

Sammendrag

Etablering av etterspørselsmodell for korte personreiser

Teknisk dokumentasjon fra estimeringen

TØI rapport 1814/2021

Forfattere: Jens Rekdal, Odd I. Larsen, Tom N. Hamre, Olav Kåre Malmin,
Nina Hulleberg, Stefan Flügel og Anne Madslie
Oslo 2021 339 sider

I perioden 2016-2018 ble det estimert og implementert nye etterspørselsmodeller for korte personreiser i Norge. Arbeidet med modellutviklingen ble gjennomført av et konsortium bestående av Møreforskning Molde AS, Transportøkonomisk institutt, Sintef, Numerika AS og TransMod AS. Denne rapporten er en teknisk dokumentasjon av estimeringsarbeidet som ledet fram til en operativ modell i 2018.

Modellsystemet består av modeller for valg av transportmiddel og destinasjon, modeller for turgenerering samt modeller for segmentering etter biltilgang. Estimeringen av hver av disse modellene er beskrevet i detalj i rapporten. Implementering av modellsystemet i CUBE var også en del av prosjektet, men er ikke inkludert i foreliggende. Dette dokumenteres i en egen rapport fra SINTEF.

Det er de senere år gjort enkelte endringer og forbedringer i noen av delmodellene. Dette er beskrevet i ulike arbeidsdokumenter, men er ikke tatt med i denne rapporten som beskriver modellene slik de forelå i 2018.

Innledning

I juli 2015 utlyste Statens vegvesen Vegdirektoratet, Jernbanedirektoratet, Kystverket og Avinor AS anbudskonkurransen «Rammeavtale for etablering av etterspørselsmodell for korte personreiser». Første prosjekt innenfor rammeavtalen var estimering og etablering av en ny etterspørselsmodell til de regionale persontransportmodellene. Bak rammeavtalen står Samferdselsdepartementet og de nevnte transportvirksomhetene, med Statens vegvesen som kontraktspart.

Møreforskning Molde AS (MFM), Transportøkonomisk institutt (TØI), Sintef, Numerika AS og TransMod AS, leverte et felles tilbud på etablering av ny etterspørselsmodell for korte personreiser. Konsortiet ble tildelt oppdraget i juli 2016.

Modellutviklingen har bestått av 5 faser, som delvis har foregått parallelt.

- Fase 1 – Etablering av datasett for estimering, implementering og anvendelse.
- Fase 2 – Estimering av modellsystemet:
 - modeller for valg av transportmiddel og destinasjon (MD-modeller).
 - modeller for turgenerering (TG-modeller).
 - modeller for segmentering etter biltilgang (BHFK-modeller).
- Fase 3 – Implementering av de nye modellene.
- Fase 4 – Implementering av brukergrensesnitt i CUBE.
- Fase 5 – Kalibrering og uttesting .

Foreliggende rapport gir en detaljert og teknisk gjennomgang av arbeidet med Fase 1 (Etablering av datasett) og Fase 2 (Estimering av modellsystemet). Fase 3 (Implementering) er kort omtalt, mens arbeidet i Fase 4 (Implementering av brukergrensesnitt i CUBE) og Fase 5 (Kalibrering og uttesting) ikke er omtalt i denne rapporten. Implementering i CUBE

er beskrevet i Sintef-rapport 2021:01297 (Tørset, Malmin og Flaata, 2021), som også fungerer som en brukerveiledning til modellsystemet.

I rapporten beskrives estimeringen av de ulike delmodellene i detalj, med struktur og estimerte parametere slik det så ut da modellsystemet ble tatt i bruk i 2018. Det har senere vært flere større og mindre revisjoner av enkeltelementer i modellsystemet. Det vil derfor være avvik mellom dagens modell (2021) og slik modellen er beskrevet i denne rapporten, både når det gjelder verdien på enkeltparametere og andre elementer. Hovedstrukturen i modellen er imidlertid beholdt slik som beskrevet i denne rapporten. Endringer etter 2018-versjonen av modellen er dokumentert i ulike arbeidsnotater til oppdragsgiver.

Rapporten er lang og til dels meget teknisk, med mange tabeller og lange vedlegg. Det anbefales derfor å bruke innholdsfortegnelsen og søkemuligheter aktivt for å finne fram til det man er interessert i.

Datamaterialet

Reisevaner

Reisevanedata (RVU-data) danner grunnstammen i estimeringen av transportmodeller med en geografisk dimensjon. Reisevanedata beskriver et lite utvalg av den norske befolkningens reisevaner. Både individuelle kjennetegn og hvilke valg intervjuobjektene (IO) har foretatt er viktig informasjon i estimeringsarbeidet. Når det gjelder individuelle kjennetegn er alder, kjønn, familietype, bilholds-ressurser og inntekt, variabler som inngår i modellene. I tillegg utnyttet informasjon om eventuelt firmabilhold, og også en del karakteristika for de reisene som er gjennomført (for eksempel størrelsen på reisefølget). Intervjuobjektet blir spurt om det er gjennomført en reise den aktuelle dagen, i tilfelle hvor mange reiser, samt hvilken transportmåte som er valgt og hvilke destinasjoner man eventuelt har reist til. En del av datamaterialet fra reisevaneundersøkelsene benyttes også i forbindelse med implementeringen av modellene i en programkode.

Ved etablering av de nye modellene er det brukt data fra den nasjonale reisevaneundersøkelsen gjennomført i 2013/14. Materialet består av intervjuer med ca. 59 000 intervjuobjekter, hvorav ca. 82 % har rapportert at de har gjennomført én eller flere reiser intervjudagen. Gjennomsnittlig antall gyldige besøk som er rapportert er ca. 1,86 per døgn.

I RVU innhentes det ikke informasjon om karakteristika ved selve reisene (reisetider, reisekostnader, avgangsfrekvenser, etc.). Denne type informasjon er likevel svært sentrale data både til estimering av de ulike modellene og i den senere anvendelsen av dem. Reisenes karakteristika i form av tidsbruk og kostnadsaspekter, gjerne kalt **LoS-data** (Level of Service), beregnes i såkalte nettverksmodeller. Forenklede versjoner av veg- og kollektiv-rutenettet er der kodet inn i et koordinatsystem og med distanser, reisetider, reisekostnader, etc. mellom punkter (noder) i geografien. Noen av disse punktene er soner hvor reiser starter og slutter. I det regionale modellsystemet brukes grunnkretser (ca. 14 000 stk. i Norge) som slike geografisk avgrenede soner hvor reiser oppstår og ender.

Nettverks- og LoS-data

Nettverkene som ble benyttet til etablering av **LoS-data** til estimeringen, besto av egne nettverk for biltransport, kollektivtransport inkl. tilbringer og for sykkel. Fotgjengere antas å bevege seg langs korteste vei i kodet vei- og sykkelnett. Det ble lagt ned betydelige ressurser i å kvalitetskontrollere LoS-dataene som produseres i de regionale modellene, bl.a. ved å sammenstille reisetider og reisedistanser for bil på et stort antall relasjoner med tilsvarende informasjon hentet fra GoogleMaps. Tilsvarende ble gangtider og ombordtider for kollek-

tivtransport fra modellen sammenlignet med informasjon fra GoogleMaps og kollektivselskapenes nettsted. Kontrollarbeidet var tidkrevende, men resultatet ble betydelig mer presise LoS-data til estimeringen enn man hadde i utgangspunktet.

For kollektivtransport ble det også kontrollert for symmetri, for å unngå feil som hindret bruk av kollektivtransport i begge retninger mellom to soner. Det ble også gjennomført et arbeid med å opprette taksttabeller for de ulike kollektivselskapene.

LoS-data for sykkel ble etablert på en mer detaljert måte enn i tidligere modellversjon, ved at man denne gang også tok med geometriske egenskaper (vertikalkurvatur) ved veglenkene, samt hvilken infrastruktur for syklende som finnes (om man må sykle i vegbanen eller kan bruke gang-/sykkelveg eller oppmerket sykkelfelt).

Sonedata

Sonedata er en tredje viktig datatype både til estimering av modeller og i senere anvendelser. Dette er data som beskriver innholdet i hver sone eller gir indikasjoner på hvilke typer aktiviteter som er gunstig å gjennomføre i de ulike sonene. Befolkning, arbeidsplasser fordelt etter næringer, antall hoteller og hytter/fritidshus, etc. er eksempler på data som ble samlet inn og benyttet. Av viktige nye/oppdaterte data kan nevnes:

- Data for fordeling av yrkesaktive bosatte på utdanning og for fordeling av arbeidsplassene (arbeidstakerne) etter kjønn, aldersgrupper og utdanning:
 - ”Produktsum” for utdanning inngår i modellene for arbeidsreiser, tjenestereiser og arbeidsplassbaserte reiser.
 - Produktsummen blir en OD-matrise mellom soner hvor det bor yrkesaktive og soner med arbeidsplasser.
 - Dette øker sannsynligheten for å reise til soner med arbeidsplasser som har en lik utdanningsprofil som i bostedssonen.
 - Det gir også en viss ekstra stivhet i destinasjonsvalget.
 - Kjønnintensitet (hhv. kvinne- og mannsdominerte arbeidsplasser) inngår også i disse modellene nesten på samme måte.
- Ny klassifisering av arbeidsplasser etter næringer basert på nye NACE-koder, men litt mer spisset inn mot de ulike reisehensikter.
- Nye data for medianinntekt i delområder ivaretar den geografiske variasjonen i inntekt
 - og mer detaljerte nasjonale tall for husholdsinntekter etter antall voksne i husholdet, familietype, alder og kjønn hever kvaliteten/presisjonen på inntektsdata mellom segmentene i modellsystemet.
- Tidligere har vi kun hatt tilgang til data for sonenes totalareal (uten vann).
 - Nye og mer detaljerte arealdata er tatt i bruk, med 12 arealkategorier.
 - Ny tetthetsvariable, med deflator basert på *bebygd* areal.
 - Tetthetsvariable (som indikator for parkeringsproblemer) inngår i tillegg til parkeringskostnader både på bostedssiden og destinasjonssiden i flere av MD-modellene (transportmiddel-destinasjonsmodellene), og også i forbindelse med modellene for biltilgang (bostedsparkering).

Estimering av modellsystemet

Modeller for valg av transportmiddel og destinasjon (MD)

De nye modellene for transportmiddel og destinasjon er estimert for 6 «reisehensikter», mot 5 i tidligere modellsystem. Disse er:

- Bostedsbaserte arbeidsreiser (til/fra arbeidsplass).
- Bostedsbaserte tjenestereiser (reiser i arbeidstiden).
- Bostedsbaserte fritidsreiser (private besøk, hyttetur, fritid, etc.).
- Bostedsbaserte reiser for henting/levering av andre (barn og andre personer).
- Bostedsbaserte private reiser (innkjøp, legebesøk, private ærend).
- Arbeidsplassbaserte rundturer (besøk foretatt fra arbeidssted og tilbake) – **NY**.

Modellen for bostedsbaserte arbeidsreiser er nå sekvensielt estimert, noe som bl.a. innebærer at den nå er «nestet» med destinasjonsvalget over transportmiddelvalget. Dette skal presumtivt gi en del «stivere» destinasjonsvalg for arbeidsreisene enn det vi hadde i tidligere modellsystem. For arbeidsreisene forutsettes det fortsatt at de reisende (som ikke har firmabil) får et skattefradrag med en marginal skattesats på 40 %, med kilometersats 1,7 kr/km (transportmiddelavhengig). Skattefradraget gis kun der arbeidsreisen er så lang at kostnaden blir høyere enn bunnfradraget.

For de bostedsbaserte tjenestereisene regnes det som før med at det er arbeidsgiver/oppdragsgiver som betaler for reisene, og modellen er dermed estimert ut fra arbeidsgivers perspektiv. Kilometerkostnad for tjenestereiser er 3,5 kr/km, men riksregulativet for bilgodtgjørelse brukes kun hvis privatbil benyttes på tjenestereiser. Hvis firmabil benyttes, forutsettes privatøkonomiske kostnader å gjelde, og disse fratrekkes mva. for tjenestereisene.

Arbeidsplassbaserte rundturer representerer en ny «reisehensikt» i det regionale modellsystemet. Arbeidsplassbaserte reiser er definert som alle reiser som med utgangspunkt i arbeidsplassen også returnerer til denne etter at besøkene er gjennomført. Dette er i utgangspunktet en relativt inhomogen reisehensikt. Den er sammensatt av reiser med alle de 5 aggregerte reisehensiktene vi ellers har modeller for, men det disse reisene har til felles er at de hovedsakelig foregår i arbeidstiden. Av de arbeidsplassbaserte rundturene i datamaterialet er det ca. 60 % arbeidsrelaterte reiser og 40 % private.

Hovedintensjonen med å få med en modell for arbeidsplassbaserte reiser er å håndtere de mest kompliserte turkjedene som ligger i RVU-materialet på en bedre måte enn vi klarte tidligere. Tidligere håndtering tok høyde for turkjeder med tre delreiser. Siden vi likevel kalibrerer mot alle besøk innebærer dette at turer til og fra bostedet vil over-predikeres litt. Når vi skiller ut de arbeidsplassbaserte rundturene kan mange av de mest kompliserte turkjedene fjernes fra kalibreringsgrunnlaget for de øvrige modellene, og antall besøk som skal treffes vil gå ned (merk at både utreisen fra arbeidsplass og returen til arbeidsplass ble regnet som besøk i tidligere håndtering). Dermed vil omfanget av reiser til/fra bostedet også gå litt ned når vi introduserer en modell for arbeidsplassbaserte reiser.

MD-modellene er detaljert redegjort for i kapittel 3 i denne rapporten. Her beskrives alle variable og parameterestimater som inngår i modellene. I arbeidet har vi vært spesielt opptatt av å samordne segmenteringen i modellene så godt som mulig. I innledende estimeringsfaser ble det laget et felles rammeverk for dette. Dette rammeverket ble strammet noe inn etter hvert i arbeidet. Tabell A viser de endelige rammene for segmenteringen i MD-modellene. Med maksimal segmentering gir alternativene totalt 4800 forskjellige segmenter (typer mennesker). Ingen av modellene har dette maksimale antallet. Flest

segmenter finnes i modellen for private reiser som har 288, og færrest i modellen for tjenestereiser som bare har 24. Forenklet kan man si at hver modell beregner litt forskjellige valgsannsynligheter for hvert segment som inngår i modellen. Dette vil gi litt forskjellig fordeling av reiser på transportmåtene og destinasjonene for hvert segment. For variablene biltilgang, kjønn, alder og familietype ligger segmenteringen i form av antall personer i segmentene i sonebefolkningsfilen (som produseres av modellene for biltilgang). For variablene reisefølge og firmabil benyttes fordelingen i RVU-materialet¹ til å fordele sonebefolkningen videre inn i de aktuelle segmenter. Segmenteringen i hver enkelt delmodell framgår i rapportens kapittel 3.

Tabell A: Maksimale antall segmenter i MD-modellene.

Variabel	Antall	Inndeling
Biltilgang	5	Definert av segmenteringsmodell biltilgang
Reisefølge	2	Alene, to eller flere sammen
Periodekort	2	Ja/nei
Kjønn	2	Mann, kvinne
Alder	12	Intervaller starter på: 13, 16, 18, 20, 25, 35, 45,50, 55, 60, 67, 70.
Firmabil	2	Ja, nei
Familietype	5	Enslig, Enslig med barn, Par, Par med barn, Andre
Segmenter totalt	4800	

Forholdet mellom tidskoeffisient og kostnadskoeffisient gir i prinsippet uttrykk for trafikantenes betalingsvillighet for en marginal reduksjon av reisetiden og kan tolkes som verdi på reisetid. Denne betalingsvilligheten kan variere med bl.a. reiseformål, reisemiddel (pga. ulike komfortaspekter) og med inntekt, men vil også kunne være ganske situasjonsbetinget. Det man i beste fall får som resultat av en estimering, er et estimat på en gjennomsnittlig tidsverdi for utvalget.

Grunnen til å vurdere implisitte tidsverdier er at man har en formening om hva som kan være realistiske verdier for disse og det relative forholdet mellom dem,, selv om det ikke finnes noen klar fasit. Man har en del annen informasjon om tidsverdier til hjelp i denne vurderingen fra spesialstudier, f.eks. de norske tidsverdiundersøkelsene.

Tabell B gir en oppsummering av implisitte tidsverdier for de modellene som er implementert. Grå rader angir tidsverdier i kroner per time, mens de andre radene er faktorer som tidsverdiene skal multipliseres med. Unntaket er siste rad, som angir hvor mange minutter hver omstigning på en kollektivreise verdsettes som.

¹ I form av parametre i inputfilen modellfaktorer.txt

Tabell B: Implisitte tidsverdier i MD-modellene.

	Arbeid	Tjeneste	Fritid		Hente/lev.		Privat	Arb.plass- basert
			RD*	NVD	RD	NVD		
Tidsverdi bilfører, kr/t	105	240	77	99	63	111	90	225
Tidsverdi bilpassasjer, kr/t	105	240	77	99	63	111	90	225
Forsinkelse (Rush-fri flyt, faktor)	2,6	1,05						
Tidsverdi kollektiv (PT):								
Ombordtid, kr/t	61	180	36	65	55	55	53	125
Forhold tilbr.tid/ombordtid	2	2,9	2	2	2	2	2	2
Forhold ventetid/ombordtid	1,8	2,3	1,8	1,8	1,8	1,8	1,8	2,44
PT omstigning (minutter)	8	7	8	8	8	8	8	8

*NVD angir normalvirkedøgn, RD restdøgn (helger, ferier mm)

Gitt den usikkerhet som ligger både i datagrunnlaget for de estimerte modeller og i de "offisielle" tidsverdier fra tidsverdiundersøkelsen, så er vi av den oppfatning at de implisitte tidsverdier i de estimerte modeller er akseptable.

Modeller for turgenerering (TG)

Når det gjelder *turgenerering (TG)* i det nye modellsystemet så gjøres dette for de *bostedsbaserte reisene* på samme måte som tidligere. Det som modelleres er antall besøk foretatt fra eget hjem fordelt på 5 formål (i tillegg til skolareiser, som ikke blir med videre). Fordi dette gir relativt demografisk homogene grupper, blir turgenerering estimert med separate modeller for de samme 5 aldersgrupper som tidligere, dvs. 13-24 år, 25-34 år, 35-54 år, 55-66 år og 67+ år. Modelltypen som benyttes kalles Hurdle-Poisson, og med denne modelltypen får man beregnet forventet antall korte reiser som foretas i løpet av en dag og samtidig en fordeling på reisehensikter.

Turgenereringsmodellene er altså modeller for beregning av forventet antall besøk og fordeling av disse på de 5 bostedsbaserte reiseformålene. Variablene er hovedsakelig demografiske kjennetegn som kjønn, alder og husholdningstype, variabler som varierer med bosted som geografiske dummyvariable og logsummer fra MD-modellene. Logsummene reflekterer også en viss demografisk segmentering i tillegg til biltilgang og hensyntar ulik tilgjengelighet til/fra de enkelte sonene.

For de **arbeidsplassbaserte rundturene** ivaretas turgenereringen av to binomiske logit-modeller. Den ene beregner sannsynligheten for å møte opp på arbeidsstedet og den andre beregner den betingede sannsynligheten for å foreta en arbeidsplassbasert rundtur gitt at man har ankommet arbeidsstedet. Sannsynligheten for en arbeidsplassbasert rundtur blir produktet av disse to sannsynlighetene. Turgenereringen tar utgangspunkt i antallet arbeidsplasser i sonene. Variablene i modellene går på alder, kjønn og utdanningsnivå samt variable knyttet til fordeling av arbeidsplassene på næringskategorier, og andre kjennetegn ved arbeidsstedsjonen som gir en viss geografisk variasjon i turgenereringen.

Segmenteringsmodeller for biltilgang (BHFk)

Som tidligere er de nye modellene for biltilgang estimert som multinomiske logitmodeller hvor alternativene er:

- Dårlig biltilgang som passasjer (ikke førerkort, ikke bil i husholdet, DBTP).
- Full biltilgang som passasjer (ikke førerkort, bil i husholdet, FBTP).
- Dårlig biltilgang som fører (førerkort, ikke bil i husholdet, DBTF).

- Full biltilgang som fører (førerkort, minst like mange biler som personer med førerkort i husholdet, FBTF).
- God biltilgang som fører (førerkort, færre biler enn personer med førerkort i husholdet, GBTF).

Man kan selvsagt sette spørsmålsteget ved hvor godt multinomiske logitmodeller passer i en valgsituasjon som dette. Har man først investert ressurser i å anskaffe seg førerkort, så har man førerkortet nærmest resten av livet. Det vil neppe være slik at man leverer det tilbake hvis det blir etablert et veldig godt kollektivtilbud, eller hvis det blir vesentlig dyrere å bruke bil i et område.

Men dette blir en litt for bokstavelig tolking av valgsituasjonen når det gjelder segmenteringen etter biltilgang. I modellsystemet har vi ikke konkrete mennesker som bor her, arbeider der, har slektninger der, barn på skole der, etc., men en gitt folkemengde i typiske befolkningssegmenter som er geografisk fordelt, og er tilknyttet en fordeling på sannsynlige arbeidsplasser, og sannsynligheter for å foreta ærend av ulike typer avhengig av hva som befinner seg i sonene rundt bostedet og transporttilbudet til dem.

Når et tiltak iverksettes, kan det oppstå krefter i modellsystemet som f.eks. gjør det mer gunstig å bo i et nabolag for litt andre befolkningssegmenter enn tidligere; det er ikke så nødvendig, eller så gunstig, som før å ha førerkort i dette området, eller å ha én bil per person med førerkort i husholdet.

Vi observerer jo at biltilgangen i f.eks. sentrale områder i Oslo med veldig godt kollektivtilbud, og med relativt ugunstige forhold for eie og bruk av bil, er betydelig lavere enn landsgjennomsnittet eller andre deler av byområdet. En vesentlig lavere andel av befolkningen har førerkort, og ikke bare i de yngste aldersgrupper, og en vesentlig lavere andel har full biltilgang. Da er det gunstig å ha et modellsystem som gjør at man kan oppnå liknende fordelinger også andre steder hvis man iverksetter tiltak som gjør det mindre gunstig/nødvendig å bruke bil. Den multinomiske modellstrukturen i segmenteringsmodellene sørger for at slike effekter ivaretas.

Selv om de tre segmenteringsmodellene har samme struktur som tidligere, er det noen nyheter i form av nye eller mer presise variable. Tetthetsvariable som indikasjon på parkeringsulempe for bostedsparkering var også med tidligere, men nye data for arealbruk har gjort det mulig å konstruere en mer treffsikker deflator for tetthetsvariablene enn før. I tillegg er det beregnet en RVU-variable for «andelen bosatte i sonene som ikke har egen parkeringsplass ved bostedet». Denne har vist seg gunstig å ha med i interaksjon med tetthetsvariablene. Det er også innhentet nye og mer presise data for husholdsinntekter i de ulike segmentene og det er beregnet en ny geografisk multiplikator for disse som gjør at vi får en viss geografisk variasjon i segmentinntektene mellom soner.

Det er en tendens i RVU at soner med høy tetthet også har relativt høye inntekter. I estimeringen er det innført interaksjonsvariable mellom tetthet og inntekt, som demper inntektseffektene i modellen i soner med høye tettheter. Det at interaksjonsvariablene blir negative for eksempelvis full biltilgang, medfører at soner med høye tettheter og høye inntektsnivåer ikke nødvendigvis skal ha den beste biltilgangen.