrive kostnadsutvikling for ulike typer produkter. Til gjengjeld er det lett å operere med en K -verdi som sammenlikner kostnadsutviklingen for ulike typer produkter. Den andre svakheten er at Wrights

modell baserer seg på akkumulert produksjonsvolum. Dette skaper en avhengighet i observasjonene som benyttes til estimering av parameteren. I statistisk estimering er det ønskelig med en uavhengighet i observasjonene, men dette er en forutsetning som mange ganger kan være vanskelig.

Crawford [2] introduserte en variant av Wrights lærekurvemodell. I Wrights modell ble det forutsatt at kostnaden pr produsert enhet avtar med akkumulert antall produserte enheter. Med akkumulert antall produserte enheter menes totalt antall enheter som er produsert fra produksjonsstart som kan være for mange år siden. I Crawfords modell opereres det med antall enheter produsert i en serie. En serie kan defineres på ulike måter. Seriene kan være bestemt av den tid en opprettholder samme produksjonsteknikk eller de kan være bestemt ut fra en gitt tidslengde. Et naturlig tidsintervall kan være et år. Dette er ikke noe dårlig utgangspunkt fordi regnskaper i alle fall lagres på årsbasis.

Crawfords lærekurvemodell blir matematisk identisk med Wrights modell angitt i likning (2.1). De variable er imidlertid definert på følgende måte:

P er kostnad pr produsert enhet i en gitt serie

P_0 er kostnad pr produsert enhet i 1-te serie

n er antall enheter

β er parameter i modellen.

Også i Crawfords modell holdes kostnad pr produsert enhet i 1-te serie som en fast størrelse og ikke som en parameter som skal estimeres, noe som gjør modellen mindre fleksibel.

Stanford Research Institute har introdusert en lærekurvemodell med to parametere. Den er gitt ved:

$$P = P_0(n + D)^\beta \quad (2.3)$$

Her er D den ekstra parameteren. For øvrig er variablene definert som i Wrights modell.

de Jong [3] har utviklet en annen variant av lærekurvemodellen til Wright. Han antar at en andel M av produksjonskostnadene er konstant og ikke påvirkes av masseproduksjon. Det gjør imidlertid den øvrige andel $(1 - M)$ av produksjonskostnadene, de Jongs modell blir da som følger:

$$P = P_0(M + (1 - M)n^\beta) \quad (2.4)$$

Parameteren M har verdier mellom 0 og 1. Parametrene M og β estimeres ut fra gitt datamateriale.

Det kan selvsagt settes opp en rekke ulike modeller for kostnadsutvikling for en enhet i en produksjonsprosess [4]. Dette kan eksempelvis også gjøres på mikronivå gitt at det eksisterer statistisk underlag. Vi er imidlertid mer interessert i kostnads- og spesielt prisutviklingen på nye nettkomponenter der produksjonsprosessen så vidt har kommet i gang. Det er derfor naturlig å ta utgangspunkt i makromodeller.

3 Modell for prognoser på pris på nettkomponenter som funksjon av produksjonsvolum

I kapittel 2 ble det satt opp modeller for kostnader på produsert utstyr. I en normal situasjon er det grunnlag for å kunne anta at det er en lineær sammenheng mellom pris og kostnad på produksjon av utstyr – eksempelvis nettkomponenter. Det må regnes med et prosentvis påslag som skal dekke andre kostnader som markedsførings-, utviklings- og overheadkostnader, og i tillegg en viss for-tjeneste.

Ut fra disse forutsetningene vil modellene i kapittel 2 også gjelde for prisutvikling på utstyr/nettkomponenter, men selvsagt med andre verdier på parametrene. Unntak vil være når produksjonsbedrifter nedfaser et utstyr fordi de ønsker å introdusere nye generasjoner av utstyret eventuelt en helt ny type utstyr på markedet. Da vil det komme inn element av taktisk prising av det gamle utstyret blant annet for å få avsetning på det nye utstyret.

Figur 3.1 viser Wrights lærekurvmodell over en lengre produksjonsperiode.

I Wrights lærekurvmodell vil kostnad (og pris) hele tiden avta omvendt proporsjonalt med akkumulert antall produserte enheter. Selv om det blir små produksjonsserier og taktisk prisnivå når et utstyr utfases, gir Wrights lærekurvmodell stadig lavere enhetskostnader og priser.

I Crawford's lærekurvmodell der det tas utgangspunkt i størrelsen på produksjonsseriene, vil kostnad og pris igjen øke ved utfasing av utstyret.

Dersom vi imidlertid er i en normalsituasjon vil både Wrights og Crawford's lære-

kurvmodeller gi forholdsvis like resultater.

Svakheten til metodene er imidlertid at de er lite fleksible, da de kun har en parameter som skal estimeres. Dette kan forholdsvis enkelt endres på ved ikke å binde seg til at lærekurven skal gå gjennom det første punktet på kurven som er produksjonskostnad (eller pris) for det første produktet. Det er også spesielt betenkelig å la lærekurven gå gjennom det første punktet og ikke de andre punktene for beregnede enhetskostnader (enhetspriser). Det første punktet er relatert til produksjon av det første produktet. Her vil det være stort slingsrom med hensyn til kostnad/pris, hvor mange tilfeldigheter vil være avgjørende.

Det vil med andre ord være atskillig bedre å bestemme at lærekurven skal gå gjennom et punkt som representerer kostnad/pris pr enhet etter at produksjonsprosessen har stabilisert seg.

Den beste løsningen og statistisk mest riktige vil være å innføre en ekstra parameter slik at en ikke bindes opp til å la lærekurven gå igjennom noen spesifiserte punkter. Verdien på denne parameteren bestemmes på optimal måte ved bruk av regresjonsanalyse. Se også [5].

En raffinering av Wrights lærekurvmodell blir da:

$$P_n = \alpha n^\beta \quad (3.1)$$

der

P_n er pris pr produsert enhet

n er totalt antall produserte enheter

α, β er parametere i modellen.

Crawford's lærekurvmodell raffineres på samme måte, men med definisjon av n som antall produserte enheter i en serie.

Anta nå at vi har observert prisen P_n med tilhørende totalt antall produserte enheter s ganger. Vi har da følgende observasjoner:

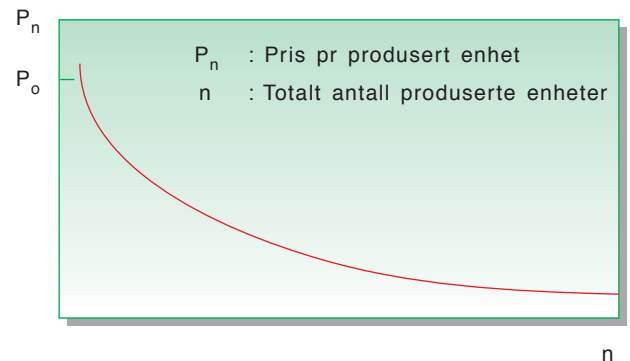
$$(n_1, P_{n_1}), (n_2, P_{n_2}), \dots, (n_s, P_{n_s}) \quad (3.2)$$

der

n_t er akkumulert antall enheter produsert ved tidspunkt $t = 1, 2, \dots, s$

P_{n_t} er enhetspris ved tidspunkt t basert på produksjon av totalt n_t enheter $t = 1, 2, \dots, s$.

Parametrene i modell (3.1) estimeres da på følgende måte:



Figur 3.1 Wrights lærekurvmodell

$$P_{n_t} = \ln \alpha + \beta \ln n_t \quad t = 1, 2, \dots, s \quad (3.3)$$

Likning (3.3) kan transformeres over til den tradisjonelle regresjonsmodell

$$y_t = \alpha^* + \beta^* Z_t \quad t = 1, 2, \dots, s \quad (3.4)$$

ved

$$\alpha^* = \ln \alpha \quad (3.5)$$

$$\beta^* = \beta \quad (3.6)$$

$$y_t = \ln P_{n_t} \quad (3.7)$$

$$Z_t = \ln n_t \quad (3.8)$$

Optimale verdier på parametrene α^* og β^* finnes ved bruk av minste kvadraters metode. Dette gir:

$$\alpha^* = \frac{\left(\sum_{t=1}^s y_t \right) \left(\sum_{t=1}^s Z_t^2 \right) - \left(\sum_{t=1}^s Z_t \right) \left(\sum_{t=1}^s Z_t y_t \right)}{s \left(\sum_{t=1}^s Z_t^2 \right) - \left(\sum_{t=1}^s Z_t \right)^2} \quad (3.9)$$

$$\beta^* = \frac{\left(\sum_{t=1}^s Z_t y_t \right) - \left(\sum_{t=1}^s Z_t \right) \left(\sum_{t=1}^s y_t \right)}{s \left(\sum_{t=1}^s Z_t^2 \right) - \left(\sum_{t=1}^s Z_t \right)^2} \quad (3.10)$$

Dermed er

$$\hat{\beta} = \hat{\beta}^* \quad (3.11)$$

$$\hat{\alpha} = e^{\hat{\alpha}^*} \quad (3.12)$$

Under forutsetning av at observasjonene er uavhengige, vil beregnede standardavvik til de estimerte parametrene $\hat{\alpha}^*$ og $\hat{\beta}^*$ være:

$$\hat{\sigma}_{\alpha^*} = \sqrt{\frac{\hat{\sigma}^2}{s} + \frac{\left(\sum_{t=1}^s Z_t \right)^2}{s^2} \hat{\sigma}_{\beta^*}^2} \quad (3.13)$$

$$\hat{\sigma}_{\beta^*} = \sqrt{\frac{s}{s \left(\sum_{t=1}^s Z_t^2 \right) - \left(\sum_{t=1}^s Z_t \right)^2}} \hat{\sigma} \quad (3.14)$$

der

$$\hat{\sigma} = \sqrt{\frac{1}{s-2} \sum_{t=1}^s (y_t - \hat{y}_t)^2} \quad (3.15)$$

$$\hat{y}_t = \hat{\alpha} + \hat{\beta} Z_t \quad (3.16)$$

De beregnede standardavvik for $\hat{\alpha}$ og

$\hat{\beta}$ blir da:

$$\hat{\sigma}_{\alpha} = \hat{\sigma}_{\alpha^*} \quad (3.17)$$

$$\hat{\sigma}_{\beta} = e^{\hat{\sigma}_{\beta^*}} \quad (3.18)$$

Ved bruk av utvidelsen av Wrights lærekurvmodell med to parametere vil de angitte formler for standardavvik være nedre grense fordi observasjonene ikke er uavhengige, da de er akkumulerte verdier.

Dersom vi imidlertid bruker Crawford's lærekurvmodell med den utvidelsen som er foreslått med to parametere, vil de observasjonene som ligger til grunn være uavhengige, og de estimerte standardavvik kan da brukes direkte.

Dersom det eksisterer rimelig mange observasjoner, vil et intervall på to ganger standardavviket til hver side for estimert parameterverdi angi et tilnærmet 95 % konfidensintervall. Dette intervallet angir da usikkerheten i de estimerte parameterverdien. Ved få observasjoner benyttes 97,5 % fraktilen ($t_{s-2, 0,975}$) istedenfor delingen med $s-2$ frihetsgrader som en multiplikator til estimert standardavvik. 95 % konfidensintervall blir da:

$$I_{\alpha, 0,95} = \left[\hat{\alpha} - t_{s-2, 0,975} \hat{\sigma}_{\alpha}, \hat{\alpha} + t_{s-2, 0,975} \hat{\sigma}_{\alpha} \right] \quad (3.19)$$

$$I_{\beta, 0,95} = \left[\hat{\beta} - t_{s-2, 0,975} \hat{\sigma}_{\beta}, \hat{\beta} + t_{s-2, 0,975} \hat{\sigma}_{\beta} \right] \quad (3.20)$$

Etter hvert som observasjonene i (3.2) fremkommer, kan lærekurvene bestemmes ved estimering av α og β . Deretter kan standardavvik og konfidensintervall beregnes. Dette gir oss mulighet til å se hvor signifikante verdiene på $\hat{\alpha}$ og $\hat{\beta}$ er.

4 Modell for prognoser på pris på nettkomponenter som funksjon av tiden

Hittil har vi sett på utviklingen av pris på en nettkomponent som funksjon av produksjonsvolumet. Det som er enda mer interessant, er å finne en modell for denne prisen som funksjon av tiden.

For å kunne bruke de utvidede Wrights og Crawford's lærekurvmodeller med to parametere er det nødvendig å føre inn en prognosemodell. La totalt (akkumulert) antall produserte enheter ved tidspunkt t være gitt ved $n(t)$. De observasjoner som er kjent, er:

$$n(1), n(2), \dots, n(s) \quad (4.1)$$

Ut fra observasjonene kan det nå lages en prognose for akkumulert antall produserte enheter. Hvis s er forholdsvis stor, kan vi bruke tradisjonelle prognosemodeller til prognostiseringen. Aktuelle prognosemodeller er:

- Enkel regresjonsmodell
- Multipel regresjonsmodell
- Holt glattingsmodell
- Holt-Winters glattingsmodell
- Tidsrekkemodeller med eller uten sesongkomponenter
- Transfermodeller
- Kalmanfiltermodeller
- Metningsmodeller.

Prognoser for akkumulert produksjonsvolum ved tidspunkt T ved bruk av den prognosemodell som passer best til dataunderlaget betegnes med

$$\hat{n}(T) \quad (4.2)$$

Ved å sette denne størrelsen inn i likning (3.1) sammen med de estimerte parametrene fra likning (3.11) og (3.12) fås:

$$\hat{P}_{(n)T} = \hat{\alpha} \hat{n}(T)^{\hat{\beta}} \quad (4.3)$$

Her er da $\hat{P}_{(n)T}$ prognostisert pris på den aktuelle nettkomponenten ved tidspunkt T .

På analog måte finnes prisprognoser for den aktuelle nettkomponent ved bruk av Crawford's utvidete modell.

5 Prisprognoser på nettkomponenter når dataunderlaget er begrenset

I de tilfeller hvor nye nettkomponenter introduseres, vil det være et spinkelt dataunderlag. Dersom det kun er en produksjonsserie og vi har en observasjon, er det ikke mulig å benytte utvidelse av Wrights og Crawford's lærekurvmodeller, som jo er basert på to parametere. Men med en gang vi har to eller flere observasjoner kan dette gjøres.

Anta at vi er i introduksjonsfasen av en ny nettkomponent og at vi har noen få observasjoner. Da vil vi kunne benytte de modeller som er utviklet i kapittel 3 til estimering av pris som funksjon av produksjonsvolum.

Når det gjelder prognose for produksjonsvolum, vil det være naturlig å ta utgangspunkt i en modell i klassen av metningsmodeller. Den mest brukte er den Logistiske modellen.

Den Logistiske modellen brukes gjerne for å beskrive startfasen av et produkts utvikling eller slutfasen når produktet nærmer seg metning. Den Logistiske modellen er gitt ved:

$$n(t) = M(1 + e^{a+bt})^{-g} \quad (5.1)$$

der

$n(t)$ er akkumulert antall enheter ved tidspunkt t

M er markedspotensialet

a, b, g er parametere i modellen.

Utgangspunktet er nå at s observasjoner, se 4.1, er kjente. Ut fra dette skal de fire parametrene M, a, b og g estimeres. Dersom s er liten, kan det være nødvendig å fastsette verdier på M og g . g kan ut fra erfaring [6] velges som et høyt tall, mens M må anslås lik et forventet markedspotensial.

Ved å foreta en lineær transformasjon kan a og b estimeres ved minste kvadraters metode.

Det er imidlertid ikke mulig å benytte minste kvadraters metode til å estimere alle fire parametrene simultant. I [6] er det imidlertid vist en rekursiv iterasjonsprosedyre som kan benyttes til å estimere alle fire parametrene i modellen (5.1).

Dermed vil:

$$\hat{n}(T) = \hat{M} \left(1 + e^{\hat{a} + \hat{b}T} \right)^{-\hat{g}} \quad (5.2)$$

angi prognoser for akkumulert produksjonsvolum ved tidspunkt T . Dette uttrykket kan så settes inn i Wrights utvidet lærekurvemodell, og vi får:

$$\hat{P}_{n(T)} = \hat{\alpha} \left[\hat{M} \left(1 + e^{\hat{a} + \hat{b}T} \right)^{-\hat{g}} \right]^{\hat{\beta}} \quad (5.3)$$

Dette uttrykket angir da prisprognose for nettkomponenter gitt at n er akkumulert produksjonsvolum.

Hvis vi isteden benytter Crawfords utvidede lærekurvemodell der \hat{n}_t^* er definert som antall enheter produsert i en serie ved tidspunkt t , vil vi få andre formler for prisprognosen.

La igjen

$$n(1), n(2), \dots, n(s)$$

være det akkumulerte produksjonsvolum. Da vil:

$$n^*(t) = n(t) - n(t-1) \quad (5.4)$$

Med andre ord, differansen mellom akkumulert produksjonsvolum ved tidspunkt t og $t-1$, angir hvor stort produksjonsvolumet er ved tidspunkt t . Av likning (5.4) og likning (5.2) får vi følgende:

$$\hat{n}(T)^* =$$

$$\hat{M} \left(1 + e^{\hat{a} + \hat{b}T} \right)^{-\hat{g}} - \hat{M} \left(1 + e^{\hat{a} + \hat{b}(T-1)} \right)^{-\hat{g}} \quad (5.5)$$

Dette uttrykket settes så inn i Crawfords utvidede lærekurvemodell og vi får:

$$\hat{P}_{n(T)} = \hat{\alpha} \left[\hat{M} \left(1 + e^{\hat{a} + \hat{b}T} \right)^{-\hat{g}} - \hat{M} \left(1 + e^{\hat{a} + \hat{b}(T-1)} \right)^{-\hat{g}} \right]^{\hat{\beta}} \quad (5.6)$$

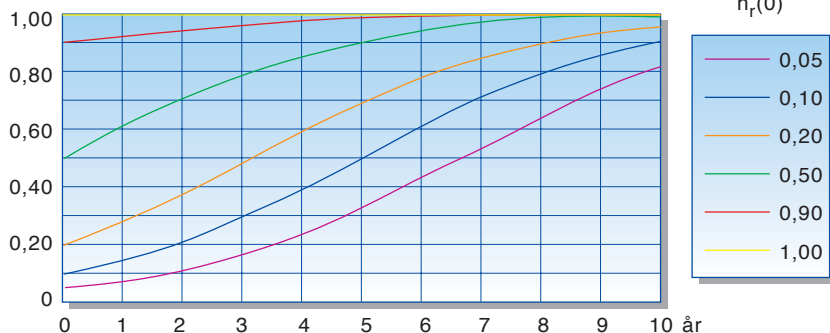
Dette uttrykket angir da prisprognose for nettkomponenter gitt at $n^*(t)$ er produksjonsvolum pr tidsintervall.

6 Praktisk fortolkning av prognosemodellen

Vi skal nå vise hvorledes prognosemodellen kan benyttes for å beskrive pris på komponenter som funksjon av tiden og se på betydningen av de parametrene som inngår.

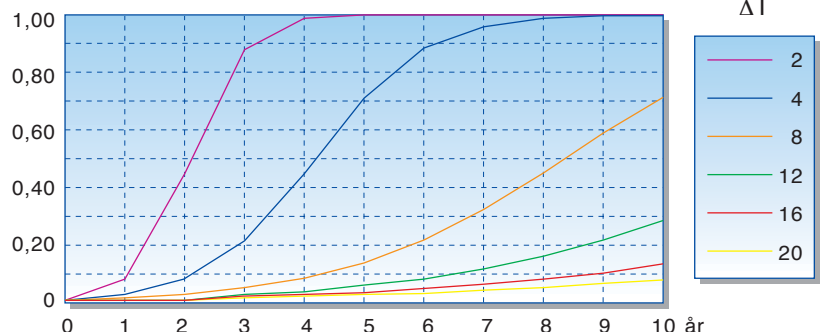
Spesielt når det dreier seg om nye komponenter som bare foreligger i form av prototyper eller hvor produksjonen er helt i startfasen er det viktig å ha en modell som en kan analysere og benytte for å gjøre gode antakelser om prisutvik-

Relativt akkumulert volum



Figur 6.1 Relativt akkumulert volum som funksjon av tiden for en konstant $\Delta T = 10$ år og med $n_r(0)$ varierende fra 0,05 til 1

Relativt akkumulert volum



Figur 6.2 Økning av relativt akkumulert volum som funksjon av tiden når $n_r(0) = 0,01$ holdes konstant og ΔT varierer fra 2 til 20 år

lingen. Det er derfor viktig å forstå hva parametrene betyr i modellen.

En viktig konsekvens som vi kan trekke av å kombinere en logistisk modell som i likning (5.1), (hvor vi for enkelhets skyld setter $g = 1$ som konstant), med en lærekurve som i likning (3.1)

$$P_{n(t)} = \alpha n(t)^\beta \quad (6.1)$$

er at det ikke vil være nødvendig å beskrive $n(t)$ som et globalt akkumulert volum, men vi kan nøye oss med den relative prognosen

$$n_r(t) = \left(1 + e^{a+bt} \right)^{-1} \quad (6.2)$$

hvor $n_r(t)$ går fra 0 til 1.

Dette betyr at vi f.eks kan benytte volumprognoser i et enkelt land eller i noen storbyer hvor anvendelsen av komponenten kan antas å være representativ for den globale utviklingen. Det betyr at en ikke trenger informasjonen om det totale globale produksjonsvolum (eller akkumulert globalt volum) som det kan være vanskelig å få data på.

Den relative volumfunksjonen $n_r(t)$ kan skrives (se appendiks):

$$n_r(t) = \left[1 + e^{\left\{ \ln(n_r(0)^{-1}-1) - \frac{2 \ln 9}{\Delta T} t \right\}} \right]^{-1} \quad (6.3)$$

hvor $n_r(0)$ er funksjonsverdien når $t = 0$ og ΔT er den tiden det tar fra funksjonsverdien $n_r(t_1) = 10\%$ til $n_r(t_2) = 90\%$, dvs $\Delta T = t_2 - t_1$.

Figur 6.1 viser nå denne relasjonen for det tilfellet at $\Delta T = 10$ år holdes konstant og $n_r(0)$ varierer fra 5 % til 100 %.

Figur 6.2 viser hvorledes $n_r(t)$ utvikler seg med $n_r(0) = 0,01$ og med ΔT varierende fra 2 til 20 år.

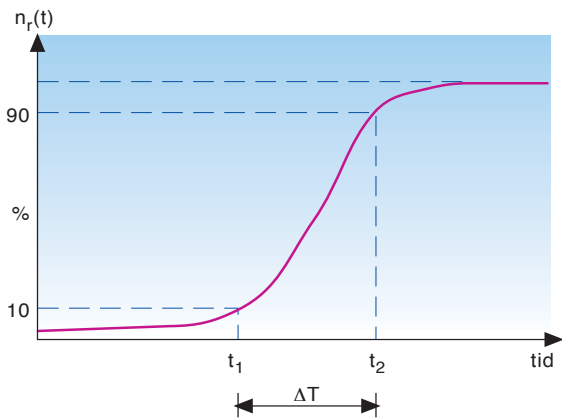
Parametrene a og b er nå uttrykt ved $n_r(0)$ og ΔT der

$$a = \ln(n_r(0)^{-1} - 1) \quad (6.4)$$

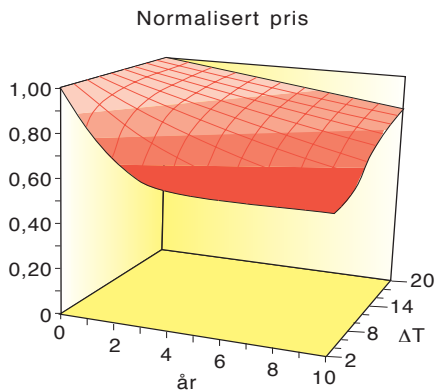
og

$$b = -\frac{2 \ln 9}{\Delta T} \approx -\frac{4,4}{\Delta T} \quad (6.5)$$

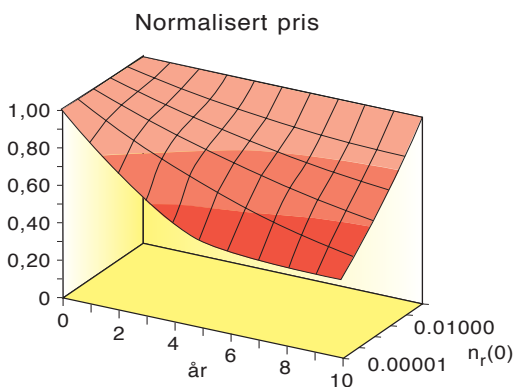
Vi ser at $n_r(0)$ (i likhet med a) bestemmer hvor på tidsaksen kurven skal ligge, mens ΔT (eller b) forteller om kurven dekker et langt eller et kort tidsforløp (dvs fra funksjonsverdien er 10 % til den



Figur 6.3 Illustrasjon av de størrelsene som inngår i likning (6.3)



Figur 6.4 Utvikling av normalisert komponentpris i tid for $n_r(0) = 0,001$ og $K = 0,95$



Figur 6.5 Normalisert pris som funksjon der tiden for forskjellige verdier av $n_r(0)$ og med $\Delta T = 10$ år, $K = 0,95$ som konstanter

er 90 % av maksimalverdien). Figur 6.3 demonstrerer denne relasjonen.

Vi vil nå vise at vi bare trenger den relative prognosefunksjonen i prisfunksjonen. La oss starte med den absolutte verdien gitt i likning (5.1) for akkumulert volum $n(t)$. Prisen pr enhet kan da skrives:

$$P(t) = \alpha n(t)^\beta \quad (6.6)$$

hvor $P(t)$ betyr $P_{n(t)}$. La oss nå sette inn $t = 0$ og vi får prisen pr enhet i referanseåret:

$$P(0) = \alpha n(0)^\beta \quad (6.7)$$

Ved å dele uttrykket (6.6) på (6.7) får vi

$$\frac{P(t)}{P(0)} = \left[\frac{n(t)}{n(0)} \right]^\beta \quad (6.8)$$

eller

$$P(t) = P(0) \left[\frac{n_r(t)}{n_r(0)} \right]^\beta \quad (6.9)$$

der de absolutte funksjonsverdiene $n(t)$ erstattes av de relative verdiene $n_r(t)$ siden

$$\frac{n(t)}{n(0)} = \frac{n_r(t)}{n_r(0)} \quad (6.10)$$

Ved å sette inn likning (6.3) i (6.9) blir det fullstendige uttrykket for pris pr enhet som funksjon av tid.

$$P(t) = P(0) \left(n_r(0)^{-1} \left[1 + e^{\left\{ \ln(n_r(0)^{-1} - 1) - \frac{2 \ln 9}{\Delta T} t \right\}^{-1}} \right]^{\log_2 K} \right)^\beta \quad (6.11)$$

som inneholder de fire størrelsene $P(0)$, $n_r(0)$, ΔT og K . Det er egentlig bare $n_r(0)$ og ΔT som må prognostiseres. Prisen i referanseåret $P(0)$ og lærekurvekoeffisientene K er mer eller mindre kjent. For å illustrere denne prisfunksjonen setter vi $P(0) = 1$ (normalisert pris) og $K = 0,95$ konstant. Figur 6.4 viser hvorledes normalisert pris varierer ved å holde $n_r(t) = 0,001$ og la ΔT variere fra 2 til 20 år.

Figur 6.5 viser likeledes hvorledes normalisert pris utvikler seg for forskjellige $n_r(0)$ ved å holde ΔT og $K = 0,95$ konstant.

7 Anvendelse av den utvidede prognosemodellen i teknisk økonomisk analyse

Det europeiske RACE-prosjektet TITAN, som står for "Tool for Introduction scenario & Techno-economic evaluation of Access Network", har valgt å benytte den utvidede prognosemodellen som grunnlag for sine beregninger [7], [8], [9]. Dette prosjektet har som målsetting å evaluere forskjellige introduksjonsscenarioer for fiber i aksessnett. Prosjektet har utviklet et PC-basert pro-

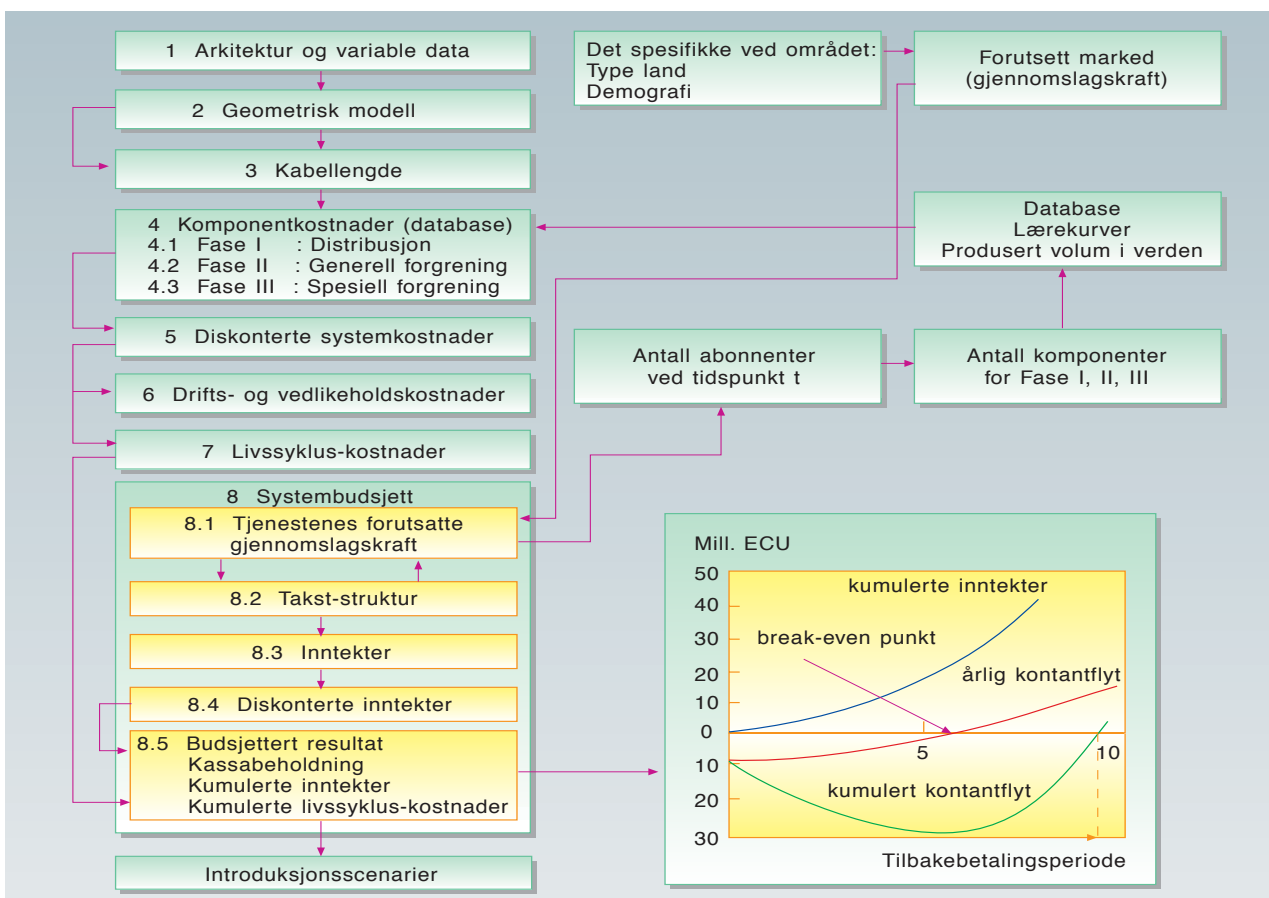
gram for beregning av lønnsomheten av de forskjellige scenarier som man ønsker å studere. Fokus er på en studieperiode av 10 år (f eks 1995–2005) og tar for seg privatmarkedet og de små bedriftskunder. Det fokuseres på aksessnett som er den mest følsomme delen av hele nettet med hensyn til kostnadene, og som dessuten er den delen av nettet som er mest avhengig av hvorledes behovet/ etterspørselen etter tjenester blir i fremtiden. Forskjellige systemalternativer studeres og beregningen som gjøres legges til grunn for planleggingen av fremtidig utbygging, med utgangspunkt i den teknologien som finnes i dag.

Prosjektet kan deles i tre deler:

- 1 Utviklingen av et fleksibelt regneprogram med en (eller flere) geometriske modeller som gir et mål på materiell, installerings- og gravekostnader, kostnader for utstyr og vedlikehold. Det tas hensyn til avskrivninger og andre økonomiske faktorer. Alle beregninger relateres til nåverdier. Det legges inn prognoser for tjenesteutvikling og inn-tjeningsmuligheter og det gjøres følsomhetsanalyser der effekten av mer eller mindre ekstreme avvik fra prognosene studeres.
 - 2 Prosjektet studerer prognoser for de tjenestene som man tror vil være de viktigste drivkreftene i privatmarkedet. Man deler tjenestene inn i båndbredde-kategorier og søker å prognostisere potensielt behov avhengig av hva tjenestene vil koste. Videre transformeres disse prognosene til behov som funksjon av tid ved hjelp av de pris-/kostnadsmodellene som er utviklet.
- Som en del av dataunderlaget for tjenestebehovsprognoser blir det gjennomført en "Delphi-studie" med ca 100 eksperter involvert. Resultatene korreleres med andre metoder for prognoser.

- 3 Den siste delen av TITAN består i å beskrive de aktuelle arkitekturer som skal studeres. Det er for eksempel FTTH (fiber til hjemmene), FTTC (fiber til grøftkantene) med kopper eller radio drop. Disse løsningene sammenliknes med tradisjonelle kopperløsninger. Det er også av interesse å studere forskjellige oppgraderinger av bestående arkitekturer, både koaksialkabelnett og parkabel-løsninger.

TITAN-prosjektet tar utgangspunkt i at alle kunder vil ha vanlig telefoni (POTS).



Figur 7.1 Flyttdiagram av de elementer som inngår i TITAN-prosjektet

Deretter antar vi forskjellige vekstrater for N-ISDN (f eks en eksponentiell vekst fra 1 % i dag til 20 % om 10 år). Aksessnett må i tillegg kunne tilby variable bredbåndstjenester fra 2 Mbit/s til ATM som er 155 Mbit/s etter behov. For moderate bitrater mindre enn 5 Mbit/s antar vi at ny modulasjons- og kode-teknologi gjør det mulig å benytte kopperpar for dette behov i store deler av nettet. Denne båndbredden vil gi en videokvalitet som tilsvarer det dagens videospillere yter.

8 Implementering av den utvidede prognosemodellen i TITAN

For å kunne utnytte den utviklede prognosemodellen i regneprogrammet har man delt programmet i tre deler.

En del av programmet, kalt "TITAN", inneholder alle beregningsfunksjoner slik som prisutviklingsfunksjonen i likning (6.11). En annen del av programmet, kalt "Main", inneholder en geometrisk modell, en liste over alle elementer som inngår, alle penetreringskurver både for tjenestetilknytninger og komponentmaterialbehov, og dessuten hele det økonomiske regnskap.

Den siste delen er en database som inneholder alle nettverkskomponenter. I tillegg inneholder databasen flere grupperinger, kalt klasser. Hver klasse inneholder en liste av elementer. For eksempel finnes det en lærekurveklasse (se tabell 8.1) hvor elementene er karakterisert ved sine K verdier (se likning 2.2). Videre er det definert en volumklasse hvor elementene er tilordnet et par av parametrene ($N_r(0)$, ΔT) (se tabell 8.2). Dessuten har databasen en drift- og vedlikeholdsklasse og en avskrivningsklasse.

Hver nettverkskomponent er tilordnet et element fra hver av klassene. For eksempel består elementene i lærekurveklassen i forskjellige K verdier slik at gammel og kjent teknologi har K verdier fra 0,9 til 1,0, mens nyere teknologi får lavere K verdier.

Volumklassene er karakterisert ved parameteren $n_r(0)$ og ΔT . Her vil for eksempel komponenter som er nye, men som en tror vil bli introdusert raskt i nettet, ha lave verdier både på $n_r(0)$ og ΔT , mens nye komponenter der utviklingen tar lengre tid får lav $n_r(0)$ verdi og stor ΔT . For eksempel vil linjekortet for ISDN ha $n_r(0)$ verdi på ca 0,01 og ΔT på mer enn 10 år.

Denne måten å beregne kostnader på utnytter de kunnskaper en til enhver tid har om nettkomponentene enten de er gamle og velkjente som kopperpar eller nye på prototypnivået. Det vil si at beregningsmetoden blir mest mulig stabil.

Dersom en ønsker å gjøre beregninger av kostnadene og andre økonomiske faktorer ved langsiktige prosjekter (f eks nettbygging), er det viktig å ha en best mulig beskrivelse av hvorledes priser/kostnader vil utvikle seg i tid. Vi har beskrevet hvorledes prognostisering ved hjelp av regresjonsanalyse anvendes hvis det er et stort nok dataunderlag for de komponentene som inngår. Dersom datagrunnlaget er lite (f eks ved nye komponenter) har vi vist hvorledes kostnadsfunksjonen kan estimeres ved å bryte den ned til de enkelte elementene som virker inn på kostnadsfunksjonen. Ved så å klassifisere de komponentene som inngår i beregningene ut fra sine egenskaper, kan en håpe å beskrive kostnadsutviklingen for prosjektet relativt bra. Som et eksempel er det vist hvorledes disse metodene blir benyttet i det strategiske planleggingsverktøyet som er utviklet i RACE-prosjektet TITAN. Flere andre RACE-prosjekter og prosjekter innen EURESCOM

ønsker å benytte disse metodene for sine kostnadsberegninger.

Referanser

- 1 Wright, T P. Factors affecting the cost of airplanes. *Journal of Aeronautic Sciences*, 3(4), 122–128, 1936.
- 2 Crawford, J R. *Learning curve, ship curve, ratios, related data*. Lockheed Aircraft Corporation, 1944.

Tabell 8.1 Lærekurveklassen med K verdier for henholdsvis optimistiske og pessimistiske anslag

Lærekurveklasse		K
I	Optiske kabler	85/75
II	Optiske passive komponenter	75/70
III	Sendere og mottakere	80/73
IV	Elektroniske komponenter	80/73
V	Innbygde komponenter	95/90
VI	Installasjoner	105/95
VII	Byggtekniske komponenter	105/95
VIII	Kopperkabler	103/98

Tabell 8.2 Volumklasse med parametrene $n_r(0)$ og ΔT

Volumklasse	$n_r(0)$	ΔT
Vol.klasse1	0,5	5,00
Vol.klasse2	0,1	5,00
Vol.klasse3	0,01	5,00
Vol.klasse4	0,5	10,00
Vol.klasse5	0,1	10,00
Vol.klasse6	0,01	10,00
Vol.klasse7	0,01	50,00

- 3 de Jong, J R. The effects of increasing size on cycle time and its consequence on time standards. *Ergonomics*, V1, 51–60, 1957.
- 4 Yelle, L E. The learning curve: historical review and comprehensive survey. *Decision Sciences*, 10(2), 302–328, 1979.
- 5 Stordahl, K, Hjelkrem, C, Bøe, J. Regresjonsmodeller. *Telelektronikk* 90(1), 87–102, 1994 (denne utgave).
- 6 CCITT. *GAS 10 Handbook. Data acquisition and forecasting methods for telecommunication planning*. 1987.
- 7 Olsen, B T et al. Tool for introduction scenarios and techno-economic studies for the access network from RACE 2087/TITAN. *Proceedings of RACE Open Workshop on Broadband Access, June 93*. Nijmegen, Nederland, 1993.
- 8 Zaganiaris, A et al. A methodology for achieving life cycle costs of optical access networks from RACE 2087/TITAN. *Proceedings of the Eleventh Annual Conference on European Fibre Optic Communications and Networks (EFOC&N), June 93*. Haag, Nederland, 1993.
- 9 Zaganiaris, A et al. Life cycle costs and economical budget of optical access networks from RACE 2087/TITAN. *Proceedings of the IEEE/5th Conference on Optical/Hybrid Access Networks, September 93*. Montreal, Canada, 1993.

Appendiks

Betrakt den relative prognosefunksjonen for volum

$$n_r(t) = \left(1 + e^{a+bt}\right)^{-1} \quad (\text{A.1})$$

Sett inn for $t = 0$

$$n_r(0) = \left(1 + e^a\right)^{-1} \quad (\text{A.2})$$

Dette gir

$$a = \ln(n_r(0)^{-1} - 1)$$

Sett inn definisjonen for t_2 og t_1 i likning (A.1) slik at $n_r(t_2) = 0,9$ og $n_r(t_1) = 0,1$.

Dette gir etter litt regning:

$$e^a e^{bt_2} = \frac{1}{9} \quad (\text{A.3})$$

$$e^a e^{bt_1} = 9 \quad (\text{A.4})$$

Ved å dele disse på hverandre får vi

$$e^{b(t_2-t_1)} = \frac{1}{9^2} \quad (\text{A.5})$$

Videre kan vi skrive:

$$t_2 - t_1 = -\frac{2 \ln 9}{b} \quad (\text{A.6})$$

som gir:

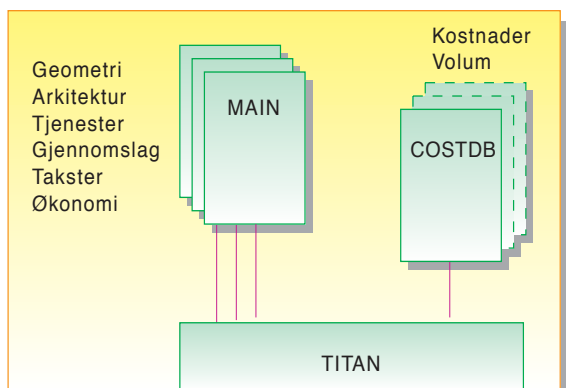
$$\Delta T = -\frac{2 \ln 9}{b} \quad (\text{A.7})$$

eller

$$b = -\frac{2 \ln 9}{\Delta T} \quad (\text{A.8})$$

a og b kan da innføres i (A.1) som gir:

$$n_r(t) = \left[1 + e^{\left\{\ln(n_r(0)^{-1}-1) - \frac{2 \ln 9}{\Delta T} t\right\}}\right]^{-1}$$



Figur 8.1 Oppdelingen av programvaren i tre deler

Scenarios for introduction of broadband services in telecommunication networks

BY ROLF B HAUGEN AND BENTE MANNSÅKER

Abstract

621.39.05

There are large activities world-wide to specify and experiment with broadband networks and services. At the moment most people believe that the broadband network of the future will be based on ATM technology, and switching manufacturers and research institutions are making detailed plans for how ATM technically is to be incorporated into next generation's systems and networks. But the crucial question for the PNOs is how to economically introduce a B-ISDN concept; it is a fact that, except for CATV, there is really no broadband service that at the moment is attractive enough to pay for the enormous investments necessary. We discuss in this paper two types of broadband services: video related services and data network services, the former type being the traditionally discussed end user services, whereas the latter is the aggregation of services from several users connected to various types of local networks. We argue that the data network services will be the driving force towards a broadband network and discuss some scenarios from broadband islands towards a B-ISDN.

1 Introduction

Broadband services have been very much discussed, and even experimented with, during the last couple of decades. Some services, like cable TV, have found their implementations in dedicated networks. Others, like picture phone and video conferences, have been introduced in the public telephone networks by digital coding (compression) of the video signals to smaller bandwidths. But, except for this recent breakthrough in compression techniques, there is still no real broadband service offered in public networks that can be said to run on sound economical terms.

In Europe, there has been a tremendous effort the last five years to specify a public broadband network as well as to initiate the development of components and subsystems that would make such networks cost effective. Several long term EC projects have been established to this end; RACE being the most influential within the telecommunications area.

By now, the evolving broadband network, B-ISDN, based on ATM technique seems to be rather well defined and agreed upon among telecom operators. But one thing is *specification* of a network, another is *cost effective* implementation; i.e. at prices that subscribers can afford. Needless to say, the subscriber loop is the overwhelming bottleneck.

Optical fibres in the subscriber loop are considered to be a prerequisite for broadband services. For *business users* fibre connections are today, economically spoken, within reach; in fact, medium and large companies might already be connected to the public network via fibre cables. This is by no means the case for *residential* subscribers and small firms. For these categories the cost of installing optical fibres are still an order of magnitude larger than with ordinary copper cables. And for residential users, most PNOs (Public Network Operators) are for political/regulatory reasons excluded from offering the most popular broadband service of today, namely cable TV. An integration of TV and tele services on the same cable might have rendered a fibre optical solution into residential homes cost effective in the near future. The RACE project TITAN is developing computerized tools for calculating (life cycle) costs of optical access networks. With this tool the network planners can extract detailed estimates of total and partial fibering of the local loop as a function of the years up to and beyond year 2000. In the computer program there are forecasting models for service demands and prices of the various network components.

Services that will be driving forces for the PNOs to start implementing broadband networks can be divided into two classes

- video related services
- data network services.

The first class contains services of the type: video telephony, video conferences, TV distribution, multimedia communications, etc. These services typically direct themselves to *one*, or at least a limited number of end users. To offer videophone or TV distribution services there must be a broadband connection all the way to the end user. That is what makes these types of services hard to offer in the public networks.

The other class of services is of the type: LAN-LAN interconnect, high capacity main frame communication (Cray), etc. These types of services direct themselves towards clusters of users that are hooked on to local networks, hence the "users" are often LANs. And it is the *aggregation* of traffic that creates a need for broadband solutions rather than the demand from single users.

To begin with, attention was directed towards the first class of users; in fact, the very notion of B-ISDN somehow leads our minds towards an end user concept. Hence, most broadband services that were experimented with throughout the eighties were of this type, and there was, and perhaps still is, an increasing frustration for not being able to bring these services to the customers at prices they would be willing to pay.

But it is the second class of services that presently "takes off" and requires solutions. Up to now, inter-networking solutions have mostly been provided by *bridges* and *routers* that are interconnected by leased lines. More recently, solutions like Frame Relay, MAN (Metropolitan Area Network), ISDN and even ATM are being discussed.

From a PNO's point of view, the latter type of solutions are by far the more interesting. Could intra-LAN traffic be transported in the switched (public) network? And which type of the above mentioned network solutions have the potential to evolve into the broadband network of the future?

These are the questions many PNOs are discussing today. They all know that there are tremendous revenue potentials in the multimedia revolution that will eventually come, a revolution that is just waiting for cheap network solutions. On the other



Figure 1 Two LANs interconnected via routers/bridges and leased lines

We see two trends concerning LANs: Within larger companies and institutions there is a need for splitting up large LANs into smaller segments, the reason being both of traffic and security origin. However, certain services, like administrative routines, will still require interworking between the segments. The other trend is to interconnect different LANs between various geographically separated departments or branches of a company, so as to build one large transparent network. Hence, in both cases the result is an ever increasing need for LAN interconnections.

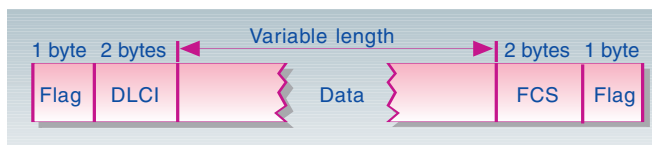


Figure 2 The frame format (HDLC) used for Frame Relay

hand, the data communications area is now an area of competition, and, needless to say, cost effective solutions always have to be found as one proceeds.

2 LAN interconnections

During the last few years we have witnessed a dramatic transition from centralized, proprietary main frame solutions to distributed network solutions with an ever increasing number of interworking PCs or UNIX workstations. The workstations have access to powerful servers over local area networks of the type Ethernet (10 Mbit/s), Token Ring (4–16 Mbit/s) and more recently FDDI (100 Mbit/s).

The number of LANs has thus increased and the traffic has exploded. The number of LANs is about 4 millions world-wide and is growing at a very high rate: according to (1) the annual growth rate in Europe is about 40 %. The growth in traffic is, however, not only due to the increasing number of users, but also to a great extent a result of new applications with their need for transport of large amount of data and short response times. The shift from main frames to network servers enlarges this trend even further. This evolution creates a need for *effective* infrastructures that, in addition to handling the transport of data, satisfy the necessary requirement of *reliability* and *security*. In this multitude of needs and demands proprietary solutions like SNA will be faded out; *open standards* become a must.

The most common means of interconnecting LANs today is by *bridges* and *routers*. Bridges, being the simpler of the two, work on the OSI link layer (2nd layer) and are limited in routing and load sharing capabilities. Routers, on the other hand, operate on the network layer (3rd layer) and work between LANs with the same *network layer*. They support various parallel paths in the network and have better security and load sharing facilities. There is a continuous dialogue between the routers in a network. Information of the traffic load in the network will thus be known by all the routers and an optimal routing can be chosen. Needless to say, routers are more expensive than bridges. Figure 1 depicts a typical network with bridges or routers.

Bridges as well as routers have been on the market for several years and represent today an important product segment for the data dealers. In (2) is made a survey of sold equipment in Norway up to 1992 and it thus gives an indication of the potential market for the PNOs to attack. The study indicates, however, that the need for capacity for the LAN interconnection is, at the moment, rather modest. This might change as soon as FDDI LANs penetrate the market.

The traffic out of LANs is typically bursty, i.e. outgoing line might be idle for a certain period followed by a period of peak traffic. Hence, leased lines might not be the most cost effective solution. As a result network solutions like *Frame Relay*, *MAN* and *ATM* are being considered. In designing these networks, however, attention has to be paid to *delay* and *throughput*. The delay should be of the same order as in the connecting LANs. Usually the requirement is set to be of the same order as the delay between two bridged LANs; this delay admits an absolute maximum of one second and a mean value of 200 ms.

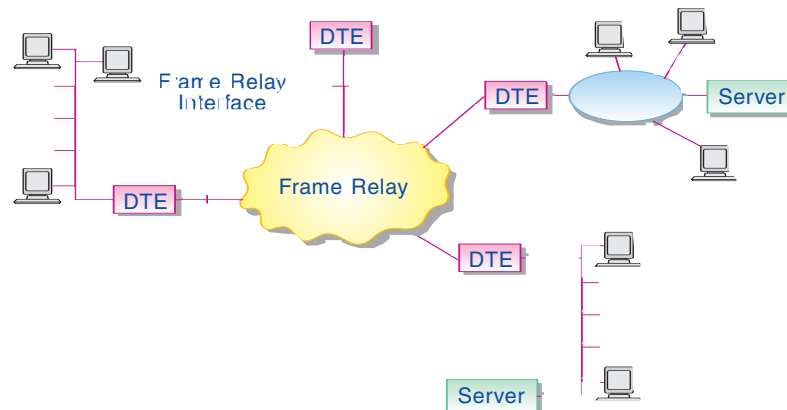


Figure 3 Local area networks interconnected with Frame Relay

3 Network structures

In this chapter we will describe some of the network candidates for interconnecting LANs.

3.1 Frame Relay

Frame Relay is a new standardized connection-oriented data service where the data is transported in variable length, HDLC based frames between end users. The service is intended to be a cost effective alternative to leased line for LAN interconnection, and might in the short term be the most attractive way of LAN interconnections. The service was first introduced in the US and was followed by several European countries by the end of 1992. Norwegian Telecom will provide Frame Relay services from January 1994.

The HDLC (High-level Data Link Control) carries, in addition to the variable data field, Data Link Connection Identifier (DLCI) that identifies the call and Frame Checking Sequence (FCS) that checks for bit errors. Since there is a dynamic allocation of bandwidth, the service is well suited for bursty traffic which is the typical traffic pattern for LAN interconnections.

A Frame Relay network is made up of network nodes and user terminals, DTE (Data Terminal Equipment). The latter connects the user to the network. DTE can be a PC, bridge, router or host computer that is provided with interface defined for Frame Relay. The standard, which is described in both CCITT and ANSI recommendations, defines signal and data transmission at link level, OSI level 2, in the interface between user equipment and network.

The first versions of Frame Relay work in a Permanent Virtual Connections (PVC) mode which indicates that the paths through the network are predefined by the network operator. When the DTE transmits frames to the network, the nodes will read the identifier, DLCI, look up in tables to find the right outgoing channel and send the frame to the next node (or end terminal). Use of several identification codes permit several parallel sessions in different directions to coexist on the same physical link. In this way a DTE can communicate simultaneously with different destinations over the same access to the network, a prerequisite for LAN users.

FCS is used to check that the frame has been transferred correctly. Frames indicated as erroneous by the FCS are discarded; means of ensuring an error-free DTE-DTE transmission has to be implemented in the higher-level protocol. If errors occur and are corrected, data must thus be retransmitted through the entire network. This obviously takes longer time than

retransmission only over the link where the error occurred. Hence, the quality of the transmission medium plays an important role for Frame Relay. An approximate limit where Frame Relay is effective is set to a BER $< 10^{-6}$ on individual links. This error rate can be met with speeds of several Mbit/s over most transmission links today; a maximum link transmission speed of 2 Mbit/s is, though, at present defined. Higher speed (2–34 Mbit/s) is considered, but it is a question whether other broadband services, both connectionless and connection-oriented, might be more suitable for this part of the market.

3.2 Metropolitan Area Network (MAN)

MAN is in this context understood as a network consisting of subnetworks that use the IEEE standard 802.6, Distributed Queue Dual Bus (DQDB), as access protocol on a dual bus (see figure 4) which also uses principles defined by ETSI.

The DQDB protocol will eventually support both connectionless and connection oriented (included isochronous) services, at the moment only the connectionless one is specified.

The interest in connectionless service stems from the need to provide cost effective services with high throughput and low delay for bit rates higher than the ones provided by Frame Relay. Bellcore thus defined the Switched Multi-megabit Data Service (SMDS) that operates on US standards for bit rates. ETSI has specified a European counterpart which is named CBDS (Connectionless Broadband Data Service). CBDS offers bit rates of 2 Mbit/s, 34 Mbit/s, 139 Mbit/s and 155 Mbit/s, i.e. bit rates compatible with existing transmission hierarchy in the public network. The connectionless service features make CBDS particularly suitable for LAN interconnections and might be attractive in a medium and long term perspective. As FDDI, with its nominal rate of 100 Mbit/s, is becoming widespread, the bandwidth requirement for LAN interconnection might come up in the region 30–60 Mbit/s which can easily be handled by CBDS.

CBDS being a service in MAN may also be offered over ATM. Offered in MAN it gives regional coverage, offered over ATM it gives unlimited coverage depending on the penetration of ATM. The service offers transmission of data units of variable length that tolerate variable delay.

Broadband connectionless services are offered in a number of field exper-

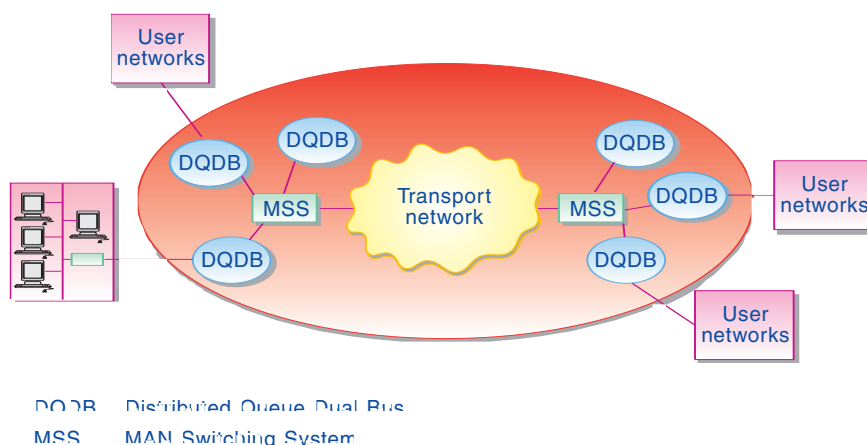


Figure 4 MAN architecture

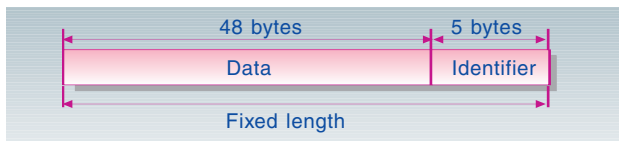


Figure 5 ATM cell

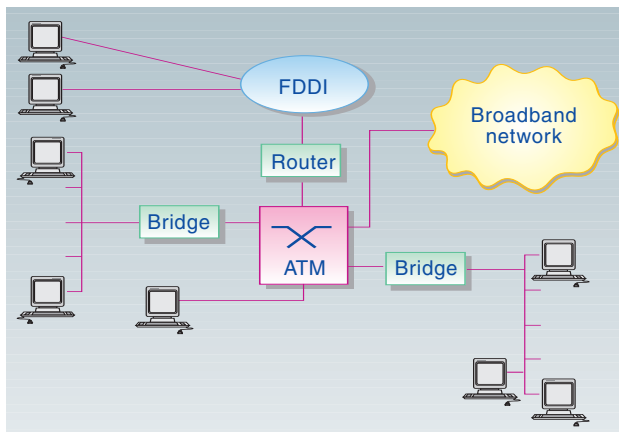


Figure 6 Use of ATM in local environment

iments in Europe and the USA. In the USA commercial service is on the point of starting. However, full exploitation will not be ready for another couple of years.

3.3 ATM (Asynchronous Transfer Mode)

ATM is a connection oriented cell-based technique which provides flexible bandwidth on demand up to 155 (or later 622) Mbit/s and integration of traffic with different quality of service characteristics. The data is transmitted in fixed packets called cells.

It is believed to be *the* broadband technique of the future and seems to come quicker than previously believed. An early specification is contained in CCITT recommendation F.811 "Broadband Connection Oriented Bearer Service", but the various RACE projects and now recently, the ATM Forum, are de facto standardization bodies.

The data manufacturers work actively through the ATM Forum to have ATM specified according to their needs. Their interest at present lies in various LAN interconnects or even substitutions for ordinary LANs. To the largest possible extent, the Forum uses CCITT standards. However, in the case where no standard has been developed, they issue their own recommendation. For the physical medium they have, in addition to the CCITT's, made specifications for multimode optical cable, shielded twisted pair and the plesiochronous hierarchy.

ATM interface cards for workstations have been on the market for more than a year. In fact, small companies in the data business were formed to exploit the ATM development, Fore, MPR

and Adaptive being three of the forerunners. The initial interfaces make use of the asynchronous byte transfer method supported by the TAXI chip for the physical transport medium. It uses the same cables as FDDI and the same coding. This allows reuse of components and existing cabling. The manufacturers claim, however, that they will also introduce Sonet/SDH solutions for the physical layer.

In figure 6 is shown an ATM switch used partly as a LAN interconnect, partly as a LAN itself. The workstations have inserted an ATM interface card based on the TAXI chip and can thus communicate directly. The next step would be to deploy an ATM cross connect in the public network to connect ATM islands of the type above.

As an indication of price: An ATM interface was initially (i.e. about a year ago) priced at USD 4,500; this has now (1993) dropped to USD 1,900, which indicates a take off in the market. An 8 x 8 switch costs about USD 30,000.

4 Access Networks

As previously mentioned, the very concept of B-ISDN was based upon the idea of providing broadband services to the end users. Several possible services have been defined and experimented with, most of them related to some form of *image services*. Video Telephony and Video Conferencing are perhaps the most discussed services, but also high quality graphics and bulk data transfer will require large bandwidths. Another important area are services related to entertainment, i.e. TV, or HDTV, distribution, video on demand, etc. Many of these services can be combined into the ubiquitous term: *multimedia*. This latter service might very well be the driving force of a B-ISDN of the future.

However, coding techniques have brought forward most of the above mentioned services over ordinary twisted pair that constitute the subscriber loop of today. Videophone services of 64 kbit/s or 128 kbit/s are already offered in ISDN, and even TV signals might be brought several kilometres over twisted pair in the experimental ADSL system of Bellcore. It remains to be seen whether the users will be satisfied with the quality thus provided in the long term.

The subscriber loop of the traditional public network is a multiple, usually triple, star configuration as depicted in figure 7. From the Main HUB (or local exchange) it typically fans out in a tree structure with hierarchical distribution points in the form of mechanical racks of cross connects. The more modern network has in addition remote subscriber units (RSU) in the form of switching machines, and in larger cities there might even be fibre optical terminating HUBs for providing high capacity links to business users.

Depending on the traffic demand, optical fibre cables might already today be cost effective out to the RSUs. But, except for large business users, fibre is not yet economical neither to the homes nor to the last distribution points (curbs). For real (non compressed) broadband services the transmission medium has to be optical fibres. Hence, fibering the very last mile out to the end users is the challenge for the PNOs the next decade, a chal-

lenge that has to be resolved before we can talk about a public broadband network!

There are several ways to deploy fibre optics into the subscriber network. The most straight forward way is to replace the distribution points by optical couplers (splitters). A splitter is a pure passive component (crystal) that splits the incoming optical beam into n sub-beams, each with $1/n$ of the original intensity. The result is a network of the type depicted in figure 8 that is commonly known as PON (Passive Optical Network).

British Telecom designed what they called TPON (Telephony over PON) overlaid the PON structure. TPON aims at being cost effective for ordinary telephone service (3). Transmission to the subscribers is based on TDMA; the first versions typically had a 20 Mbit/s multiplexer in the local exchange. Two alternatives were considered: Fibre To The Home (FTTH) and Fibre To The Curb (FTTC).

In Europe, Deutsche Telekom has several FTTC projects in the previous East Germany that are very much based on the above mentioned principles (4). Several FTTH pilot projects do also exist (5), but the cost of the in-house optical terminals is presently too high to give hope for even getting close to copper solutions. However, ETSI is now specifying digital interfaces (V5) for optical terminations inside local exchanges. This makes the multiplexers superfluous, and the optical subscriber loop more economically interesting; in fact, FTTC is now within reach!

In RACE, there is an ongoing project called TITAN (Tool for Introduction strategies and Techno-economic studies for the Access Network) that considers various economical aspects of an optical subscriber loop for residential and small business users. TITAN's scope is the calculation of discount system cost, operation, maintenance and powering costs, life cycle costs and the overall financial budget. Their aim is to compare different optical access networks as well as those of the ordinary twisted pair network of today. The costs of all network components are calculated using the extensive, integrated cost database developed within RACE, containing data gathered from many European sources. The database includes unit prices for components, installation, civil works, etc., and the calculations are performed for various network structures, depending on subscriber density, splitting ratios, number of splitting stages, and coupler locations. Assumptions are made concerning demand for different demographic situations, number of users and traffic. The services considered are: N-ISDN, broadcast TV channels and narrowcasting channels offering tele-working, tele-shopping and tele-interaction (tele-learning and tele-entertainment).

TITAN plans to demonstrate cost trends for the various network block elements over the study period 1995–2005. They will also propose a methodology for estimating operation and maintenance costs that allow cost comparisons of different access networks and enable the PNOs to study the effect of various tariff policies on cash flow. TITAN will finally propose a set of recommendations for a reasonable implementation policy.

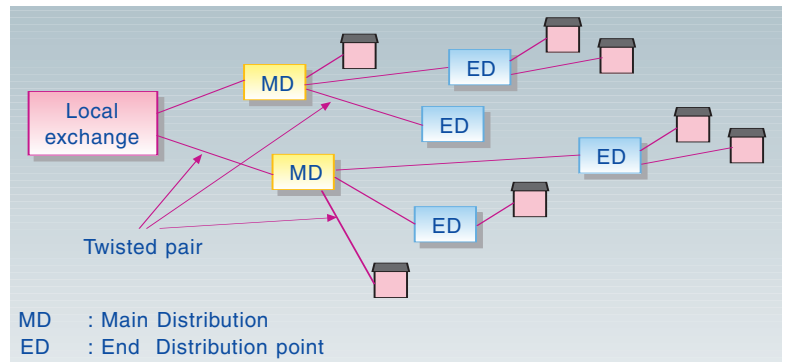


Figure 7 Traditional network structure of subscriber loop

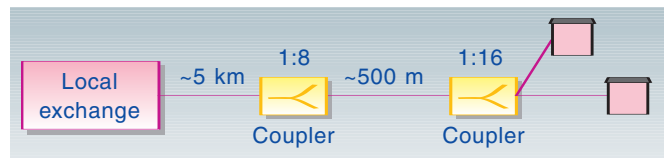


Figure 8 A typical subscriber loop based on PON

5 Evolution towards ATM

There are obvious different paths to broadband networks. One thing is the willingness to pay for higher bandwidth, another is the demographical situation in each country or region. Clusters of universities in one and the same urban area will probably require different intermediate term solutions than do small companies scattered sparsely over a wide area. However, the leading star is an ATM based broadband network, and the PNOs need to make their own strategic way from the present network to B-ISDN of tomorrow, the challenge is to find the most optimal way.

5.1 Scenarios with Frame Relay

Step 1

We know from (2) that a typical business user in Norway today with branches in different geographical locations uses X.25 or leased line with capacity less than, or equal to, 64 kbit/s to connect their LANs. Hence, there does not seem to be a great demand for network solutions with access bandwidths larger than 2 Mbit/s for the next planning period.

Our typical user will furthermore have a PABX connected to the public network and possibly a need for a dedicated 2 Mbit/s service for video conferences. All these services are connected to the access network via a multiplexer (static or statistical). If the total need for bandwidth is high enough he might have a fibre cable to the local exchange where the different services will be directed to the appropriate network.

Since the need for data communication is basically between a limited number of sites (branches) with a bandwidth demand of

less than 2 Mbit/s, Frame Relay will be an economical solution. And if a customer wants more random open communication with many other customers, it will be an economical solution to lease a logical channel via a *router service*. This will provide an open connectionless service for the customer. In fact, this may be a suitable solution for electronic mail services.

In figure 9 is depicted a network structure that we believe to be a cost effective solution for the short and medium term for the type of business infrastructure that we have in Norway. And, as we will see, it has a structure that easily evolves into an ATM network.

The traffic in the network will continue to grow as a result of a general increase in data applications but also from the increasing use of graphics and visualization in applications. This will

gradually move the customer towards higher bandwidths and drive the networks towards their limiting capacity.

In the core network there will be a continuous upgrading with SDH based optical systems and cross connects for a flexible handling of the various services. SDH based solutions will probably grow further out in the regional networks and also in the local loop for medium and large business users.

Step 2

There will be an overlay of ATM nodes (cross connects and switches) in the core network. This is partly due to a demand for interconnecting FDDIs and local ATM-LANs, and partly a need for interconnecting Frame Relay networks in a flexible way. Figure 10 visualizes the next step in the evolution. In this scenario the ATM network is transparent for the Frame Relay services. This means that the various functions for the service have to be performed in the Frame Relay nodes (FR), and ATM is solely responsible for transmission between the FRs. Conflicting situations between the overload controls in the two networks might, however, arise in this scheme.

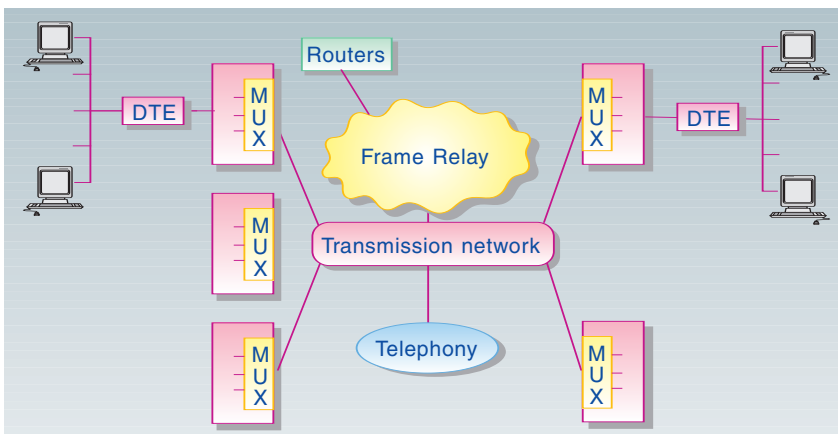


Figure 9 Frame Relay overlay: a first step towards a broadband network

Step 3

The next step would be to introduce InterWorking Units (IWU) between the Frame Relay and ATM networks. This can be done by associating each DLCI with a virtual channel (VCC) in ATM. Then each Frame Relay connection can also be identified in the ATM network, the individual users can be given the same QoS and overload control can be harmonized.

Step 4

At this stage the Frame Relay nodes are replaced by ATM nodes. This situation will arise when the customers start to use new multimedia applications that need larger functionality from the LANs. ATM based local area network will frequently be the interface to the public network. Users that want to stick to Frame Relay will be provided with ATM adapters. Since both services are connection oriented and operate in a packet mode, such adapters are assumed to be relatively cheap.

The connectionless router based service can stay in the network until the traffic takes off. It might then be changed into an open CBDS server. This service will have the same type of functionality as the router service, but offer higher throughput. The combination of a permanent virtual link service for the main traffic, and an open connectionless service for the more random switched traffic gives a flexible structure for the early broadband networks.

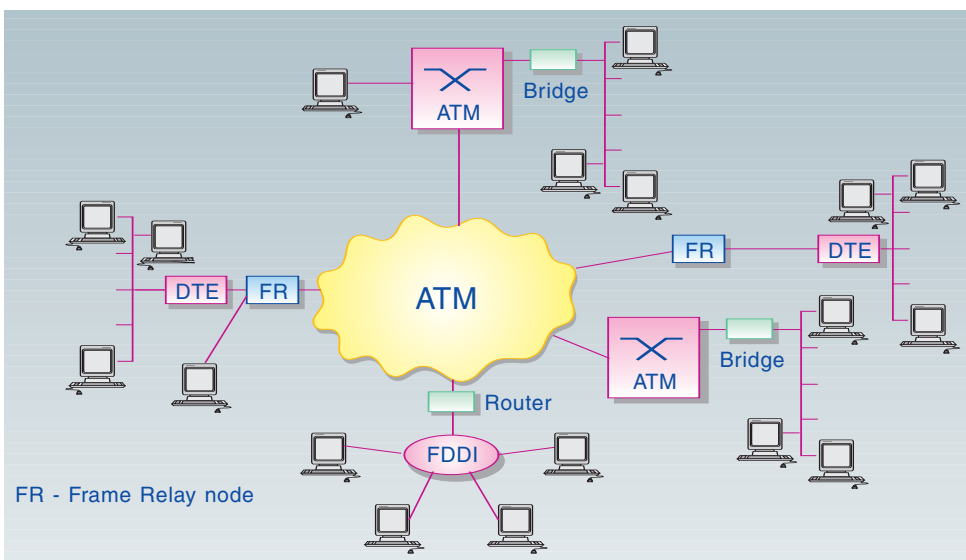


Figure 10 Step 2: Coexistence between Frame Relay and ATM

It should be added that this scenario is assumed to be valid for business users. For residential users the situation is more unclear. There is, however, an ongoing development for ATM over PON. Whether such a solution will be allowed to carry TV channels and whether it will prove to be cost effective compared to coax cables remains to be seen.

5.2 Scenarios with MAN

Possible early users of MAN will be universities, high-technology companies, petrochemical companies, etc. For these branches there is often use of common resources like super computers and distribution of new video based services from the universities.

Step 1

There are islands of communications using MANs. As the number of these centres in Europe increases they want to interconnect. The first step would be to interconnect the connectionless networks using routers. The use of SDH based cross connect would belong to this phase. These leased line services can also be connected by SDH based cross connects. The situation is basically described in figure 4.

Step 2

The next step would be to interconnect the MANs with ATM. At this stage the transport mechanism in ATM is based on permanent connections, i.e. virtual path VP. The use of VP does not necessitate add-on functions in ATM to interconnect the MANs. All the MANs have a logical connection between themselves in a meshed topology.

Step 3

At this stage a connectionless (CBDS) server is introduced in the ATM network. The CBDS server in ATM has some of the same functions as the switching system in MAN; i.e., on the basis of an E.164 address it can route the connectionless packet through the ATM network to the user.

MAN also supports a connection oriented service. This is only a leased transmission line service within a limited distance. MAN can also be used as an access network to ATM. The bus is then used as a concentrator of traffic from different subscribers that only have small traffic interest between them. This may be an early cost effective way to connect subscribers to a network in high density areas.

Step 4

The next stage would be a B-ISDN network based on ATM enhanced with CBDS. Although DQDB and ATM have the same cell structure the network structures and functions are quite different. When the users move to ATM based network with full switching capabilities, the MAN networks will be a bottleneck. It is therefore hard to see how a MAN in a flexible way can evolve into a general access network of B-ISDN. However, if CBDS is the only service requested, the MAN can serve as the access network.

6 Conclusions

There is in this report indicated two possible scenarios for introduction of a public broadband network. The first scenario, starting with routers and frame relay and evolving towards an ATM network, is believed to be a cost effective strategy for small to medium populated countries. The second scenario, use of MAN technology, might be a cost effective and convenient choice for countries with clusters of universities, large companies, institutions, etc. that have the need for switched broadband services between each other.

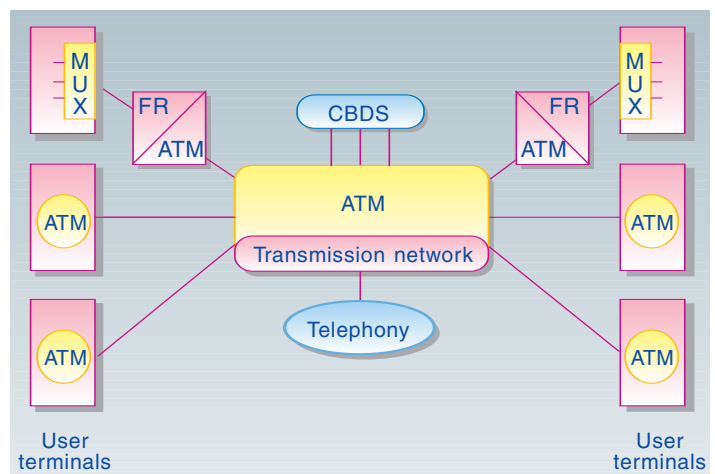


Figure 11 Step 4: Broadband network for business users



Figure 12 MAN interconnect via ATM based on virtual path

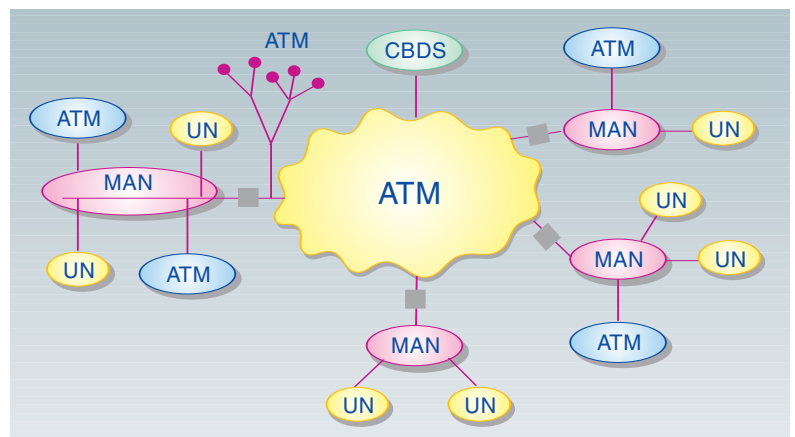


Figure 13 Broadband networks evolving from MANs

References

- 1 Mannsåker, B et al. *Overall strategic studies, Task 1.2: Telecoms service trends*. 1992. (Eurescom P111.)
- 2 Svendsen, B. High speed data services for LAN interconnection in Norway. *Proc. IBC&N*, 1993.
- 3 Faulkner, D. Optical networks for local loop applications. *Journal of lightwave technology*, 7, 1741–1751, 1989.
- 4 Stegmeier, H. Progress in optical information technology. *Proc. Telecom 91, Technical Symposium*, 1991.
- 5 Grotjohann, H. *Fibre-to-the-home project with PTT Telecom: OLN-P*. AT&T brochure.

Status

International research and standardisation
activities in telecommunication

Editor: Endre Skolt



Table of contents

	page
Introduction	
<i>by Endre Skolt</i>	183
Signal processing	
<i>by Gisle Bjøntegård</i>	184
Introduction	
COST	
ITU-T	
ISO/MPEG	
ETSI	
Data networks and open system communications	
<i>by Berit Svendsen and Mette Røhne</i>	186
Introduction	
ETSI	
EURESCOM	
ITU-T	
Teletraffic and dimensioning	
<i>by Harald Pettersen</i>	188
Introduction	
ITU-T	
ETSI	
ATM Forum	
EURESCOM	
RACE	
COST	
Terminal equipment and user aspects	
<i>by Trond Ulseth</i>	191
Introduction	
ETSI9	
ITU-T	
EURESCOM	
COST	
Telecommunications Management Network	
<i>by Ståle Wolland</i>	194
Introduction	
ISO/JTC	
A presentation of EURESCOM	
<i>by Endre Skolt</i>	196
Introduction	
Objectives	
The organisational structure	
Members	
Results achieved	
Work Programme 1994-98	
References	198

Rapporteurs of Teletronikk Status section

Study area	Editor
Terminal equipment and user aspects	Trond Ulseth
Service definitions	Ingvill H Foss
Message Handling and Electronic Directory	Geir Thorud
Health care informatic standards	Sigurd From
Signal processing	Gisle Bjøntegård
Radio communications	Ole Dag Svebak
Intelligent Networks	Endre Skolt
Telecommunications Management Network (TMN)	Ståle Wolland
Languages for telecommunication applications	Arve Meisingset
Transmission and switching	Bernt Haram
Teletraffic and dimensioning	Harald Pettersen
Data networks	Berit Svendsen
Security	Pål Spilling
Broadband	Inge Svinnset

Introduction

BY ENDRE SKOLT

This is the second issue of *Teletronikk* which includes a status section on international telecommunication research and standardisation activities. The previous issue had reports on intelligent networks, radio communications, service definitions, broadband, and transmission and switching. In this issue of *Teletronikk* papers on the following areas are included:

- Terminal equipment and user aspects
- Teletraffic and dimensioning
- Signal processing
- Data networks
- Telecommunication management network.

The paper on terminal equipment and user aspect focuses on the communication aspects of multimedia, including both user-to-user communication and user-to-host/server communication. Results from current research and standardisation will make it possible to offer services in public switched networks (PSTN), ISDN and B-ISDN. Important results are accomplished on voice and audiovisual terminals, conference control services, and standards for regulatory purposes. The paper is provided by Mr Trond Ulseth.

Teletraffic and dimensioning may be defined as the statistical description of the use of resources in the network, methods to structure and to size system and network resources to meet service quality requirements, etc. The paper covers teletraffic studies in ATM networks, B-ISDN, and ISDN. Traffic studies related to intelligent networks (IN) and universal personal telecommunications (UPT) are also included. An important work item is traffic engineering which includes topics such as traffic measurements, network performance, dimensioning, algorithms for traffic controls, and traffic characterisation. Results on traffic control, resource management and congestion

control are also reported. The paper is provided by Mr Harald Petersen.

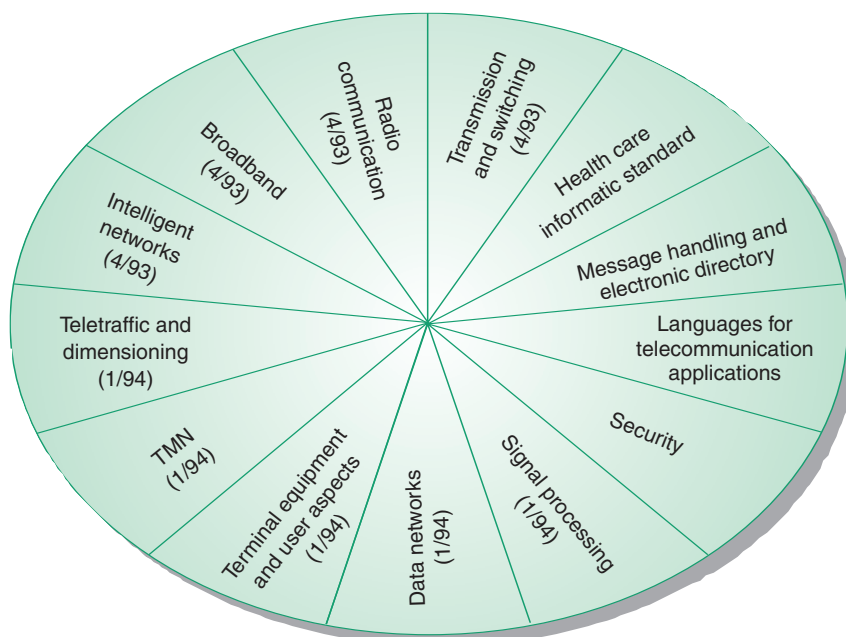
The paper on signal processing focuses on image compression. Compression techniques have made it possible to fit live video into low capacity networks originally designed for other purposes. Standardisation work on applications such as compression of live video with video phone quality on analogue telephone lines, teleconferencing quality in basic access ISDN, "at least VHS quality" in a 2 Mbit/s channel and "TV quality" in a 4 Mb/s channel, are reported. The paper is provided by Mr Gisle Bjøntegård.

In the area of data networks, the paper addresses the current status on Frame Relay. Frame Relay is a high speed data service (64 kbit/s–2 Mbit/s) suited for interconnection of local area networks. A Frame Relay bearer service will be available in ISDN, B-ISDN and dedicated networks. Important work items are service description, protocol specifications, interworking and congestion management. The presentation is provided by Ms Berit Svendsen and Ms Mette Røhne.

Telecommunications Management Network (TMN) is an architectural concept for communication and processing of management information for tele- and data communications networks. In a paper provided by Mr Wolland the status of the current work going on in the area of Open Systems Interconnection Management is presented. The presentation focuses on the result achieved in ISO/IEC.

In the previous issue of this status section we gave a brief presentation of the new structure of ITU. The last paper in this section is an introduction to the research and development institute EURESCOM, which is jointly owned by 26 European Public Network Operators.

006
001.89:654



Signal processing

BY GISLE BJØNTEGÅRD

006
621.391

Introduction

Signal processing has a large number of applications. In this presentation we will focus on image compression. The new compression techniques have made it possible to fit live video into networks originally intended for other purposes – speech or data transmission with low capacity. In this context the following applications are studied and achievements will be reported in more detail below:

- Compress live video with “video phone quality” to fit into a modem used on an analogue telephone line
- Compress live video with “teleconferencing quality” to fit into basic access ISDN (2B)
- Compress live video with “at least VHS quality” to fit into a 2 Mb/s channel
- Compress live video with “TV quality” to fit into a 4 Mb/s channel.

The development of compression techniques have largely been undertaken by standardisation bodies like ITU and ISO. NTR has actively taken part in the development related to all the applications mentioned above.

This summary presents the outcome and status of the work in the following research and standardisation organisations:

- European Co-operation in the Fields of Scientific and Technical Research (COST)
- International Standardization's Organization (ISO)
- International Telecommunication Union Standardisation Sector (ITU-T)
- European Telecommunication Standards Institute (ETSI).

Table 1 Main standards in the area of video compression

Name	Organisation	ID	Applications
JPEG	ISO	IS-10918	Compression of still pictures
H.120	CCITT	H.120	Video conferencing. 2 Mbps
H.261	CCITT	H.261	Video telephony, Video conferencing. 64–2000 Kbps
MPEG1	ISO	IS-11172	Video storage. 1.5 Mbps
MPEG2	ISO	IS-13818	Digital TV
H.262	ITU-T	H.262	TV quality video on ATM networks

COST

The COST 211 project (continued through COST 211bis and COST 211ter) has played an important role in developing the techniques used in all relevant standards today. In the following we have summarised some important achievements made by COST.

- The H.120 standard from the early eighties was completely developed in COST 211.
- The H.261 standard for video conferencing was largely based on inputs from COST 211bis.
- ISO/MPEG1 standard (see below) uses the same basic elements as H.261.
- COST 211ter made significant contributions to ISO/MPEG2.
- The ongoing development of the ITU-T very low bitrate video phone standard will use the model developed by COST 211ter as a basis for their work. The main job of developing this model is done by NTR.
- For the coming ISO/MPEG4 standard (see below) the proposal based on another COST model is expected to be an important input.

NTR had the chairmanship of the simulation group in COST 211ter during the last two years – which turned out to be a very important period for preparing methods for new video compression standards.

ITU-T

The H.261 standard was developed in CCITT/SGXV Working party XV/1. It was finalised in 1990 and is intended for bitrates $p \cdot 64$ kbit/s. $p=1,30$. The main participants in this development were national telecom operators. The chairman of the experts group was Mr. Sakae Okubo from NTT, Japan.

In 1991 CCITT/SGXV Working party XV/1 initiated a new experts group with the same chairman and with the mandate to develop a video compression method for ATM (Asynchronous Transfer Mode) networks. From the beginning it was decided to do this development in co-operation with ISO/MPEG. The meetings have been combined and CCITT/SGXV – now ITU-T SG15 – have decided that the resulting standard shall have common text with the ISO/MPEG2 standard. The contributions from ITU in this combined work have largely been focusing on issues of special interest for real time communications like coding/decoding delay and handling of the type of packet loss that may occur in ATM networks. The resulting standard is intended for full TV quality and bitrates of a few Mb/s (typically 2 – 8 Mb/s).

The standard will be called H.262 and will be finished late 1994.

In 1993 ITU-T SG15 initiated an experts group on low bitrate video phone. The scope of this group is to make standards for complete video phones to operate on modems over analogue telephone lines. The work will cover video compression, audio

compression modems and terminal aspects. The standard will be developed in two phases:

- A short term standard complete in 1995. The compression methods will be "fine tuned" versions of already available standards. The video coding method used here will largely be based on work done by NTR through the COST 211ter project.
- A long term standard. This work will be made in co-operation with MPEG4 and it is expected that totally new methods will be developed for audio as well as video compression. A standard is expected in 1998/99.

ISO/MPEG

The complete name of the group is ISO-IEC JTC1/SC29/WG11. This means that WG11 is the working group where the basic work is done. This working group is also called "Moving Pictures Experts Group" (MPEG). As a working group it is rather large – at the moment 200 – 250 participants to the meetings.

MPEG1 or ISO/IEC 11172 was completed in 1992. It is a standard for "Coding of moving pictures and associated audio". It contains separate parts for video, audio and multiplex. The bitrate is defined to be around 1.5 Mb/s. The resulting video quality is approximately "VHS quality". The audio part may have full broadcast quality.

MPEG2 was started in order to develop a standard with full TV quality. It is largely based on the same coding elements as MPEG1 but extended to include the full TV format. NTR has contributed actively to this standard by introducing new coding elements particularly suited for the TV interlace format. Typical bitrates for MPEG2 is in the range (2–10) Mb/s. However, MPEG2 may also use HDTV formats and thereby give HDTV quality with bitrates in the area (15–20) Mb/s. Thus the foreseen work on HDTV in MPEG3 was never initiated. The status of the standard now is Committee Draft – CD (November 93). After balloting the final standard (IS) is expected to be ready late 1994. The name of the standard will then be: ISO/IEC 13818.

A new work item has been started recently. It is named MPEG4 and will deal with new compression methods aiming at bitrates down to "a few tens of kbit/s" for live video. The current plan is to have a standard by 1998/99.

ETSI

ETSI is not playing a significant role in the development of video standards. ETSI NA5 has to some extent had the responsibility of co-ordinating European inputs to other bodies. However, this is becoming more difficult as the main participants become manufacturing companies rather than telecom operators. An important task of ETSI is to make ETSs based on standards already developed in other bodies (see above).

Data networks and open system communications

BY BERIT SVENDSEN AND METTE RØHNE

006:681.324
006:681.327.8

Introduction

The most widespread data communication service offered today by public network operators is the X.25 packet mode service.

Frame Relay is a high speed data service (64 kbit/s–2 Mbit/s) and is especially suited for interconnection of local area networks. Frame Relay is a simplified data service compared to X.25 and requires less processing in the network. Frame Relay is transferring data transparent in layer 2 frames without functions for sequence control, retransmission and error correction inside the network.

In this summary the latest results in the following three organisations are presented:

- International Telecommunication Union – Telecommunication Standardisation Sector (ITU-T)
- European Telecommunication Standards Institute (ETSI)
- European Institute for Research and Strategic Studies in Telecommunication (EURESCOM).

ETSI

NA 2, NA 5 and SPS 5 are the main sub technical committees doing standardisation work on data networks and data services.

NA 2 is responsible for network standards required to introduce a European Frame Relay service. NA 2 has also been responsi-

ble for packet mode service in ISDN. NA 2 has finalised the Packet Handler Interface (PHI) specifications and the packet mode service in ISDN. NA2 is currently working with service descriptions for Frame Relay service in ISDN and dedicated Frame Relay networks. The study is also focused on a Remote Frame Handler Interface (RFHI).

The following list of documents is the most important results:

- ETS 300 007 'Integrated Services Digital Network (ISDN); Support of packet mode terminal equipment by an ISDN'
- ETS 300 099 'Integrated Services Digital Network (ISDN), Specification of the Packet Handler Access Point Interface (PHI)'
- Draft RFHI 'Integrated Services Digital Network (ISDN) specifications of the Remote Frame Handler Interface'.

Three different drafts on service descriptions

- 'Frame Relay Service – General description'
- 'ISDN Frame Relay Bearer Service – Service Definition'
- 'Frame Relay Data Transmission Service – Service Definition'.

The plan is that NA 2 will finish the draft service definitions and will start to work on Frame Relay Network-to-Network Interface during 1994.

In NA 2 France Telecom, Alcatel Bell, Siemens and Swiss PTT have taken an active part.

NA 5 is responsible for the service definition of the Frame Relay service to be offered in B-ISDN and Frame Relay interworking. The work is focused on the support of Frame Relay in B-ISDN and interworking between B-ISDN and other networks.

The following documents are available as drafts:

- 'Draft B-ISDN – Frame Relay Bearer Service – Service Definition'
- 'Draft support of Frame Relaying Bearer Service in B-ISDN and Frame Relay interworking between B-ISDN and other networks'.

The plan is that NA 5 will finish these drafts in 1994. Nokia, France Telecom, Swiss PTT and Ellemtel have been the most active contributors.

The protocols at the User-to-Network interface for Frame Relay will be standardised by SPS 5. The standardisation will be based on the corresponding recommendations made by ITU-T.

EURESCOM

There is still one important item of Frame Relay that is not solved. This is the requirements for network-to-network interface so that Frame Relay networks from different operators can interwork. Therefore, in order to assist and enforce these studies in Europe, Project P301 Frame Relay Network-to-Network Interface definition was started in April 1993.

Table 2 Standards in the area of Frame Relay in NA2 and NA5

ID	Organisation	Name
ETS 300 007	ETSI NA 2	Support of packet mode terminal equipment by an ISDN
ETS 300 099	ETSI NA 2	Specification of the Packet Handler Access Point Interface (PHI)
Draft RFHI	ETSI NA 2	Specifications of the Remote Frame Handler Interface
ETSI/STC NA2(93)07	ETSI NA 2	Frame Relay Service – General description
ETSI/STC NA2(93)08	ETSI NA 2	ISDN – Frame Relay Bearer Service – Service Definition
ETSI/STC NA2(93)09	ETSI NA 2	Frame Relay Data Transmission Service – Service Definition
Draft ETS DE/NA-53204	ETSI NA 5	Support of Frame Relaying Bearer Service in B-ISDN and Frame Relay interworking between B-ISDN and other networks
Draft ETS DE/NA-53207	ETSI NA5	B-ISDN - Frame Relay Bearer Service – Service Definition

The first deliverable with definition and models of different scenarios for Network-to-Network Interface will be finalised early 1994. The project will end in June 1994.

P301 focus their work on Network-to-Network interface on signalling, addressing, management, service parameters, and QOS.

The partners in P301 are Norwegian Telecom, Helsinki Telephone Company, Telia Research, and CSELT.

ITU-T

SG 11 is responsible for protocol specifications for Frame Relay, SG 13 (former XVIII) is responsible for Frame Relay service specifications and SG 7 is responsible for data communication networks

SG 7 issued a revised X.25 recommendation, and X.25 is now standardised with access speed 2 Mbit/s. Facsimile transmission over PSPDN (Public Switched Packet Data Network) is specified in recommendations X.5, X.38 and X.39. SG 7 is also responsible for interworking, Frame Relay offered in data networks and multicast service.

SG 7 is standardising interfaces and procedures applicable for the purpose of providing a gateway function in a PSPDN to a private PSDN (Public Switched Data Network), which is an interconnect mechanism different from X.75. Related to the Frame Relay service offered in dedicated networks FRDTS (Frame Relay Data Transmission Service), SG 7 will study and make specifications necessary for interworking arrangements, Network-to-Network Interface and User-to-Network Interface. This will also include signalling for FRDTS.

In SG 7 the following recommendations or standards are available:

- Draft recommendation X.35 'Interface between a PSPDN and a private PSPDN which is based on X.25 procedure and enhancements to define gateway functions that is provided in PSPDN'
- Recommendation X.6 'Multicast service definition'.

Further work in SG 7 will be to develop the necessary recommendations to support the set of multicast services specified in Recommendation X.6 and to further study interworking cases specific to public data networks.

X.35 will probably be approved in February 1994, and the Frame Relay Internetworking specifications (this will be covered in X.7x series) are planned to be approved in November 1994.

X.3x covering interworking between Frame Relaying Bearer Service/Frame Relaying Data Transfer Service and X.25/X.31 is planned to be approved in November 1994. It will have the same structure as X.31.

The working on Frame Relay User Network Interface covering both semipermanent and switched access is planned to be approved in November 1994. X.31 is planned to be revised in

November 1994 with the addition of semi-permanent H-channel access capabilities.

In SG 11 the protocol specifications for Frame Relay, Q.922 and Q.933, were finished in 1992 and 1993.

Concerning the User-to-Network Interface for ISDN – FMBS (Frame Mode Bearer Services) the following results are available:

- Q.922 'ISDN Data Link Layer for Frame Relay'
- Q.933 'Digital Subscriber Signalling System No. 1 (DSS1) Signalling Specifications for Frame Mode Bearer Service'.

Recently, SG 11 started to work on a Q.933 based signalling for a Frame Relay Network-to-Network Interface.

SG 13 has been working with the Frame Relay Bearer Service in ISDN, and SG 13 has made the service specification as well as specifications on areas related to this service: Congestion management, Network-to-Network Interface and interworking. SG 13 has also been working with Frame Relay in B-ISDN and has made a recommendation, I.365.1, that specifies the Frame Relaying Service Specific Convergence Sublayer.

Specifications for Frame Mode Bearer Service in SG 13:

- I.370 'Congestion Management for the ISDN Frame Relaying Bearer Service'
- I.233.1 'ISDN Frame Relaying Bearer Service'
- I.372 'Frame Relaying Bearer Service Network-to-Network Interface requirements'
- I.555 'Frame Relaying Bearer Service interworking'
- I.365.1 'Frame Relaying Service Specific Convergence Sublayer'.

Table 3 Standards in the area of Frame Relay in SG 11 and SG 13

ID	Organisation	Name
Q.922	ITU SG 11	ISDN Data Link Layer Specification for Frame Mode Bearer Services
Q.933	ITU SG 11	Digital Subscriber Signalling System No. 1 (DSS 1). Signalling specification for Frame Mode Bearer Services.
I.233.1	ITU SG 13	ISDN Frame Relay Bearer Service
I.370	ITU SG 13	Congestion Management for the ISDN Frame Relaying Bearer Service
I.372	ITU SG 13	Frame Relaying Bearer Service Network-to-Network Interface requirements
I.555	ITU SG 13	Frame Relay Bearer Service interworking

Teletraffic and dimensioning

BY HARALD PETERSEN

006
621.39

Introduction

The field of teletraffic, defined as the "Application of the Theory of Probability to Telecommunication Research, Engineering and Administration", is concerned with statistical description of the processes of arrivals and releases of demands for resources in a network, analysis of resource sharing systems under random traffic load, methods to structure and to size system and network resources to meet quality of service requirements, including methods for measurements and forecasting.

International activities in teletraffic research and standardisation are more intense than ever. The technology is changing, introducing higher flexibility in the networks, development of networks integrating services with a wide range of performance and bitrate requirements, mobility and intelligent network capabilities require development of new methods and standards on traffic control functions in the networks and dimensioning procedures. Focus in the research activities has since the late eighties been on teletraffic issues in ATM networks.

ATM networks have the flexibility to support various services and applications and offer transfer capacity adapted to the needs. This flexibility combined with relative few physical limitations necessitate traffic and congestion control functions within the network to protect the network, and hence the quality of service offered to the other users. Furthermore, variations in the traffic streams have to be utilised to optimise use of network resources in order to achieve low costs of communication.

The work going on in the following research and standardisation organisations is presented:

- International Telecommunication Union – Telecommunication Standardization Sector (ITU-T)
- European Telecommunications Standards Institute (ETSI)
- ATM Forum
- European Institute of Research and Strategic Studies in Telecommunication (EURESCOM)
- Research into Advanced Communications for Europe (RACE)
- European Co-operation in the Fields of Scientific and Technical Research (COST).

ITU-T

Within ITU-T, SG 2 has the general responsibility for traffic engineering. SG 13 is responsible for B-ISDN resource management including traffic and congestion control.

ITU-T SG 13

The main work within ITU-T on traffic control in B-ISDN has so far been done in SG 13, resulting in Recommendation I.371 "Traffic Control and Congestion Control in B-ISDN" being approved in March 1993.

In addition to the general framework with traffic descriptors and parameters, and definitions of various traffic and control functions, the first version of I.371 contains only mechanisms for peak cell rate handling. Enhancements of I.371 are necessary. A

second version dealing with the remaining issues related to peak cell rate definition and operation is scheduled to be in a frozen state in March 1994. The topics to be decided include separate or aggregate enforcement of user cells and Operation and Maintenance (OAM) cells, definition of peak cell rate and Cell Delay Variation (CDV) granularities in the ATM layer, and translation table between peak cell rate and peak emission interval. Furthermore, it is suggested to revisit the specification given in I.371 of traffic contracts with two different cell loss ratios using the Cell Loss Priority (CLP) bit, proposing to change to specifying the peak cell rate of the CLP = 0 and CLP = 1 streams separately.

The task of identifying additional multiplexing schemes and specification of relevant parameters and procedures is being studied. The parameters Sustainable Cell Rate and Burst Tolerance are being discussed. Frozen state of this new enhancement of I.371 is expected in 1995 at the latest.

ITU-T SG 2

SG 2 has four main areas of responsibility – routing, numbering, network management, and service quality of networks. The latter area includes traffic engineering, the responsibility of Working Party 2/3. This survey will only cover the traffic engineering activities.

N-ISDN: The E.700 series of Recommendations provide guidelines on traffic engineering for ISDN. It consists of four main groups of recommendations: ISDN traffic modelling (E.710 ...), ISDN traffic related network performance objectives (E.720 ...), ISDN dimensioning methods (E.730 ...) and ISDN traffic measurements (E.740 ...). This constitutes a common framework for both N-ISDN and B-ISDN. The existing traffic engineering recommendations for N-ISDN are incomplete, particular in the area of packet switching. Due to lack of reliable data the traditional quantitative methods have shortcomings in forecasting the demand for new services arising with ISDN, UPT and IN. New and revised recommendations in this area are scheduled for 1995 and 1996.

Traffic Engineering for B-ISDN: SG 2 started work on traffic engineering for B-ISDN at the end of 1991. Recommendations in the E.700 series will be produced on the following topics:

- Traffic characterisation
- Algorithms for traffic controls
- Dimensioning
- Traffic network performance
- Traffic measurements.

Both the call setup and release phase and the information transfer phase will be dealt with. Procedures and parameters related to both connection level and cell level will be elaborated. Guidelines on dimensioning of network elements will be produced, starting with dimensioning of links between network nodes. The impact of buffer sizes on link dimensioning will be considered.

Guidelines will be provided to the network operators on how to use traffic and congestion controls for the operation of the networks. Initially, the focus will be on the Connection Admission Control (CAC).

IN & SS 7: The deployment of IN based services will require the development of new traffic engineering methods and recommendations. Revised and new grade of service parameters, traf-

fic models, traffic and congestion control mechanisms, resource allocation and dimensioning methods for IN structured networks and associated SS No. 7 networks will be recommended in the new E.71N series and in revisions of E.723 and E.733. Signalling traffic issues for B-ISDN are included in the work-plan.

Mobile and UPT: Traffic engineering aspects of networks supporting mobile and UPT services will be covered in the E.750 series of Recommendations to be finalised in the 1994 – 96 time frame. The first recommendations in the series have already been published. The series addresses initially circuit-switched traffic and common channel signalling traffic, whereas packet-switched connections will be dealt with later on. The teletraffic engineering problems related to the interchange of messages required by UPT are addressed. Modelling of the traffic processes in the user and the control plane, grade of service requirements, dimensioning methods and measurements methods will be covered in the E.750 series.

Miscellaneous: Traffic engineering for Public Switched Telephone Networks (PSTN) requires continuing development to improve the efficiency of network operation. There is ongoing work on dimensioning circuit-groups with multi-slot bearer services. Traffic measurements necessary for SS No. 7, N-ISDN, B-ISDN, IN and UPT are addressed.

ETSI

The work on traffic control in B-ISDN within ETSI is done by the Traffic Control and Resource Management (TCR) group within NA5. This activity goes on in parallel with the corresponding activity within ITU-T SG 13, the same topics being discussed aiming at harmonised European positions.

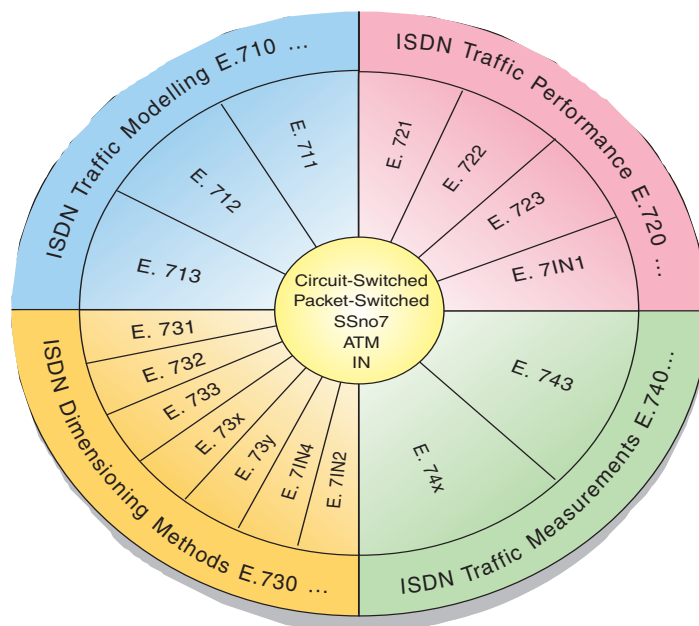
Release of the ETS “Requirements for traffic control and resource management of VP/VC cross-connected network” (VP = Virtual Path, VC = Virtual Channel) is scheduled for October 1994, and the ETS “Traffic Control and Congestion Control in B-ISDN” endorsing the second version of I.371 is expected to be released in December 1994.

ATM Forum

The ATM Forum is an organisation with the objective of accelerating the deployment of ATM products and services through rapid convergence of interoperability specifications and promotion of industry co-operation. The Traffic Management Group is responsible for traffic control and network performance issues. The results so far are included in the ATM Forum User-Network Interface 3.0 specification and the ATM Forum Broadband Interconnection Interface (B-ICI) 1.0. The specifications are based on ITU-T Recommendation I.371, however, on some points these specifications are more detailed. As an example, the parameters Sustainable Cell Rate and Burst Tolerance are defined within the ATM Forum UNI Specification, while they are still being discussed within ITU-T. Furthermore, reference traffic loads are defined for use when measuring the performance of a B-ICI link.

EURESCOM

There are traffic engineering activities in the fields of ATM and IN within the EURESCOM institute.



The E.700 series of Recommendations

Traffic issues in ATM VP cross-connect networks have been considered in the finalised project P105 “European ATM Network Studies”, resulting in deliverables on Connection Acceptance Control (CAC), Network Parameter Control (NPC) and Dimensioning. The CAC and NPC procedures specified are based on the declared peak cell rate, and are important input to the specification of the European ATM Pilot. Within the follow-up project P302 “European switched VC-based ATM Network Studies” (planned for the period September 1993 – November 1995) the activity on resource allocation is aiming at defining more source traffic descriptors allowing for statistical multiplexing gain in the network. Guidelines for implementation and evaluation of the corresponding Usage Parameter Control (UPC), NPC and CAC procedures will be prepared. Possible benefits with use of long buffers and corresponding buffer management schemes will also be studied.

Due to the planned development of IN and the rapid implementation of services, the need for appropriate methods and tools for dimensioning the IN infrastructure is recognised as very important. The objective of the project P308 “Methods for Tools to Dimension IN Networks”, planned for the period from September 93 to mid 95, is to determine general methods in order to give guidelines to the network operators on how to dimension the different elements of an IN network (including the SS no 7 network). The topics include traffic models, congestion and load control, flow control parameters, dimensioning methods and performance analysis. The quality of service aspects include both the delays of the call set-up phase and the probability of query rejection due to congestion in IN network elements.

RACE

The programme for Research and technology development in advanced communications in Europe, the RACE programme, is divided into two main phases, Phase I from 1988 to 1992 and Phase II from 1992 to 1995. There are a lot of projects within the RACE programme having activities on traffic engineering issues. In this survey we will focus on the main projects which, to the author’s knowledge, have influence on the development of the ATM traffic field.

R1022 ATD: “Technology for ATD” was one of the largest projects within the RACE I programme with ALCATEL as prime

contractor. The project had a big traffic group of experts from network operators, industry and universities. A high level of collaboration on ATM traffic issues was established having a strong influence on the development of this field. A lot of technical reports were produced dealing with ATM traffic source models, comparison of buffering schemes in switching blocks, performance evaluation of switching block architectures and priority disciplines, network performance studies, analysis of statistical multiplexing gain, development and evaluation of UPC and CAC mechanisms, and preparations of performance tests on the ATM testbed (RATT) developed in the project. The traffic studies have in some areas supported design choices made in the development of the equipment.

R2061 EXPLOIT: "Exploitation of an ATM Technology Testbed for Broadband Experiments and Applications". The aims of EXPLOIT are to provide a platform for pilot applications and broadband experiments, interconnection to other broadband platforms through development of mappers and interworking units, and performing traffic experiments to validate traffic models and dimensioning methods in the areas of traffic control, resource management and network performance analysis. EXPLOIT is a follow-up project in RACE II of the R1022 project with mainly the same partners, and the EXPLOIT testbed is based on the RATT. The main objectives for the traffic group have shifted from mostly theoretical work within R1022 to performing experiments to validate and refine the methods developed. The project duration is from 1992 to 1994, with a possible extension into 1995.

R1012 BLNT: "Broadband Local Network Technology", a project within the RACE I programme run mainly by Siemens with the objective to develop a local switch. The development was supported by studies on traffic control and performance aspects, like analysis of multi-stage switching networks. The control functions implemented were influenced by the studies. The implementation of a Spacing Policier was developed where a leaky bucket based peak rate policing is combined with a spacing function.

R1014 ATMOSPHERIC was the third switching system project within RACE I, with Ericsson as main partner. ATM traffic issues were also part of this project, dealing with techniques for network resource allocation and management. The results were used as input to the building block specifications.

R1083 PARASOL: "ATM Specific Measurement Equipment". ATM traffic has to be described by its behaviour on several levels (connection, burst, and cell level) which differ in time scale by many orders of magnitude. This complexity, together with the strict requirements on cell loss ratio in ATM networks, makes traditional simulation techniques insufficient in analysing ATM network traffic. Development of ATM measurement equipment is therefore of major importance for validation and testing of the performance and the traffic characteristics of ATM systems and control functions. The main objective of the PARASOL project has been to develop an ATM specific measurement equipment consisting of both an ATM traffic generator and an ATM traffic analyser, providing functions such as measurements of cell loss, cell delay and cell delay variation. A flexible traffic model used

within the traffic generator was developed to allow specification of traffic load to be as realistic as possible.

R2024 BAF: "Broadband Access Facilities" is an ongoing RACE II project. The objective of the project is to contribute to the development of broadband access facilities at conceptual, technical and application level for small businesses and residential subscribers. The project will realise an access network based on shared fibre technology using a passive optical network (PON) with a tree structure. One of the main issues is the traffic performance activity dealing with bandwidth management and traffic concentration aspects. A Medium Access Control (MAC) protocol for the PON will be developed and analysed in the project. The suitability for all possible B-ISDN services is addressed, in particular Constant Bit Rate (CBR) services with stringent requirements on cell delay and cell delay variation.

R2032 COMBINE: "Composite Broadband Interworking and End-to-End Models" is a RACE II project investigating the interworking of ATM based networks with other networks. Two DQDB – ATM interworking units and a connectionless server will be developed and attached to the EXPLOIT testbed mentioned above. Protocols and networks will be modelled and analysed to predict the end-to-end performance in heterogeneous network environments, using both analytical and simulation models. Traffic experiments related to interworking will be performed to assess the performance and functionality of the interworking units and to validate the simulation results obtained.

R2060 CIO: "Co-ordination, Implementation and Operation of Multimedia Services" is a RACE II project with the main objective to specify and implement two advanced multimedia tele-services as prototypes on various end-systems. The end-systems will be integrated in selected advanced communication experiments. The services will be tested and demonstrated in a heterogeneous network environment. A subtask is devoted to performance analysis and measurements of transport protocols in order to obtain Quality of Service (QOS) predictability and tune the protocol parameters to improve performance. The results will be utilised to guide the resource reservation schemes.

COST

The objectives of the COST 242 project "Methods for Performance Evaluation and Design of Multiservice Broadband Networks" are to co-ordinate and advance European research in the field, identify performance and network design problems, and provide solutions by mathematical models, computer tools or optimisation algorithms. The project is a follow-up of the COST 224 project completed in 1991, is gathering researchers from all over Europe, and has significant influence on the development of traffic issues in ATM networks. The COST 224 ended up with a book published: "Performance evaluation and design of multiservice networks". Within COST 242 queuing models for end to end performance are being developed. Traffic control strategies are investigated with focus on VBR connections and the relationships between traffic descriptors, connection admission control and parameter enforcement. Network design, routing strategies, network architecture and dimensioning are being considered. High speed LAN, MAN and interconnection of heterogeneous networks are evaluated.

Terminal equipment and user aspects

BY TROND ULSETH

Introduction

When reporting standardisation and research activities in the area of terminal equipment and user aspects, a definition of this area is required. In the first issue of this status section¹ 14 study areas were identified. The reports under the heading terminal equipment and user aspects will present status for work on terminals and applications for attachment to fixed networks, i.e.

- Public Switched Networks (PSTN)
- Integrated Services Digital Networks (ISDN)
- Broadband Networks (B-ISDN).

A key word in the area of terminal equipment today is multimedia. Multimedia is not a well defined term. A simple definition is that it involves at least two different information types, e.g. text, graphics, still picture, audio and video (moving image). In simple words, the research and standardisation activities on terminal equipment cannot focus on a single medium only, but has to take into consideration the fact that two or more media might be used in one application.

Today, most of the suppliers of computers/PC offer multimedia packages. These are standalone applications. The research and standardisation activities reported focus on the communications aspects of multimedia, being user-to-user communication or user-to-host/server communication.

The standardisation activities mainly focus on ISDN as the communication systems, but a lot of results are relevant for broadband systems. However, terminals for attachment to the Public Switched Telephone Network (PSTN) are still the largest market, and there are standardisation activities in this area too.

Most of the research activities focus on

- Terminal equipment for broadband applications
- Terminals for people with special needs (disabled, elderly people, etc.).

To avoid overlap with other reports, the presentation under this heading will not include

- Radio communication (mobile) user equipment
- Terminal equipment supporting Message Handling Systems (MHS)
- Signal processing
- Service definitions.

ETSI

In ETSI the Technical Committee Terminal Equipment (TC TE) is responsible for standardisation of terminal equipment, whereas the Technical Committee Human Factors (TC HF) is working on user aspects.

At present 8 STCs are reporting to TC TE. Two of these (TE3 and TE6) are working on MHS and directory systems. The works of these STCs are covered by another presentation.

ETSI Technical Assembly has decided to organise work in some areas into projects. So far, 4 projects are established, one of these is Multimedia. The responsibility for this project is given to TC TE. A Project Manager (Mrs A Hernandez, Telefonica, Spain) is appointed. To assist TC TE and the Project Manager, STC TE10 (Multimedia Planning and Co-ordination) has been created. Mr B Schuurink, Royal PTT Telecom, The Netherlands has been elected chairman of this STC. It is important to note that STC TE10 is responsible for planning and co-ordination, while the standardisation work shall be carried out in the existing STCs. However, where no STC is willing to carry out the work, STC TE10 may establish a work item.

So far the ETSI Multimedia Project Manager has analysed the present ETSI Work Programme. A total of more than 200 work items have been identified as relevant for multimedia standardisation. Most of these belong to TC TE. The multimedia project is therefore the largest of the ETSI projects.

For multimedia standardisation activities STC TE1, STC TE2 and STC TE4 are the most important STCs within TC TE. Of these, only STC TE1 has established work items addressing multimedia so far. Most of these work items are on multimedia and hypermedia² retrieval services, and have been established after an analysis made by Project Team 28V. The report of this Project Team is published as ETR 084 Multimedia & Hypermedia Information Retrieval service (M&HIRS) Investigation of candidate architectures for M&HIRS.

STC TE4, which is responsible for standardisation of Voice and Audiovisual Terminals, has a number of work items in the

Table 1 ETSI standards on voice and audiovisual terminals

ID	Name
ETS 300 143	Integrated Services Digital Network (ISDN); Audiovisual Services, Inband Signalling Procedures for Audiovisual Terminals Using Digital Channels up to 2048 kbit/s
ETS 300 144	Integrated Services Digital Network (ISDN); Audiovisual Services, Frame Structure For a 64 to 1920 kbit/s Channel and Associated Syntax for Inband Signalling
ETS 300 145	Integrated Services Digital Network (ISDN); Audiovisual Services, Videotelephone Systems and Terminal Equipment Operating on One or Two 64 kbit/s Channels

¹ *Endre Skolt: International research and standardisation activities in telecommunication. Teletronikk 89(4), 113-127, 1993.*

² *An explanation of the term Hypermedia can be found in Liestøl, G, The digital video album, Teletronikk 89(4), 31-35, 1993.*

Table 2 ITU Recommendations for conference control

ID	Name	Status
T. 120	Introduction to the T series audiovisual recommendations	Under study
T. 121	Audio Graphics Conferencing (AGC)	Under study
T. 122	Multipoint Communication Service for audio graphic and audio visual conferencing	Approved
T. 123	Protocol stacks for audiographic and audiovisual teleconference applications	Approved
T. 124	Generic conference control for audiovisual services	Under study
T. 125	Multipoint communication service protocol specification	Approved by SG 8 to be sent on ballot

audiovisual area that will be important for multimedia applications. The following standards have been approved:

Other standards covering the audio aspects, and applications such as

- audiographic conferencing in the ISDN
- videoconferencing in the ISDN
- Multipoint Control Units (MCU)

are under study.

For multimedia standardisation there is a need to identify a common platform. The ITU-T T.120 series recommendations could be this platform because they take into consideration both audiovisual applications and provide protocols for telematic applications and data communication. An introduction to these recommendations can be found in the ITU-T part of this presentation. This topic will be on the agenda for the next meetings of the ETSI committees involved.

Another topic of great commercial interest is standards for regulatory purposes. Directive 91/263/EEC regulate the rules for type approval of telecommunications terminal equipment in Europe. The type approval is based on Common Technical Regulations (CTR). ETSI is responsible for providing the technical base for CTRs, TBRs (Technical Base for Regulations). TBR 8 (Terminal requirements for ISDN 3,1 kHz telephony terminals) has been on Public Enquiry, and is now ready for national vote. Norwegian Telecom Research has provided the rapporteur for TBR 8.

The access TBRs for ISDN Basic Rate Interface (TBR 3) and ISDN Primary Rate Interface (TBR 4) will be ready for submission on Public Enquiry by end of March 1994. The European Commission is considering in interim arrangement until these TBRs are approved.

There are three STCs reporting to TC HF. STC HF1 is working on Human Factor issues in general, while STC HF2 is working on issues related to people with special needs. STC HF3 is working on usability evaluation.

Most of the results of TC HF are technical reports or guidelines. Reports on human factor aspects of Multimedia telecommunication (DTR/HF 01016), and on end-user procedures for ISDN videotelephony (DTR/HF 01018) are under preparation. A draft standard on pictograms for videotelephony (prETS 300 375) is on Public Enquiry.

ITU-T

Within ITU-T, several Study Groups are working on terminal equipment and user aspects. These are:

- SG 1 (Service definitions) on some user aspects (Human Factors)
- SG 8 (Terminal for telematic services) on terminals and protocols for telematic services
- SG 12 (End-to-end transmission performance of networks and terminals) on speech transmission characteristics and quality aspects moving image communication
- SG 14 (Modems and transmission techniques for data, telegraph and telematic services)
- SG 15 (Transmission systems and equipment) on system aspects/terminals for audiovisual communication.

SG 8 is probably the most important study group in ITU-T for standardisation of terminals and applications. One important activity is the work on a set of recommendations for conference control (T.120 series of recommendations). The work covers both point-to-point conferences and multipoint conferences. These recommendations are:

The structure of this set of recommendations is as shown in Figure 1.

The following application recommendations are planned so far:

- T.SI Still image for the conference (JPEG)
- T.AVC High Level Audio and Video Control
- T.BWC Bandwidth Arbitration and Control
- T.TDC Transparent User Data Channel
- T.CFAX Multipoint Facsimile for AVC and AGC³
- T.RESV Reservations for AVC and AGC
- T.PRO Application Profiles for AVC and AGC
- T.CBFT Multipoint Binary File Transfer.

EURESCOM

The owners of EURESCOM are the Public Network Operators in Europe. It is therefore logical that the main activities of EURESCOM are on networks. However, one of the study areas of EURESCOM is Services. Terminal equipment is not addressed, but user aspects are considered in some of the projects.

³ AVC and AGC – Audiovisual Conference and Audiographic Conference.

Terminals are essential for the success of new applications and services. This is recognised by EURESCOM. The present project on strategic studies (P311) will therefore analyse the need for increased EURESCOM activities in the area terminals and applications.

There is also COST work on telecommunications services for disabled people (COST 219 and COST 220). The objectives are to stimulate activities in the field, and to harmonise several systems for communication between hearing and speech impaired persons. Liaison between COST 220 and ETSI TC TE to analyse the need for standardisation in this area is established.

COST

COST 237 perform research on multimedia techniques and their application to distributed information services based on existing telecommunication networks and services. The results will be demonstrated in a selected application. COST 237 will be terminated in February 1997. Another COST project of relevance is COST 211 on coding of video signals in multimedia services.

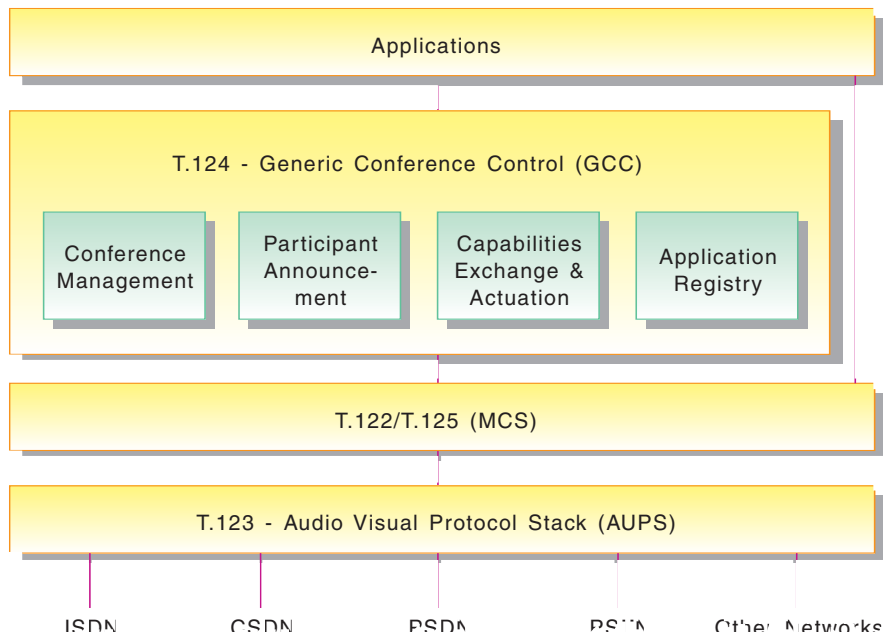


Figure 1 Structure of ITU T.120 recommendation series

Telecommunications Management Network

BY STÅLE WOLLAND

006
654.01:65

Introduction

Telecommunications Management Network (TMN) is an architectural concept for communication and processing of management information for tele- and data communications networks. The objective is to be able to implement management capability over a distributed management network (logically and/ or physically) in a standardised way so as to be system and vendor independent. Interoperability and re-use of management resources are also important objectives.

This summary presents some of the work going on in ISO/IEC. Future presentations will address some of the other work going on in ITU-T, ETSI, Network Management Forum and RACE. There is also management work going on in ANSI, ECMA, IEEE, TINA-C and other bodies.

ISO/JTC

The standardisation of Open Systems Interconnection (OSI) Management is performed within JTC1/WG21/WG4. The work was originally directed at management of Open System Interconnection resources.

The applicability of the work to areas such as telecommunications services has long been recognised and there is now harmonisation and co-ordination between the two areas ISO/IEC OSI Management and ETSI / ITU-T TMN. This harmonisation is the responsibility of ITU-T SG VII.

There are now International Standards in 17 management areas (in addition to the ISO/IEC designation the corresponding ITU-T Recommendation is also indicated):

General

- OSI Reference Model – Part 4: Management Framework (ISO/IEC 7498-4:1989 | CCITT Rec. X.700)

OSI Management Information Service

- Systems Management Overview (ISO/IEC 10040:1992 | ITU-T Rec. X.701)

Structure of Management Information

- Structure of Management Information – Part 1: Management Information Model (ISO/IEC 10165-1:1993 | CCITT Rec. X.720)
- Structure of Management Information – Part 2: Definition of Management Information (10165-2:1992 | ITU-T Rec. X.721)
- Structure of Management Information – Part 4: Guidelines for the Definition of Managed Objects (ISO/IEC 10165-4:1992 | ITU-T Rec. X.722)

Management Communications

- Common Management Information Service (Incorporating Amendments 1 & 2)(ISO/IEC 9595:1991 | ITU-T Rec. X.710)
- Common Management Information Service – Amendment 4: Access Control (ISO/IEC 9595:1991/AMD4:1992 | ITU-T ?)
- Common Management Information Protocol – Part 1: Specification (Incorporating Amendments 1 & 2) (ISO/IEC 9596:1991 | ITU-T Rec. X.711)
- Protocol Implementation Conformance Statement Proforma (IS 9596-2 | X.712).

Fault Management

- Systems Management – Part 4: Alarm Reporting Function (ISO/IEC 10164-4: 1992 | ITU-T Rec. X.733)
- Systems Management – Part 5: Event Report Management Function (ISO/IEC 10164-5: 1993 | ITU-T Rec. X.734)
- Systems Management – Part 6: Log Control Function (ISO/IEC 10164-6 | ITU-T Rec. X.735)

Security Management

- Systems Management – Part 7: Security Alarm Reporting Function (ISO/IEC 10164-7: 1992 | ITU-T Rec. X.736)
- Systems Management – Part 8: Security Audit Trail Function (ISO/IEC 10164-8: 1993 | ITU-T Rec. X.740)

Accounting Management

None.

Performance Management

None.

Configuration Management

- Systems Management – Part 1: Object Management Function (ISO/IEC 10164-1: 1993 | ITU-T Rec. X.730)
- Systems Management – Part 2: State Management Function (ISO/IEC 10164-2: 1993 | ITU-T Rec. X.731)
- Systems Management – Part 3: Attributes for Representing Relationships (ISO/IEC 10164-3: 1993 | ITU-T Rec. X.732)

There is work going on in a large number of other areas. The state of progress varies from Working Draft (WD) through Committee Draft (CD) to Draft International Standard (DIS).

Areas of ongoing work include the following:

(Extended) Systems Management (SM) Architecture

Extended SM Architecture specifies additional models to Management Framework and SM Overview. Current major topics are Management Domains and Management Knowledge Management.

Management Domains Architecture (WD nn | X.nnn), Systems Management Overview – Amendment 1: Management Knowledge Management.

SM Information

In this area there is ongoing work on:

Generic Management Information (DIS 10165-5 | X.723), General Relationship Model (CD10165-7 | X.nnn), Requirements and Guidelines for ICS Proformas (DIS 10165-6 | X.724).

SM Communications

In this area there is ongoing work on:

Application Context for System Management with Transaction Processing (CD 11587.2 | X.nnn), Extended Functionality for System Management Communication (WD nn-nn | X.nnn), In addition there is preparatory work on a WD on Managed Objects for Upper Layers.

SM Functions

In this area there is ongoing work on:

Management Domains Management Function (nn-nn | X.749), Management Knowledge Management Function (CD 10164-16 | X.750), General Relationship Management Function (10165-nn | X.747), Change Over Function (WD10165-17 | X.751), Test Management Function (DIS 10164-12 | X.745), Confidence and Diagnostic Test Categories (DIS 10164-14 | X.737), Objects and Attributes for Access Control (DIS 10164-9 | X.741), Usage Metering Function (DIS 10164-10.2 | X.742), Metric Objects and Attributes (DIS 10164-11 | X.739), Summarization Function (DIS 10164-13 | X.738), Response Time Monitoring Function (nn-nn | X.748), Scheduling Function (DIS 10164-15 | X.746), Software Management Function (WD X | X.744), Time Management Function (WD nn-nn | X.743), Command Sequencer for System Management (nn-nn | X.nnn).

A presentation of EURESCOM

BY ENDRE SKOLT

006

Introduction

EURESCOM is an acronym for European Institute for Research and Strategic studies in Telecommunications and is a collaborative research and development institute jointly owned by 26 public network operators (PNOs) in Europe and established in March 1991.

In order to meet the future challenges in the liberalised telecommunication environment, the traditional PNOs have joined forces to improve their efficiency and productivity. One reason for establishing EURESCOM is the need for interoperability of national services to enable provision of pan-European services. New services developed and implemented jointly will obviously save costs. Liberalisation will generate a large number of competing service providers and network operators, and PNOs will learn when to compete and when to co-operate. Being part of a research community dealing with short and medium term projects will give valuable experience.

The PNOs spend large sums of money every year on research and development, and an extended co-operation would therefore benefit all, in particular the small and medium sized companies. In 1993 Norwegian Telecom's membership expenses were approximately 5 million NOK. The EURESCOM budget was 230 million NOK. Most important are the strategic advantages. Norwegian Telecom participates in a large research community with access to the latest trends in telecommunication technology and systems, and the possibility to influence the direction of research.

In this paper we will present the objectives of EURESCOM, outline the organisational structure and give examples of results that have been achieved so far. The work programme for 1994–98 will also be introduced.

Objectives

The overall objective of EURESCOM is to support the development and provision of harmonised pan-European public fixed telecommunications networks and services and to show the importance and usefulness of this provisioning. The objective comprises

- Enabling the development of harmonised strategies for the planning and the provisioning of future European public networks and services
- Stimulating and carrying out short term, medium term and long term research projects
- Stimulating and co-ordinating field trials
- Contributing to European and world-wide standardisation.

In many ways the PNOs will be competitors, and that may be seen as an obstacle to the co-operation within the framework of EURESCOM. However, EURESCOM will act in accordance with the following principles:

- The activities of EURESCOM will not hinder or restrict in any way competition among its members
- The institute will not speak on behalf of its members unless asked to do so
- The members are not bound by the recommendations of the institute.

The organisational structure

EURESCOM is an institute having the legal status of a German GmbH (limited liability company) with its permanent seat in Heidelberg, Germany. Based on the general structure of German companies, EURESCOM has the following bodies:

- A General Assembly (GA), in which all members of the GmbH are represented and exercise their voting rights in accordance with the established rules
- A Board of Governors (BoG) as a permanent management and supervisory body. The BoG of EURESCOM has nine members, elected by the GA. The current board is chaired by Mr F Thabard of France Telecom and members from Deutsche Telekom, BT, STET, Telefonica, PTT NL, Telecom Finland, Telecom Portugal and TeleDanmark.
- The GmbH staff, headed by a director, comprises 20 members and is located at the EURESCOM headquarters in Heidelberg.

The structure of EURESCOM is illustrated in Figure 1.

Members

All European PNOs providing national and/or international public fixed telecommunications networks that have special or exclusive rights for the provisioning of network infrastructures or basic services may become members of EURESCOM. All members must belong to countries within the geographical area

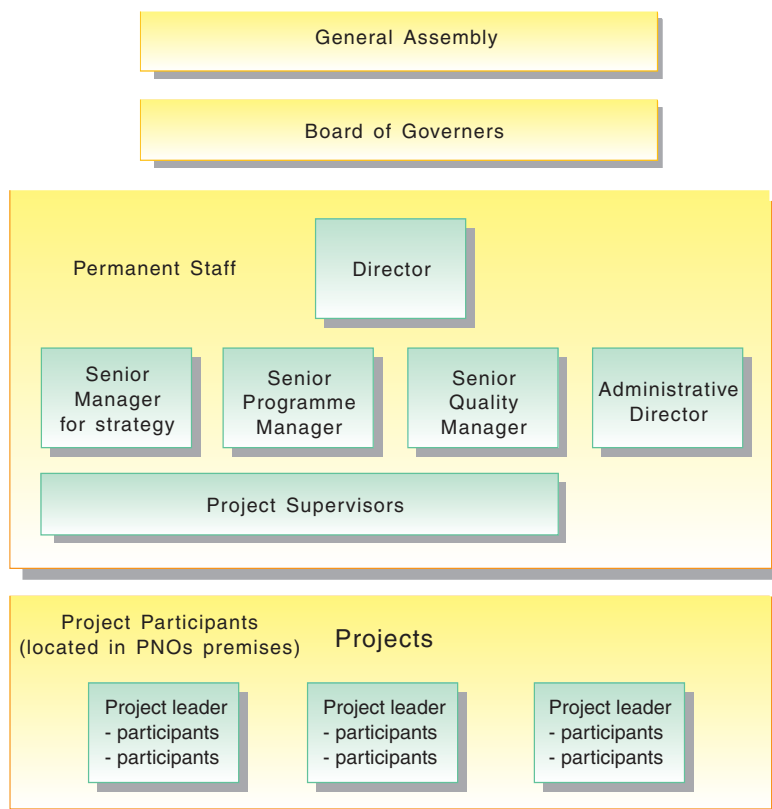


Figure 1 Organisational structure of EURESCOM

of CEPT, and they must have strong interests in ensuring integration of the European network infrastructure, interoperability of services across Europe and commit themselves to support advanced collaborative research and development activities. Most of the members belong to the category of “old” public network operators, however, newcomers like Mercury, UK have joined.

The European PNOs which are currently members of EURESCOM are shown in Table 1.

Results achieved

Since EURESCOM was established three work programmes have been launched resulting in about 50 projects. Although EURESCOM is still in the initial phase, the majority of these projects have brought significant results giving the owners valuable input to their development of European services and networks. The work programme is grouped into six work areas:

- 1 Strategic studies
- 2 Infrastructure and switched networks
- 3 Intelligent networks
- 4 Telecommunications Management Networks
- 5 Telecommunication services
- 6 Software requirements and practices

In the following we will highlight important results.

In the work area of telecommunication services major achievements are:

- Specification of a multimedia retrieval service, including the supporting network capabilities and terminal requirements
- An overview of the authentication methods applied in telecommunications, analysis of the pros and cons of these, and a guide for developers of services that need authentication
- Scenarios for the construction and exploitation of a pan-European directory service based on the X.500 recommendation.

In the area of Intelligent Networks (IN) the following achievements have been made:

- Specification of a pan-European Virtual Private Network (VPN) service
- Specification of a pre-Universal Personal Telecommunication (UPT) service
- Contribution to the definition of pan-European IN services
- Contributions to the development of future IN architectures.

In the Telecommunications Management Network (TMN) work area:

- TMN organisational model
- Management services definition and evolution
- TMN security requirements
- Contributions to the development of the TMN X-interface for the management of transmission and signalling systems.

The infrastructures and switched networks work area includes results as follows:

- Specifications of a pan-European Asynchronous Transfer Mode (ATM) Virtual Path handling system, intended for use by the ATM Pilot Group
- Technical recommendations for the METRAN Project, suggested for application during the METRAN implementation phase
- Specification of Abstract Test Suites for ISDN end-to-end testing, to be used not only by EURESCOM shareholder/parties, but possibly by manufacturers and non-European PNOs as well
- Exploration of optical network issues and reliability of existing and future optical networks.

Concerning the area of Software requirements and practices, two fact finding studies have been carried out:

- Common software practices, which concludes that EURESCOM must concentrate on areas where the PNOs can

Table 1 European PNOs which are members of EURESCOM

Company	Country
Association of Telephone Companies in Finland	Finland
Entreprise des Postes et Télécommunications	Luxembourg
Österreichische Post und Telegraphenverwaltung	Austria
British Telecommunication plc	United Kingdom
Swiss Telecom PTT	Switzerland
TeleDanmark A/S	Denmark
Deutsche Bundespost Telekom	Germany
France Telecom	France
Hungarian Telecommunications Company Ltd.	Hungary
General Directorate of Posts and Telecommunications	Iceland
STET Societa Finanziaria Telefonica p.a.	Italy
Telefones de Lisboa e Porto (TLP) S.A.	Portugal
Mercury Communications Ltd.	United Kingdom
Companhia Portuguesa Radio Marconi S.A.	Portugal
Koninklijke PTT Nederland N.V.	The Netherlands
Norwegian Telecom	Norway
Hellenic Telecommunications Organisation S.A.	Greece
BELGACOM	Belgium
Telia AB	Sweden
SPT TELECOM s.p.	Czech Republic
Slovak Telecommunications State Enterprise	Slovak Republic
Telecom Eireann	Ireland
Telefonica de Espana S.A.	Spain
Post and Telecommunications of Finland	Finland
Telecom Portugal S.A.	Portugal

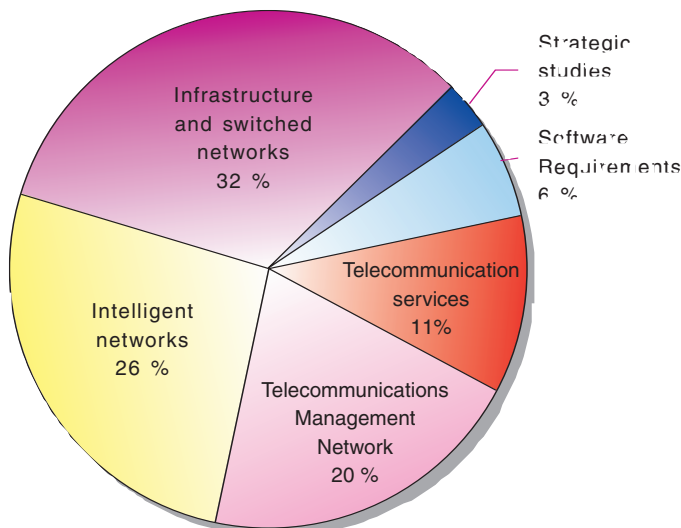


Figure 2 Distribution of manpower resources among work areas

have common interests, e.g. computing environments for TMN systems

- Requirements on databases for telecom services. It has been concluded that no existing products are customised for telecommunications. In the area of charging and mobility specific requirements are identified.

Work Programme 1994–98

The General Assembly of EURESCOM has recently approved a work programme for 1994–1998. It is stressed that this work programme shall reflect and serve the interest of the users, which encompass operational and support departments, commercial departments, corporate planning departments and research departments. Therefore, it has been important for EURESCOM to find a balanced approach in what type of projects should be defined and the resources spent. The work programme for 1994–1998 indicates the following project distribution in time:

- 1 Short term projects (0–3 years time frame using 40% of the manpower resources)
- 2 Medium term projects (3–5 years time frame using 40% of the manpower resources)
- 3 Long term projects (5–7 years time frame using 20% of the manpower resources).

The distribution of manpower among the six work areas is shown in Figure 2.

The following outlines new activities in the work programme for 1994–98:

- Strategic studies. In this work area there is only one new project, which aims to approve, maintain and progress the vision and scenarios for pan-European telecommunications. It will last until late 94. The plan is to allocate 100 man months per year for strategic studies for the period 1995–98.
- Telecommunication services. In this work area several new projects have been defined, such as “Definition of a pan-European IC card for authentication”, “A user centred approach to the specification of telecommunication services” and “Enhanced pan-European supplementary services and

broadband services for residential users”. It is decided to extend and continue the pan-European directory service project that has been running in the previous work programme.

- Intelligent networks (IN). In this work area only two new projects have been proposed. The first one is “Pre-study on service creation procedures and environment / requirements for pan-European IN-services for medium term needs” and the second is “Specification of tools to dimension IN networks”. There already exists four projects in the previous work programme that will be continued in 1994; “Evolution of IN”, “IN and TMN service testing”, “Enabling pan-European services by co-operation between PNOs’ IN platforms” and “Methods for tools to dimension IN networks”. It is stressed that in the 1994–98 period it is important to focus on management aspects of IN applying the TMN concepts which soon will constitute the operational basis for national networks. In addition, the development and provisioning of broadband and multimedia services should be reflected in the IN architecture.
- Telecommunications Management Network (TMN). New projects are “TMN management of pan-European Freephone service”, “TMN guidelines”, “Specification of TMN X-interface for charging, billing and accounting” and “Pan-European TMN- experiments and field trial support”. Many of the TMN-projects that have been running in the previous work programme have been finalised. However, “Management of IN-based services”, “TMN guidelines and information model” and “Security architecture” will be continued. All the new and continuing projects in this work area are aiming to support implementation of applications in demonstrators, and pilots and field trials.
- Infrastructures and switched networks. Several projects will be continued from the previous work programme: “Participation of EURESCOM in the METRAN project”, “B-ISDN-Signalling Project”, “Frame Relay Network-to Network Definition”, and “Reliability and Quality Measurement System”. New projects in this work area are “Protection, performance and synchronisation of pan-European SDH network”, “End-to-end testing for pan-European ATM network”, “Radio access technologies”, “Methodology and tools for ISDN network integration and traffic route testing” and “Optical networking”.
- Software requirements and practices. A new project in this work area is “Design of management service”, which is a continuation of the project “TMN operations system platform”. However, the focal point of this work area is expected to be in the projects “Software platforms and interfaces” and “Software quality assurance”, which both will be continued from the previous work programme.

References

- 1 General Framework for the EURESCOM Work Programme, April 1993.
- 2 EURESCOM Work Programme 1994–98 (Approved by the General Assembly 20 October 1993).
- 3 EURESCOM Internal Memo, Subject: Article on EURESCOM, 3 December 1993, Kevin Fogarty.
- 4 TF93/289-1 (TF-notat), Orientering om EURESCOM, 10 mars 1993, Helge Godø.