

## Sammendrag:

# Utvikling av regionale modeller for persontransport i Norge

## Innledning

Samferdselsmyndighetene i Norge har engasjert Transportøkonomisk institutt (TØI) og Møreforskning Molde AS (MFM) til å etablere et sett med landsdekkende kortdistanse transportmodeller som blant annet skal benyttes i arbeidet med nasjonal transportplan (NTP). Modellene skal dekke reiser inntil 10 mil én vei. Grunnlaget for arbeidet er blant annet en nasjonal reisevaneundersøkelse gjennomført i Norge i 2001 (RVU2001), og en rekke andre geografisk fordelte data fra samme tidsrom. RVU2001 og de øvrige data er stedfestet til grunnkretser som geografisk enhet. Det er vel 13500 slike grunnkretser i Norge. RVU2001 omfatter intervjuer med ca 21000 personer. I tillegg har PROSAM (samarbeidet for bedre transportprognoser i Osloregionen) i samme periode intervjuet 8800 personer bosatt i Oslo og Akershus fylke (PRVU01/02). Denne undersøkelsen er også benyttet i modell-estimeringen. Samferdselsmyndighetene har, i samarbeid med SINTEF, også utviklet et sett regionale nettverksmodeller som skal levere og motta data fra transportmodellene. Nettverksmodeller og nettutlegging er ikke omtalt i denne rapporten.

Arbeidsopplegget for prosjektet forutsatte at man skulle estimere felles modeller for hele landet, som videre skulle implementeres og tilpasses spesifikt i 5 landsdeler som tilsvarer Vegvesenets regioninndeling. Dette ble bestemt fordi datagrunnlaget i de to RVUer til sammen ikke var stort nok til å estimere egne modeller for hver av de 5 landsdelene, og fordi det dessuten er vesentlig mindre ressurskrevende å estimere et felles sett modeller. En ulempe er at man ved å estimere felles modeller må gi avkall på å håndtere forhold som er spesifikke for hver landsdel og som dessuten kan variere betydelig internt i hver landsdel (kø/reisetidspunkt, med mer). For å bedre ivareta behovet for samferdselsmyndighetenes daglige planlegging er programmeringen av modellene uavhengig av soneinndelingen. Dette gjør at en planlegger relativt lett kan definere og spesifisere mindre modellområder enn en region, for eksempel fylker eller grupper av kommuner, og kjøre modellene på disse. Dermed kan man unngå å kjøre modellen på en hel region med 5000-6000 soner hvis man skal studere et prosjekt eller en prosjektpakke med en klar geografisk avgrensning når det gjelder influensområde.

## Oversikt over modellsystemet

En modellkjøring vil vanligvis starte med ta man tar ut matriser med reisetilbud fra nettverksmodellene, "Level of Service" (LoS)-matriser. Dette er matriser med

---

Rapporten kan bestilles fra:

Transportøkonomisk institutt, Postboks 6110 Etterstad, 0602 Oslo

Telefon: 22 57 38 00 Telefax: 22 57 02 90

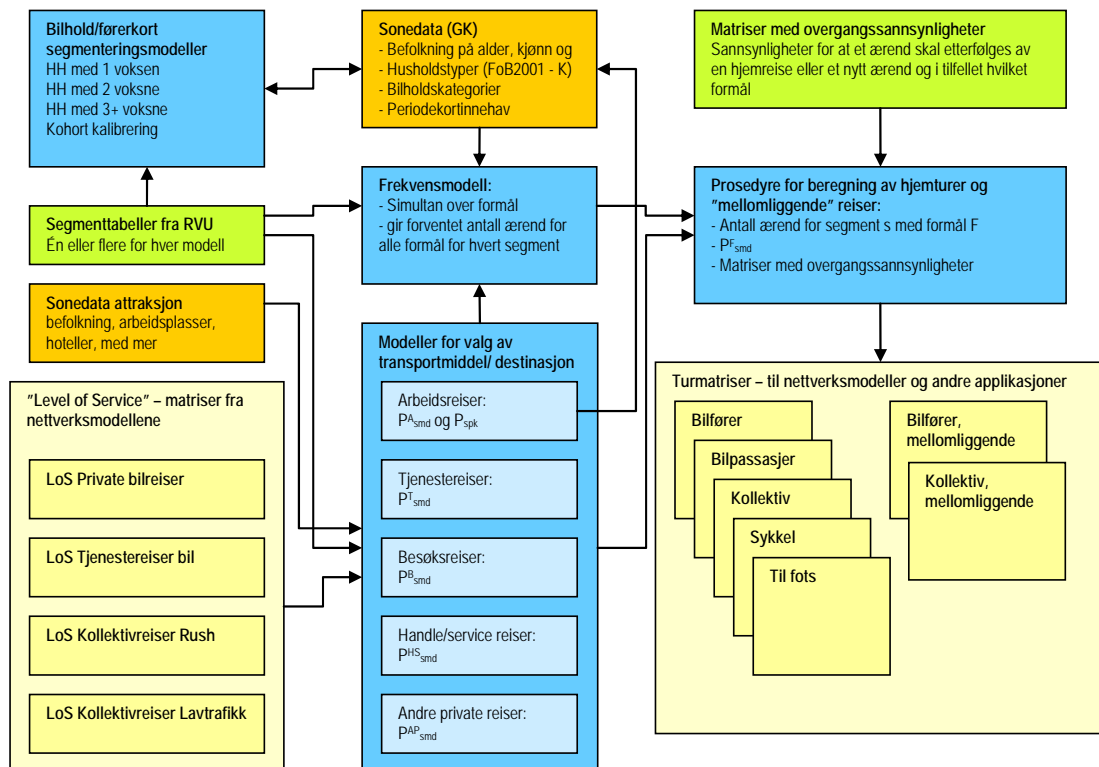
informasjon om reisetider og kostnader mellom alle soner i modellområdet og for alle transportmåter. Disse spesifiseres som inputdata sammen med de øvrige data som inngår i en modellkjøring; demografiske *sonedata til generering av turer*, *sonedata for attrahering av turer* og ulike *segmenttabeller* fra RVU2001, som gir gjennomsnittsverdier for befolkningssegmenter når det gjelder variable det ikke finnes sonedata for.

Modellkjøringen starter med *bilholds- og førerkortmodellene* som segmenterer de demografiske data (befolkning i grunnkretser på alder, kjønn og husholdstype) etter bilhold og førerkortinnhav. Det er egne separat estimerte bilholds- og førerkortmodeller for tre ulike husholdstyper, og befolkningen fordeles på 5 segmenter etter biltilgang, avhengig av om personen selv har førerkort og av hvor mange biler det er i husholdet i forhold til antall førerkort. Modellene er kalibrert mot prognoser for førerkortinnhavet i år 2010, 2015, 2020, 2025, 2030 og 2040. Bilholds- og førerkortmodellene er først og fremst viktige i langsiktige trafikk- og transportprognoser.

De 5 modellene for valg av transportmiddel og destinasjon (én for hvert reiseformål) er bosteds- og rundturbaserte og benytter segmenteringen fra bilholds- og førerkortmodellene som input. Valg av transportmiddel og destinasjon er estimert simultant, som multinomiske logitmodeller for de fleste reisehensikter. I modellen for *arbeidsreiser* ivaretas også valg mellom periodekort og enkeltbillett for kollektivtransport, og dette er formulert slik at informasjonen om periodekortinnhavet kan benyttes direkte av de 4 øvrige reisehensiktene.

*Modellene for reisefrekvens* er logitmodeller estimert i kombinasjon med Hurdle-Poisson. En standard Poisson-modell angir sannsynligheten for at en begivenhet skal opptre 0,1,2,3,... ganger i løpet av et gitt tidsrom. Hurdle-Poisson innebærer i dette tilfellet at sannsynligheten for 0 forekomster avviker fra det man har med standard Poisson. Modellene er estimert simultant for alle reisehensikter og gir forventet antall ærend for hver reisehensikt. I frekvensmodellen inngår bl a log-summer fra transportmiddel- og destinasjonsmodellene som variable, hvilket innebærer at reisefrekvensene påvirkes av transporttilbudet og biltilgjengeligheten. Siden modellene for transportmiddel og destinasjon er basert på rundturer med bare én hoveddestinasjon, mens frekvensmodellen beregner alle ærend som er gjennomført, er det laget en *prosedyre som ivaretar rundturer med mellomliggende reiser*. Prosedyren samordner informasjonen fra transportmiddel- og destinasjonsmodellene og frekvensmodellen slik at systemet produserer korrekt antall utreiser, hjemreiser og mellomliggende reiser, basert på matriser med "*overgangssannsynligheter*". Dette er sannsynligheter for at et ærend med et gitt formål etterfølges av en hjemreise eller et nytt ærend, og i tilfelle hvilket formål. Resultatet av en modellkjøring er turmatriser for rene tur/retur reiser og utreiser, mellomliggende reiser og returer for reiser med flere enn én destinasjon.

Figur 1. Oversikt over modellsystemet



TØI-rapport 766/2005

## Segmenteringsmodeller for bilhold og førerkortinnehav

De demografiske data tilgjengelig for modellsystemet fordeler befolkningen i grunnkretsene på alder (5 års intervaller), kjønn og 3 husholdskategorier (hushold med én voksen person fra 18 år, med to voksne personer og med tre eller flere voksne personer). Bilholdet og førerkortinnehavet i husholdene er et svært viktig aspekt ved individers muligheter til å foreta visse typer reiser, samtidig som utviklingen av biltilgjengeligheten over tid sannsynligvis er en avgjørende faktor for veksten i biltrafikken. Vi ønsker derfor at våre transportmodeller skal ivareta effekter av endret biltilgang og førerkortinnehav. Hensikten med disse modellene er dermed å segmentere befolkningen ytterligere for å få informasjon om befolkningsgruppens biltilgjengelighet. Vi har valgt å lage modeller som segmenterer befolkningen på alder og kjønn i sonene videre inn i 5 gjensidig utelukkende segmenter med ulik biltilgang. Resultatene fra disse modellene kan benyttes direkte i transportmiddel- og destinasjonsmodellene. De fem segmentene er:

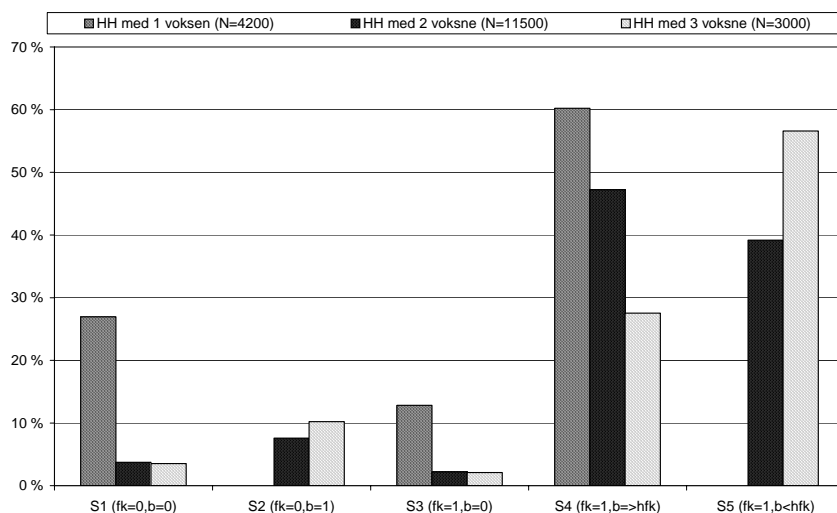
- S1: Personer uten førerkort og ingen biler i husholdet (ikke tilgang til bil som fører, dårlig tilgang til bil som passasjer).
- S2: Personer uten førerkort, men med en eller flere biler i husholdet (bare biltilgang som bilpassasjer).

- S3: Personer med førerkort, men uten biler i husholdet (dårlig tilgang til bil)
- S4: Personer med førerkort og minst like mange biler som førerkort i husholdet (full biltilgang)
- S5: Personer med førerkort og færre biler enn førerkort i husholdet (delvis/god biltilgang)

Som vi ser, har personer som tilhører segment S1 og S2 ikke førerkort, og personer som tilhører segment S1 og S3 har ikke bil. For personer med førerkort, er totalt antall førerkort og totalt antall biler i husstanden avgjørende for biltilgangen. Data fra RVU2001 er benyttet som estimeringsgrunnlag, sammen med data som beskriver respondentenes bostedsgrunnkrets. Materialet er splittet i tre deler avhengig av antall personer i husholdet som har fylt 18 år.

Figur 2 viser at fordelingen på de 5 bilhold/førerkortsegmentene varierer betydelig mellom de tre husholdstypene. I hushold med én voksen er andelen med full biltilgang 60 %, mens den er knappe 50 % i hushold med 2 voksne og knapt 30 % i hushold med 3 og flere voksne. I hushold med én voksen person er dessuten andelen uten biltilgang nær 30 % (eldre og yngre aleneboende). I hushold med 3 og flere voksne er andelen med delvis biltilgang nær 60 %.

Figur 2 . Fordeling av personer på bilholds- og førerkortsegmenter etter husholdstype

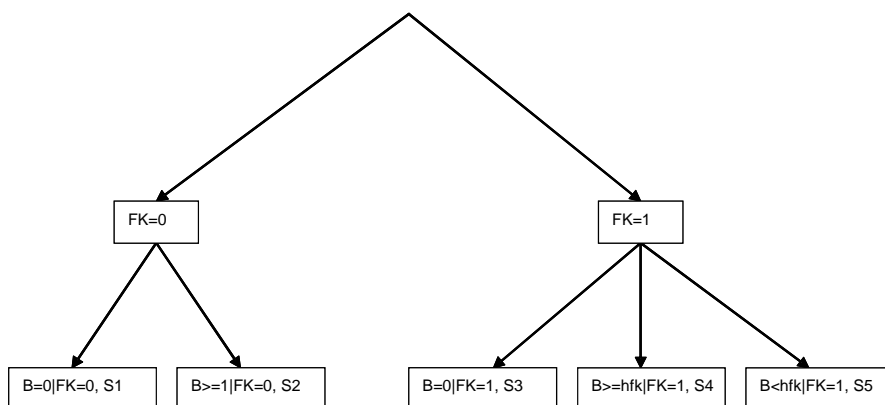


TØI-rapport 766/2005

Den metoden som er valgt innebærer å definere situasjonen som tre atskilte problemer; førerkortinnehavet, biltilgang for personer med førerkort, og biltilgang for personer uten førerkort. Samtidig er modellene for disse tre situasjonene estimert simultant. Førerkortinnehavet er en binær situasjon (fk=1 eller fk=0). Biltilgang for personer med førerkort representeres ved tre gjensidig utelukkende alternativer, ingen biler (b=0), bil men færre biler enn førerkort i husholdet (b>0, b<hfk), og bil og like mange eller flere biler enn førerkort i husholdet (b>0, b>=hfk). Når vi for personer uten førerkort ikke trenger å ta hensyn til antall førerkort og antall biler blir situasjonen som vist i figur 3. For hushold med bare én voksen person vil ikke S2 og S5 være aktuelle, da det i slike hushold ikke er naturlig å ha bil hvis man ikke har førerkort (S2), og fordi det maksimalt kan være én person med førerkort i slike hushold (S5). For hver av de tre husholdskate-

goriene er det estimert modeller som fordeler befolkningen på de 5 bilhold/fører-kortsegmenter.

Figur 3. Struktur i modeller for bilhold og førerkortinnehav



TØI-rapport 766/2005

Modellene er estimert med en sannsynlighetsmaksimeringsprosedyre kodet i GAUSS. Modellene er spesifisert med "nyttefunksjoner" som består av alternativspesifikke konstanter og variable som beskriver bostedssone, individuelle kjennetegn og husholdskarakteristika. De viktigste variablene er alder, kjønn, familietype, befolkningstetthet og inntekt. Alder og kjønn inngår både formulert som dummyvariable og som kontinuerlige variable.

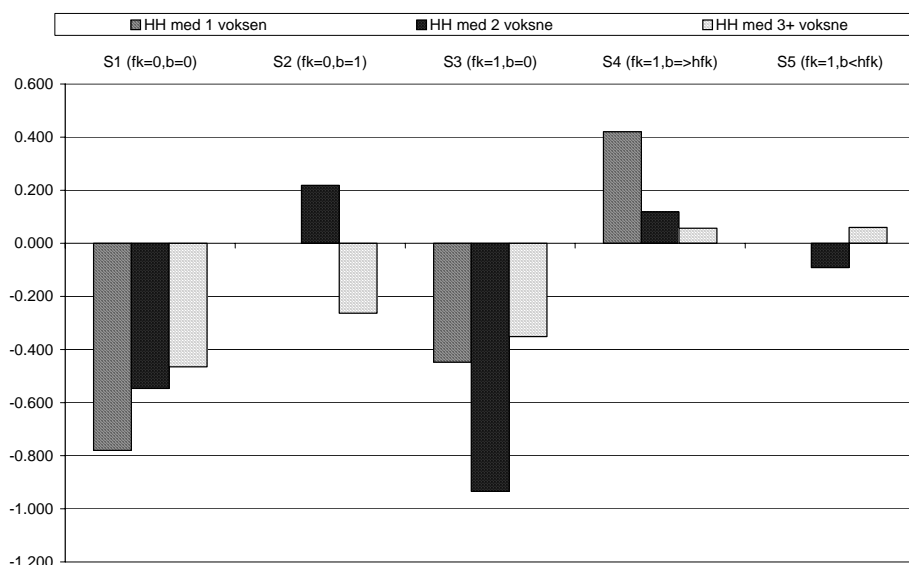
Vi har ikke hatt tilgang til variable som beskriver parkeringsforholdene ved bostedet direkte. Befolkningstetthet er benyttet som en proxyvariabel for dette, formulert både som kontinuerlige variable og som ulike dummyvariable. Tanken er at det i områder med høy befolkningstetthet ofte kan være høyere generaliserte kostnader knyttet til bilhold enn i områder med lav tetthet. Denne variabelen vil imidlertid også fange opp andre faktorer, som bedre kollektivtilbud, kortere avstander til aktiviteter, med mer, og dermed mindre behov for bil i områder med høy tetthet. Det er en klar sammenheng mellom befolkningstetthet og både bilhold og førerkortinnehav i våre modeller. I områder med høy befolkningstetthet er sannsynligheten for full biltilgang og førerkort lavere, og sannsynligheten for delvis eller ingen biltilgang høyere enn ellers.

Husholdsinntekt inngår i estimeringen som kontinuerlige variable. Det er forutsatt at den marginale bilen har en fast årlig kostnad på NOK 10 000, som trekkes fra husholdsinntekten. Den marginale bilen er ofte mye eldre enn gjennomsnittsbilen og har dermed lavere verdi. Eldre biler har sjelden kaskoforsikring. I alle modeller har inntektsvariabelen et tillegg for storbyområder (de 4 største byene i Norge), som gjør at inntektseffekter dempes i disse områdene.

Figur 4 viser inntektselastisitetene i de tre modellene beregnet med utgangspunkt i den fordeling en hadde på hushold og segmenter i datagrunnlaget for estimeringen. Det som vises i figuren er prosentvis forskyvning mellom segmentene som følge av 1 % økning i inntektene. Størrelsen på endringene er svært avhengig av

fordelingen på segmentene i utgangspunktet (jfr fig 2). Det er et generelt trekk at en finner store endringer for alternativer med få personer og mindre endringer for alternativer med mange personer.

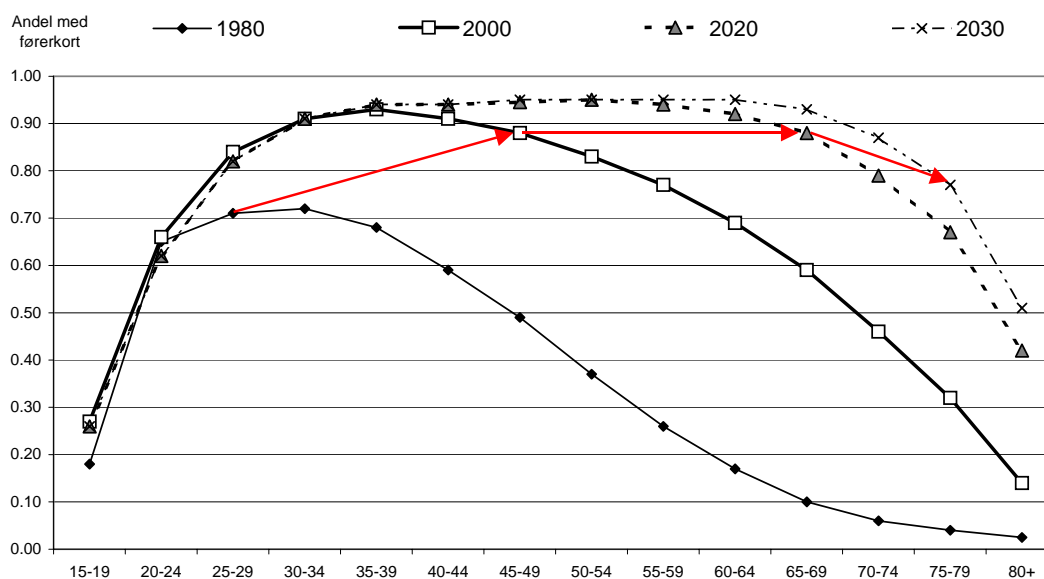
Figur 4. Inntektsvirkninger i bilholds- og førerkortmodellene. Prosent endring i segmentets størrelse ved 1% alminnelig inntektsvekst.



TØI-rapport 766/2005

I implementeringen kalibreres modellene mot prognoser for førerkortinnhavedet i år 2010, 2015, 2020, 2025, 2030 og 2040 beregnet ved kohorteffekter, dvs hvordan førerkortinnhavedet endrer seg for de enkelte fødselskull. For kvinner er disse kohorteffektene illustrert i figur 5.

Figur 5. Førerkortinnhavedet for kvinner i ulike aldersgrupper i 1980 og 2000. Forventet førerkortinnhavedet i 2020 og 2030.



TØI-rapport 766/2005

Det er få personer som taper førerkort før fylte 70 år. En realistisk modell innebærer derfor stigende kohortkurver fram til 70 års alder eller lenger. Kohortkurvene framkommer ved å binde sammen punkter i ulike periodekurver som vedrører samme generasjon (kohort). De røde pilene angir kohortkurven for førerkortandel for kvinner som i 1980 var mellom 25 og 29. I 2000 var disse kvinnene mellom 45 og 49 år, og førerkortandelen hadde økt fra 70 % 20 år tidligere til nærmere 90 %. Denne andelen holder seg tilnærmet konstant frem til 2020, før den synker fremover mot 2030. Da er kvinnene i alderen 75-79 år. Andelen førerkort er likevel større enn den var for de samme kvinnene i 1980. Når modellene skal benyttes til langsiktige prognoser vil kalibreringen sørge for å ivareta disse effektene.

## Modeller for valg av transportmiddel og destinasjon

### Data

Datamaterialet til estimeringen av disse modellene er basert på tre ulike kilder. *Reisevanedata* (fra RVU2001 og PRVU01/02) gir informasjon om valgt transportmåte og valgt destinasjon for rundturer fra bosted til én hoveddestinasjon og tilbake til bosted, om kjennetegn ved informantene som har gjennomført rundturene, og om informantenes hushold. *Sonedata* gir informasjon om hva som befinner seg av attraksjoner (arbeidsplasser, befolkning, med mer) i mulige destinasjoner, og *LoS-data* (Level of Service / transportstandard) gir informasjon om reisetider og kostnader forbundet med besøk i de ulike destinasjoner. All informasjon som er benyttet er stedfestet til grunnkretser som geografisk enhet. Grunnkretser er den mest detaljerte standard geografiske inndelingen i Norge (ca 13500 stk i hele landet). Siden modellene vil reflektere kvaliteten på de tre datakilder er det lagt en del ressurser i å ”vaske” dataene.

Utgangspunktet for reisevanene i estimeringsgrunnlaget er informantenes redogjørelse av alle gårdsdagens reiser, med valgt transportmiddel, valgt destinasjon, reisetidspunkt, reisetid, med mer, for alle bevegelser utenfor bostedet. Dette skal gi opphav til en sammenhengende sekvens med reiser som, for de aller fleste informantene, starter i bostedet om morgenen og ender i bostedet om ettermiddagen eller kvelden. Ved gjennomgangen av datamaterialet ble det identifisert en del usammenhengende reisesekvenser og andre feil, som enten ble rettet opp eller forkastet (17 % av de ca 93000 delreisene ble forkastet). De sammenhengende reisesekvensene som starter og ender i bostedet ble gjort om til rundturer med én eller flere destinasjoner. Det ble i alt ca 35000 slike rundturer, hvorav vel 75 % bare har én destinasjon. Rundturer med to og flere destinasjoner ble behandlet spesielt for å finne ut om én av destinasjonene kunne benyttes som hoveddestinasjon. Rundturer som ”lignet” på en ren tur/retur reise ble akseptert til å inngå i estimeringsgrunnlaget, men spesifisert som en tur/retur, mens mer kompliserte rundturer ble forkastet. Forkastingen var basert på visse kriterier knyttet til tid tilbrakt på hver destinasjon, avstand fra bosted til hver destinasjon, og faktisk reiseavstand for hele rundturen sammenliknet med avstand til og fra bostedet til potensielle hoveddestinasjoner. Ca 35 % av rundturene med to og flere destinasjoner ble akseptert og disse ble tilordnet en variabel som indikerer at én eller flere sekundære destinasjoner er tatt bort fra disse rundturene. Resten av rundturene ble vurdert som for komplisert og forkastet. Årsaken til den relativt

konservative behandlingen av disse rundturene er knyttet til problemene ved å spesifisere korrekte LoS-data til de rundturer som har flere destinasjoner når disse ikke likner på en ren tur/retur.

Når observasjoner med manglende stedfesting, uaktuelle transportmåter og reisehensikter, eller andre dataproblemer er fjernet, omfatter estimeringsgrunnlaget i alt ca 23000 rundturer. For hver rundtur trekkes det et tilfeldig utvalg ("random sampling of alternatives") på maksimalt 249 destinasjoner fra mengden soner innenfor 10 mil fra bostedet. Disse destinasjonene brukes som alternativer til den destinasjon som faktisk er valgt, og i datafilen kobles sonedata for attraksjon og LoS-data både til den valgte og til de uttrukne, alternative destinasjoner.

## Reisehensikter og transportmåter

Det er estimert bosteds- og rundturbaserte modeller for de fem reisehensiktene:

- Arbeidsreiser
- Tjenestereiser
- Handle/servicereiser
- Private besøksreiser og andre private reiser

Estimeringsarbeidet har vært en relativt omfattende prosess med prøving og feiling, feilsøk og korrigeringer av datamaterialet. Vi har fremdeles et inntrykk av at spesielt kollektivtransport er relativt skjematisk behandlet i nettverkene, i hvert fall i deler av landet. På den annen side har hver korrigering som er gjennomført, og som har gitt nye LoS-matriser, gitt relativt moderat påvirkning på modellestimatene.

Modellene er estimert for transportmåtene bilfører, bilpassasjer, kollektivtransport, sykkel og gang (til fots). Flest reiser er gjort som bilfører, og denne andelen varierer fra 47 % (besøksreiser) til 69 % (tjenestereiser). Det er færrest sykkeltureturer i materialet, fra 3 % for handle/servicereiser til 8 % for besøksreisene. Det er videre slik at en relativt stor andel av reisene ikke har kollektivtransport tilgjengelig (hele 60 % for noen reiseformål), noe som tyder på at kollektivnettet ikke er så detaljert kodet som man skulle ønsket. I datamaterialet er det signifikante forskjeller mellom kvinner og menn både når det gjelder transportmiddel-fordeling og avstandsfordeling. Relativt sett er kvinner oftere bilpassasjerer enn menn og de benytter hyppigere kollektivtransport eller går til arbeid. I tillegg reiser de i gjennomsnitt kortere enn menn, dette er spesielt tydelig for arbeidsreisene og tjenestereisene.

## Hierarkisk struktur

For alle reiseformål har vi testet ut ulike trestrukturer. Generelt har vi fått som resultat at varianter med destinasjonsvalget rangert over transportmiddelvalget fikk en logsumkoeffisient som ikke var signifikant forskjellig fra 1. Variantene med transportmiddelvalget over destinasjonsvalget hadde en noe bedre log-likelihood-verdi enn de multinomiske modellene, men kostnadskoeffisienten var betydelig lavere i tallverdi og signifikans, noe som gav betydelig høyere implisitte



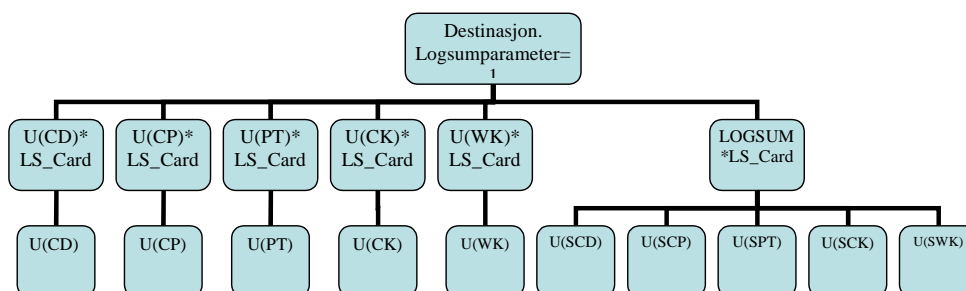
tidsverdier. I og med at vi allerede i utgangspunktet hadde tidsverdier som var i overkant av forventet, valgte vi å benytte multinomiske modeller framfor strukturerte ved valg av transportmiddel og destinasjon for tre av reisemåtene (tjenestereiser, besøksreiser og handle-/servicereiser). Modellen for andre private reiser er strukturert med destinasjonsvalget over transportmiddelvalget. Denne modellen var spesielt vanskelig å estimere med tilfredsstillende resultater. Årsaken er trolig at dette er en relativt lite homogen reisehensikt som bl.a. omfatter fritidsreiser, rekreasjonsreiser, hente/bringe reiser og en del andre, mindre hyppige reisehensikter.

Modellen for arbeidsreiser har en noe mer komplisert struktur enn de andre modellene, da vi her har valgt å estimere transportmiddel/destinasjon og periodekortinnehav simultant. Den enkle (og mest vanlige) måten å håndtere periodekort for kollektivtransport på, er uten tvil å benytte månedskortpris/antall reisedager som kostnad for dem som har periodekort (til alle aktuelle destinasjoner). Dette er en praksis som har flere ulemper, hvorav den viktigste er at man ikke får tatt hensyn til at kollektivtransport gjerne vil oppfattes som gratis på registreringsdagen. De som har periodekort vil også gjerne bare være sporadiske brukere av bil på arbeidsreisen, og det er derfor rimelig å anta at de ikke kan utnytte rabattmuligheter når det gjelder eventuelle bompenger og ferger.

## Periodekort

I modellen for arbeidsreiser har vi derfor valgt å behandle reiser med periodekort som et eget nest med en egen logsumparameter. Dette innebærer at vi får 5 ekstra nyttefunksjoner, én for hver reisemåte. Nyttefunksjonene her (SCD, SCP, SPT, SCK, SWK) er nestet med logsumparameteren LS\_CARD. Nyttefunksjonene uten periodekort (CD, CP, PT, CK, WK) multipliseres også med samme logsumparameter for å bringe dem på samme nivå (et "dummy-nest"). Dette er nødvendig for å få korrekte generiske parametre. Modellen er multinomisk i valg av destinasjon og transportmiddel.

Figur 6. Modellstruktur for arbeidsreiser



TØI-rapport 766/2005

Kostnaden for en arbeidsreise ved bruk av periodekort settes til månedskortpris/22. Denne kostnaden legges til alle transportmåter under nestet med periodekort, og tilordnes en egen parameter. Kollektivtransport har kun denne kostnaden under nestet med periodekort, mens de øvrige transportmåter i tillegg

er tilordnet de ordinære kostnadene forbundet med disse. Internt i nestet for reiser med periodekort vil dermed kollektivtransport fremstå som gratis, og mellom nestene vil kollektivtransport være billigere med periodekort enn uten. Samtidig vil det være dyrere å reise med andre transportmåter hvis man har periodekort. Det viser seg at parameteren til prisen på periodekort blir høyere enn parameteren for andre kostnader (en krone anvendt på månedskort "vurderes" til kr 1.30), noe som kan skyldes at kjøp av periodekort innebærer et relativt stort kontantutlegg, samtidig som det alltid vil være en viss usikkerhet knyttet til hvor mange reiser det faktisk vil benyttes for (pga sykdom, risiko for at kortet mistes osv). Det er flere fordeler ved å operere med et eget nest for periodekort, bla:

- Valgbetingelsene blir så langt mulig korrekt spesifisert og vi bryter samtidig en del av korrelasjonen mellom reisetid og kostnad. Prisen på periodekort påvirker destinasjonsvalget og valget mellom kort/ikke kort for en gitt destinasjon, men for dem som velger periodekort, fremstår kollektivtransport som gratis i forbindelse med transportmiddelvalget.
- Vi unngår problemet med spesifisering av kostnaden for kollektivtransport for alternative destinasjoner og for alternative transportmidler, siden reiser med periodekort er egne alternativer.
- Vi kan differensiere kostnaden for bilførere og bilpassasjerer etter hvorvidt disse transportmåter benyttes sporadisk eller ikke. Dette bidrar også til å bryte opp korrelasjoner.
- Vi får automatisk en modell for periodekortinnehav som kan benyttes i forbindelse med øvrige transportmidler dersom arbeidsreisemodellen kjøres først.
- Det blir mulig å analysere effekter av å endre prisdifferansen mellom enkeltbillett/klippekort og periodekort med modellen. Prisforskjellen vil påvirke både kortinnehav og reisemiddelvalg.

## Tidsverdier

I alle modellene finner vi at kvinner har høyere tidsverdier enn menn, om enn i varierende grad. Størst forskjell er det i modellene for tjenestereiser, der tidsverdiene for kvinner er nesten dobbelt så høye som for menn. For besøksreiser og handle/servicereiser finner vi forskjell i tidsverdi mellom kvinner og menn kun for bilførere, mens vi i kategorien "andre reiser" også har det for bilpassasjerer. For tjenestereiser og arbeidsreiser finner vi forskjeller mellom kvinner og menn både for bilførere, bilpassasjerer og kollektivreisende. Differensiering av tidsparametrene ble i utgangspunktet forsøkt som en proxy for inntektsforskjeller, men forskjellen ble motsatt av hva man da skulle forvente. Det er flere mulige forklaringer på dette:

- Kvinner har gjennomgående et strammere tidsbudsjett enn menn, noe også tidsverdistudier bekrefter.
- Kvinner oppfatter gjennomgående ulempen ved å kjøre bil som større enn det menn gjør.

- Kvinner kan gjennomgående ha bedre muligheter til å skaffe arbeid nær hjemmet enn menn. I så fall vil forskjellen i tidsparametere reflektere at soners attraktivitet som mål for arbeidsreiser for hhv menn og kvinner ikke er godt nok spesifisert.
- Kvinner og menn har ulike typer tjenestereiser, hvor kvinners reiser generelt er kortere.
- Gjennomsnittsavstanden til valgte reisemål generelt er kortere for kvinner enn menn.

Generelt er det slik at de implisitte tidsverdiene som er beregnet virker noe høye for alle modellene, med unntak av arbeidsreiser. Viktigste grunnen til dette kan være at reisetidene som er brukt ved estimeringen er beregnet i en nettverksmodell hvor det ikke forekommer kapasitetsproblemer. Tidene er også beregnet med utgangspunkt i 85 % av skiltet hastighet for alle vegtyper. Begge forhold vil bidra til en undervurdering av reisetiden, spesielt for veglenker av mindre bra standard når det gjelder kurvatur, stigning og vegbredde og for bygater og veglenker med lyskruss.

Tabell 1. Implisitte tidsverdier i modellene (NOK/t)

	Arbeid	Tjeneste	Besøk	Handle/serv.	Annet
<b>Bilfører (CD):</b>					
Mann		135	63	66	198
Kvinne		242	89	87	253
Tillegg storby		58			
Tillegg helg			-19		-84(m), -77(k)
Mann under 50 år	55				
Mann, 50 år +	73				
Kvinne, under 50 år	84				
Kvinne, 50 år +	106				
Tillegg destinasjon Oslo i rush	32				
<b>Bilpassasjer (CP):</b>					
Bilpassasjer			61	55	
Mann	92	64			198
Kvinne	154	147			253
Tillegg storby		58			
Tillegg helg			-19		-84(m), -77(k)
<b>Kollektiv (PT):</b>					
Ombordtid			33	28	92
Ombordtid, mann	38	56			
Ombordtid, kvinne	48	125			
Tillegg helg					-32
Gangtid	43	168	61	57	133
Ventetid ved 5 minutters ventetid	60	369	70	84	124
Ventetid ved 30 minutters ventetid	24	150	29	34	51
Ventetid ved 60 minutters ventetid	17	106	20	24	36
Overgang (NOK pr overgang)		39		16	

TØI-rapport 766/2005

Hvis vi ved datauttak til estimering systematisk har undervurdert reisetider med bil, vil dette ved modellestimering bli kompensert ved en høyere tallverdi på parametere for reisetid med bil. Tilsvarende kan i noen grad også gjøre seg gjeldende når det gjelder reisekostnader, hvor vi ettertid kan konstatere at data kunne vært bedre når det gjelder ferger og bompenger. I tillegg er det korrelasjon mellom tid og kostnad, som gir en tendens til at en av parametrene fanger opp den samlede effekten av avstand.

Systematiske skjevheter når det gjelder data for reisetider og reisekostnader som er benyttet for estimeringen vil altså også kunne påvirke parameterestimater og implisitte tidsverdier. Dette kan være særlig alvorlig når modellen senere anvendes med kvalitetssikrede data som innebærer mer eller mindre systematiske endringer i forhold til de data modellene er estimert på.

Tidsverdiene er spesielt høye for kategorien andre private reiser, noe som trolig henger sammen med at reiseformålet er lite homogent og spenner fra lange rekreasjonsreiser til korte hente/bringe reiser. Mange av disse reisene foregår i helgene, og dummyvariable i interaksjon med tidsvariablene for bil reduserer tidsverdiene i helgene betraktelig, noe som kan ha sammenheng med at folk generelt har slakkere tidsbudsjett i helgen.

## Avstandsfordeling

Ved estimeringen av modellene har vi hatt visse problemer med å treffe med observert avstandsfordeling. Dette er vanskelig når en modell skal dekke hele avstandsintervallet fra noen meter til 20 mil for en tur/retur reise, samtidig som nyttefunksjonene er lineære i koeffisienter og variable (spesielt når avstandsfordelingen i datamaterialet ikke er tilnærmet lineær). Tendensen er at modellene underpredikerer korte og lange reiser og overpredikerer de mellomlange. Vi har testet ut en del ikke-lineære transformasjoner av tidsvariablene, trinnvis lineære spesifikasjoner, samt ulike interaksjoner mellom variable uten at det gav tilfredsstillende resultater sett i sammenheng med den økning i regnetid som slike formuleringer vil medføre. Vi har derfor lagt inn enkelte avstandsdummys i modellspesifikasjonen, noe som har ført til en markert bedring av modellenes evne til å gjenspeile datamaterialets avstandsfordeling, men dette er variable som eventuelt kan fjernes ved anvendelser. Dummyvariable for avstandsbånd anbefales satt til null i forbindelse med implementeringen. Programmet som kjører modellen har nå visse muligheter for kalibrering av avstandsfordeling.

## Andre hovedtrekk

Vi vil ikke gå inn på alle variable som har vist seg signifikant i de ulike modellene, men nevner noen fellestrekk for modellene:

- Ventetid i kollektivtransport er en variabel som var problematisk i forbindelse med estimeringen. Dette skyldes at man i nettverksmodellen, hvor denne variabelen beregnes, får ut halvparten av tiden mellom avgangene, summert over antall påstigninger (ev. en maksimal ventetid pr påstigning) og ikke det som er folks faktiske ventetid. Vi testet ut et utall varianter, som åpen og skjult ventetid, ventetid pr påstigning, ulike lineære transfor-

masjoner, sammenslått gangtid og ventetid osv, for å få ventetiden signifikant. Vi endte opp med en formulering basert på kvadratroten av total ventetid. Ved dette får man redusert effekten av lange ventetider, samtidig får man en ventetidsvekt som varierer med størrelsen på ventetiden.

- En dummy som angir at turen har minst én sekundær destinasjon (flere besøk underveis) øker sannsynligheten for bruk av bil. Denne medfører også at rundturer med mer enn én destinasjon får økt andel bilførerturer slik modellen er implementert.
- De fem ulike kategorier for førerkort/bilhold i husholdningen er viktige forklaringsvariable i alle modellene. Generelt fører god eller full biltilgang til økt sannsynlighet for bruk av bil som fører, samtidig finner vi at det i hushold med færre biler enn førerkort ofte er kvinnene som taper ”kampen om bilen”.
- Det er lagt inn enkelte dummys for alder og type region (storby osv), men dette er forsøkt begrenset for å unngå altfor mange segmenter i de endelige modellene.
- Vi har begrenset med informasjon om parkeringsforhold knyttet til den enkelte reise. I stedet har vi generert noen indekser som sier noe om arbeidsplass tettheten i destinasjonene, med det for øye at disse skal si noe om hvor vanskelig (og kanskje dyrt) det er å finne parkeringsplass. For arbeidsreiser og tjenestereiser finner vi at de to høyeste arbeidsplass tetthetene reduserer tilbøyeligheten til å kjøre bil, med størst ulempe knyttet til indeksen for høyest arbeidsplass tetthet. Denne kategorien arbeidsplass tetthet virker også negativt inn på bilbruken for handle/servicereisene og andre private reiser, mens vi ikke finner noen slik sammenheng for besøksreisene.
- I modellen for arbeidsreiser er det estimert en ekstra parameter for reisetid med bil når destinasjonen er i Oslo. Denne vil delvis kompensere for det forhold at reisetider med bil er beregnet for ubelastet nett og således vil undervurdere faktisk reisetid.

## Modeller for valg av reisefrekvens

Datagrunnlaget for estimering av modeller for reisefrekvens er personfilen fra RVU2001, påkodet informasjon om antall besøk foretatt registreringsdagen. Dette er gjort for de 5 reisehensikter det er estimert transportmiddel- og destinasjonsmodeller (MD-modeller) for, i tillegg til skolereiser. Skolereisene er ellers behandlet relativt skjematisk i modellsystemet, med gravitasjonsmodeller som fordeler barn og ungdom i grunnskolealder til nærmeste barne- eller ungdomsskole. Fordi kollektivtilbudet for skolereisene ikke er kodet i nettverkene er det for disse reisene ikke laget noen modell for valg av transportmiddel. Skolereisene tas likevel med i frekvensmodellene.

Frekvensmodellene gir forventet *antall besøk* for hver reisehensikt. Konseptet går i korthet ut på å estimere en Hurdle-Poisson-modell for totalt antall besøk i kombinasjon med en logit-modell som splitter totalt antall besøk på ulike formål (6 i alt). Poisson er en fordeling for ikke-negative heltall, som egner seg godt når ut-

fallsvariabelen er av nettopp denne typen. Hurdle-Poisson kan benyttes når f eks 0 har en avvikende sannsynlighet fra en standard Poisson-fordeling, noe som tydeligvis er tilfelle for antall besøk. Med dette konseptet spesifiseres én nyttefunksjon for hver reisehensikt. Denne formuleres med alternativspesifikke konstanter og variable som beskriver kjønn, alder, familietype, bostedstype, samt en logsum fra korresponderende MD-modell. I estimeringen vektet nyttefunksjonene med antall besøk som er gjennomført for hver reisehensikt. Inntekt inngår ikke direkte i disse nyttefunksjonene, men kommer inn fra bilholdsmodellene via logsummene fra MD-modellene. Koblingen mellom den vektete logitmodellen og Hurdle-Poisson skjer gjennom formulering av en samlet logsum fra logitmodellen for fordeling på besøk med ulike reisehensikter. Denne samlede logsum inngår i fordelingsfunksjonen i Hurdle-Poisson modellen. Med utgangspunkt i denne kombinerte modellen kan vi beregne forventet antall besøk for de seks reisehensiktene.

Estimeringsarbeidet startet med forsøk på å formulere én samlet modell for alle befolkningssegmenter. Dette krevde imidlertid en dyptgående segmentering med dummyvariable for ulike aldersgrupper. Det ble derfor estimert uavhengige modeller av samme type for 5 ulike aldersgrupper (13-24 år, 25-34 år, 35-54 år, 55-66 år og 67 år og mer). Parameteren knyttet til logsummene fra MD-modellene har størst tallverdi i modellene for de yngste og spesielt de eldste aldersgruppene. Siden logsummene i stor grad reflekterer transportressursene til segmentene (egne transportressurser så vel som generell tilgjengelighet til attraktive destinasjoner med ulike transportmåter), reflekterer dette at de midlere aldersgrupper har et mindre elastisk transportbehov enn de yngste og eldste. Situasjonen når det gjelder tilgjengelighet er altså mer avgjørende for reisefrekvensene for de yngste og eldste. Tabell 2 viser modellenes fordeling på besøk for de 6 typer av reisehensikter som er definert. På nasjonalt nivå ser vi at situasjonen er svært ulik mellom aldersgruppene, og dette er noe av bakgrunnen for at inndelingen av materialet i aldersgrupper viste seg nødvendig.

Tabell 2. Fordeling på besøk med ulik reisehensikt etter modell/aldersgruppe. Virkedøgn, hele landet. Prosent av alle reisene i aldersgruppen.

Formål/alders	13-24	25-34	35-54	55-66	67+	Totalt
Arbeid	12	26	31	28	3	24
Tjeneste	3	9	11	9	1	8
Handle/service	19	24	24	29	53	26
Besøk	15	10	7	11	14	10
Andre private	26	29	27	23	30	27
Skole	24	3	1	0	0	6

TØI-rapport 766/2005

Totalt for hele landet treffer modellene mot datamaterialet med under 1 % avvik, mens det er noe større avvik når man går inn i de 5 regionene. Dette har gjort det nødvendig å introdusere et mindre antall regionale dummyvariable.

## Prosedyre for beregning av "mellomliggende" turer

Modellene for valg av transportmiddel og destinasjon er estimert på data som reflekterer rene tur/retur-reiser til én hoveddestinasjon. Mer komplekse rundturer

med flere destinasjoner er enten forenklet (hvis turen likner på en ren tur/retur-reise) eller forkastet fra datamaterialet for estimeringen av disse modellene. I materialet for estimeringen av frekvensmodellene er imidlertid alle ærend tatt med, også ærend som er gjennomført gjennom komplekse rundturer med flere destinasjoner.

Hvis resultatene fra transportmiddel- og destinasjonsmodellene og frekvensmodellene ble benyttet direkte ville modellsystemet produsert for mange delreiser, da noen av besøkene i virkeligheten foretas i forbindelse med ”mellomliggende” turer og ikke rundturer. For å korrigere for dette er det utarbeidet en metode som i prinsippet skal gi et tilnærmet korrekt antall delreiser, dvs en prosedyre som også tar hensyn til rundturer med mellomliggende destinasjoner.

Prosedyren benytter informasjon fra transportmiddel- og destinasjonsmodellene og frekvensmodellene på en slik måte at systemet produserer det riktige antall utgående reiser, mellomliggende reiser og hjemreiser, basert på matriser med ”overgangssannsynligheter”. Dette er sannsynligheten for at en reise med hensikt ”i” etterfølges av en hjemreise eller fortsetter til en annen destinasjon med formål ”j” og deretter hjemreise.

Matrisene med overgangssannsynligheter kan beregnes med utgangspunkt i RVU-data. For alle delreiser som ikke er hjemreiser registreres reisehensikten og hensikten for neste reise. Alle slike par telles opp og legges i en tabell, slik at sannsynlighetene kan beregnes.

Dette opplegget må baseres på visse forenklinger for å være beregningsmessig håndterbart. Ser vi på RVU-data for rundturer som starter og ender i eget hjem, vil det være et varierende antall besøkte steder i slike rundturer. Den første forenkling innebærer at alle slike rundturer forutsettes å skje med enten ett eller to besøkte steder. Den andre viktige forutsetning er at sannsynligheten for valg av destinasjonen for 2. besøk i en rundtur med to destinasjoner ikke avhenger av tid og kostnad ved reise fra destinasjon 1 og destinasjon 2, men av de respektive destinasjoners attraktivitet i forhold til reiser til/fra eget hjem.

I implementeringen beregnes matriser for mellomliggende turer bare for bilfører og kollektivtrafikk, dette for å spare regnetid og lagringsplass.

## Programmering/implementering

De ulike delmodeller er programmert i C++. Siden programmet vil benyttes for modeller som kan ha mer enn 6000 soner, er det lagt spesiell vekt på at det skal gå raskt. I tillegg er det lagt vekt på at koden skal være enkel å vedlikeholde. Programmet er plattformuavhengig, dvs det kan kjøres i kombinasjon med TRIPS/-CUBE eller EMMA. Resultatfiler (matriser) skrives ut på TRIPS-format og det vil senere legges inn en opsjon som gir EMMA-format.

Alle parametre og konstanter som inngår i modellen leses fra filer, og det er således mulig å endre på disse.

Det er lagt opp til håndtering av relativt store datamengder i internminnet, og for store regioner anbefales minimum 2 GB internminne for de maskiner som skal

kjøre modellen. I tillegg er det selvsagt en fordel at klokkefrekvens på prosessor og kapasitet/hastighet på harddisk er så kraftig som mulig.

Programmets regnetid bestemmes nå primært av antall soner og, for den enkelte sone, av antall naboer (dvs antall soner innenfor en omkrets på 10 mil). Beregningstiden for to modellområder med like mange soner kan derfor variere forholdsvis mye avhengig av gjennomsnittsstørrelsen på sonene.

Det er i løpet av vinteren 2005 og våren 2006 gjort en del endringer i programmet som kjører etterspørselsmodellen. Man har nå visse muligheter til å kalibrere på turlengdefordeling og det er også gjort visse endringer når det gjelder behandling av ventetiden på ferger. Programmet skriver nå også ut en fil med aggregerte rammetall fra beregningen. Denne er benyttet for å kalibrere modellene for de enkelte regioner mot tilsvarende rammetall beregnet ved oppblåsing av RVU 2001 til totalpopulasjon. Det er også rettet noen "bugs" som ble oppdaget i løpet av denne perioden.

## Konklusjoner fra foreløpig testing

Modellen gir virkedøgntrafikk for personreiser en "normal" virkedag, dvs virkedager utenom ferietider (sommer, jul og påske) og for reiser kortere enn 10 mil én vei. En kontroll mot vegtrafikktegninger bør fortrinnsvis gjøres mot tilsvarende data. Siden tellinger dekker veger med svært ulike karakteristika, kan enkle omregninger av modelltall til ÅDT lett bli misvisende. Det beste sammenligningsgrunnlag er direkte tellinger for lette biler på virkedager, f.eks et gjennomsnitt for månedene september til november. Disse tellinger bør imidlertid korrigeres for at en andel av de lette bilene går i varetransport, som ikke dekkes av modellen, samt at en andel av trafikken vil være lange reiser som skal dekkes av modellen for lange reiser. Denne andelen vil kunne variere mye fra tellepunkt til tellepunkt. Man skal også være oppmerksom på at "feil" når det gjelder modellberegninger av vegvalg med TRIPS eller EMMA, kan gi avvik i forhold til tellinger. Dette er spesielt viktig i byområder, hvor det tidvis er køer som påvirker vegvalget.