

Simultan myldretids-trafikmodel for Hovedstadsområdet

- Teoretiske overvejelser og praktiske erfaringer ved opbygning af en simultanmodel

Af Morten Agerlin, Anders Nyvig A/S

Trafikmodel for Hovedstadsområdet

Anders Nyvig A/S har for en lang række interessenter i Hovedstadsområdet opstillet en trafikmodel, som beskriver person- og erhvervsbiltrafik, kollektiv trafik samt cykeltrafik i Hovedstadsområdet.

Siden den første model blev opstillet i 1990, er der løbende sket en udvikling af modellen. Modellen er i flere trin blevet gjort mere detaljeret i kraft af finere zoneinddelinger og mere fintmaskede trafiknet.

Modellens udvikling og brug følges af den såkaldte brugergruppe, som udgøres af følgende interessenter:

- Energi- og Miljøministeriet
- Trafikministeriet
- Københavns og Frederiksberg kommuner
- Københavns, Roskilde og Frederiksborg amter
- Hovedstadsregionens Statistikkontor (HSK)
- Ørestadsselskabet og A/S Øresund
- Vejdirektoratet
- HT og DSB
- Institut for Planlægning, DTU
- Anders Nyvig A/S

De hidtidige versioner af modellen har beskrevet døgntrafikken og har været opbygget som flertrinsmodeller.

Det seneste skud på stammen er en model for myldretidstrafikken. Denne model er

opbygget som en simultanmodel, hvor turproduktion, destinationsvalg og transportmiddelfordeling sker simultant. Endvidere er der for at beskrive myldretidssituationen indføjet kapacitetsbegrænsninger i vejvalget.

Trafikmodelstrukturer

En traditionel 4-trins trafikmodel er normalt defineret ved nedenstående 4 processer:

1. **Turgeneration og -attraktion:** Totalt antal rejser til og fra den enkelte zone beregnes.
2. **Fordeling på rejsemål:** Ovenstående totaler fordeles ud på mål-/udgangszone hvorved der dannes en total rejsematrix.
3. **Transportmiddelfordeling:** Den totale rejsematrix fordeles på de i modellen indgående transportmidler.
4. **Rutevalg:** For hvert af transportmidlerne fastlægges ruten mellem alle zonepar.

Begrebet simultanmodel dækker over at flere eller færre af ovenstående processer udføres simultant i stedet for i den sekvens som er angivet ved nummereringen.

I Hovedstadsmodellens myldretidsmodel er det de tre første processer, som udføres simultant. Da myldretidsmodellen benytter kapacitetsbegrænsninger i vejnettet indgår simultanmodellen og rutevalget i en iterativ proces, som først stopper når simultanmodellens matricer passer til rutevalgsmode llens kapaciteter

på vejstrækningerne. Her er der altså heller ikke tale om en traditionel sekventiel proces.

Når der i det følgende tales om simultanmodeller, er det baseret på erfaringer ved ovenstående definition - altså en simultan udførelse af turgeneration og -attraktion, fordeling på rejsemål samt transportmiddelfordeling.

Fordele/ulemper ved simultanmodeller

Nedenfor er i kort form listet en række fordele og ulemper ved at benytte simultanmodeller i forhold til mere traditionelle sekventielle/fletrins modeller.

Fordele ved simultanmodeller:

- Alle variable får indflydelse på både rejseantal, rejsemål og transportmiddelvalg.
- Simplere og hurtigere beregningsgang når modellen først er opstillet - der er færre trin som skal beregnes.
- Når modellen er opstillet, er man uafhængig af diverse fremskrivninger af grundlagsmatricer med deres eventuelle fejl og mangler. Man kan direkte på baggrund af rejse- og byplanvariable i en fremtidig situation estimere antallet af rejser mellem modelzoner.
- Modellen kan bruges direkte for nye zoner, som ikke har trafik i dagens situation - f.eks. en Ørestad.

Ulemper ved simultanmodeller:

- Alle variable påvirker alt - det kan være svært at gennemskue og realitetsvurdere årsager og virkninger i modelresultaterne.

- Der er mange frihedsgrader i modellen - man forsøger at beskrive meget med få modeludtryk. Modellen kan 'eksplodere' i uheldige tilfælde og give urealistisk høje tal.

Fordele/ulemper ved forskellige typer simultane modeller

Principielt kan en hvilken som helst formeludtryk benyttes til opstilling af en simultanmodel. Kendte eksempler er gravitationsmodeller og forskellige varianter af såkaldte direkte efterspørgselsmodeller, hvor antallet af ture T er en funktion af socioøkonomiske variable S og transportudbudsvariable U :

$$T = f(S) * g(U).$$

Nedenfor er beskrevet fordele og ulemper ved to direkte efterspørgselsmodeller som har været testet i forbindelse med opbygningen af myldretidsmodellen for Hovedstadsområdet.

Baumol-Quandt modeltype:

$$T = K \cdot \prod_x S_x^{a_x} \cdot \prod_y U_y^{b_y}$$

altså en generel produktform af de uafhængige variable opløftet til en potens.

Log-lineær model:

$$T = K \cdot \prod_x S_x^{a_x} \cdot \prod_y e^{b_y \cdot U_y}$$

hvilket kan omformuleres til:

$$T = K \cdot \prod_x S_x^{a_x} \cdot \exp\left(\sum_y b_y \cdot U_y\right)$$

altså en eksponentiel transformation af et lineært udtryk.

Principielt er det kun den eksponentielt transformerede del af sidstnævnte udtryk, som i statistikken betegnes en log-lineær model.

Fordele ved log-lineær model

- Ved kalibrering af modellen estimeres det observerede totale rejseantal altid korrekt, dvs. man er sikret mod generelle over-/underestimationer i forhold til observerede rejseantal.
- Mere teoretisk, statistisk korrekt idet modellen som forudsætning har heltallige og poisson-fordelte afhængige variable.
- Let at estimere

Ulemper ved log-lineær model

- Passer ikke ind i gængs turrate-tankegang
- Komplicerede elasticitets-udtryk

Fordele ved Baumol-Quandt model

- Passer ind i gængs turrate-tankegang
- Simpelt elasticitets udtryk (tilnærmet lig α 'er og β 'er)

Ulemper ved Baumol-Quandt model

- Ved brug i Hovedstadsmodellen førte den til en konsekvent overestimation af kollektivrejser.
- Kan ikke håndtere variable der kan antage værdien nul med en negativ potens som f.eks. antal omstigninger ved en kollektivrejse eller variable defineret ved differencer.
- Modellen ses ofte løst ved at linearisere ved en logaritmering af formeludtrykket. Dette bør undgås, idet det medfører en konsekvent underestimation, da logarit-

men/eksponentialfunktionen er asymmetriske transformationer.

Styring af modelestimationen

De fleste statistiske modeller er baseret på forudsætninger om uafhængighed mellem modellens variable, deraf ofte kaldt modellens uafhængige variable.

I virkelighedens verden er situationen desværre ofte den modsatte - en lang række af de variable man ønsker skal indgå i modellen er indbyrdes afhængige. Det kan både være internt indenfor ét transportmiddel, hvor f.eks. rejsetid i bil ofte vil være stærkt korreleret med kørselsomkostninger i bil og det kan være de enkelte transportmidler imellem som f.eks. rejsetid i bil og rejsetid i bus.

Måske kunne man leve med denne overtrædelse af nogle af modellens teoretiske forudsætninger, men man vil ofte komme ud for, at man ikke kan få de ønskede fortegn på modellens estimerede konstanter. F.eks. vil modellen måske lægge hele afstandsafhængigheden i en rejsetidsvariabel, mens en omkostningsvariabel, som er stærkt korreleret med rejsetiden derfor får en lettere tilfældig vægt i modellen og muligvis et forkert fortegn.

En kendt metode til at undgå korrelationer for variable mellem transportmidler, er at vælge et transportmiddel som basis og dernæst beregne de øvrige i forhold hertil, enten i form af differencer eller i form af brøker.

Ofte udpeges bilen som basis og der kan herefter defineres pseudovariabel som f.eks. rejsetid i bus minus rejsetid i bil eller som rejsetid i bus divideret med rejsetid i bil.

Som regel vil man herudover blive nødt til at benytte en variabel som en total målestok. Hvis man f.eks. definerer en variabel som rejsetid i bus divideret med rejsetid i bil, har man brug for en variabel der modellerer den totale tids-/afstandsafhængighed, som f.eks. bilrejsetiden eller luftlinieafstanden.

Indbyrdes afhængigheder indenfor det samme transportmiddel er sværere at komme udenom på en teoretisk forsvarlig måde.

Her har vi i Hovedstadsmodellen også benyttet os af at definere pseudovariabel. Vi har således for hvert transportmiddel defineret en servicevariabel baseret på rejsetider og en omkostningsvariabel baseret på omkostningsvariable.

På basis af tilgængelige analyser, modeller og adfærdsstudier har vi sammenvejet de forskellige bidrag i f.eks. rejsetid til én variabel. F.eks. defineres servicevariabelen for kollektiv trafik som:

$$K_{\text{service}} = - (\text{bustid} + 0.65 * \text{togtid} + 1.75 * \text{gangtid} + 2.5 * \text{ventid} + 5 * \text{omstign} + 1 * \text{skjulven})$$

Herved kommer man teoretisk set ikke udenom afhængighederne, men i praksis "ses" afhængighederne ikke når parametrene estimeres og det er betydeligt lettere at sikre sig de ønskede fortegn på de enkelte oprindelige variable på denne måde.

Når man på denne måde "øver vold" på de estimerer den statistiske model af sig selv ville sætte på de oprindelige variable, må man være forberedt på at få en model, som tilpasser kalibreringsdata dårligere, men man får forhåbentlig en anvendelig model, som reagerer "rigtigt" på de indgående variable.

Inddragelse af erfaringer fra andre modeller eller adfærdsstudier

Ovennævnte metode er også en måde at inddrage erfaringer fra andre modeller og adfærdsstudier på eller ligefrem overføre delmodeludtryk fra andre steder - med de farer det indebærer.

Erfaringer fra udvikling af den simultane myldretidsmodel

Til simultanmodellen er anvendt en log-lineær model. Firmaet har selv i 1994 afprøvet en sådan model i en trafikmodel for Odense, men derudover kender vi ikke til anden brug af denne modeltype indenfor trafikmodeller.

Modellen er kalibreret på baggrund af observerede rejsemønstre i 1992. Modellen er således kalibreret på baggrund af geografisk variation i variablene og ikke en historisk variation, som ville være mere korrekt - sådanne data er imidlertid ikke tilgængelige. Modellen beskriver bil- og kollektivtrafik (bus og tog).

I første omgang var modellen er baseret på nedenstående to formeludtryk, et for hvert af de to modellerede transportmidler:

$$T_{ij}^{\text{bil}} = \exp(a^{\text{bil}} + b^{\text{bil}} * \text{GENZONE}_i + c^{\text{bil}} * \text{ATTZONE}_j + d^{\text{bil}} * \text{BILVAR}_{ij} + e^{\text{bil}} * \text{KOLLVAR}_{ij})$$

$$T_{ij}^{\text{koll}} = \exp(a^{\text{koll}} + b^{\text{koll}} * \text{GENZONE}_i + c^{\text{koll}} * \text{ATTZONE}_j + d^{\text{koll}} * \text{BILVAR}_{ij} + e^{\text{koll}} * \text{KOLLVAR}_{ij})$$

hvor T_{ij}^k er antal ture fra zone i til zone j med transportmiddel k, $a^k - e^k$ er konstanter specifikke for transportmiddel k, GENZONE_i er zonevariable for zone i, ATTZONE_j er zonevariable for zone j, BILVAR_{ij} er rejsevariable knyttet til bil-

rejser mellem zone i og j og $KOLLVAR_{ij}$ er rejsevariable knyttet til kollektivrejser mellem zone i og j.

Idéen er her, at $KOLLVAR$ i bilmodellen skal beskrive den reduktion i bilrejserne, som konkurrencen fra kollektivtrafikken bidrager med og vice versa. Imidlertid er flere af variablene i $KOLLVAR$ stærkt korrelerede med rejsevariable i $BILVAR$, hvorfor der er problemer med at opnå de korrekte fortegn på variablene.

I denne modelformulering kommer man ud over disse korrelationer ved først at definere 4 pseudovariabeler beskrivende rejseservice for bil og kollektiv og omkostninger for bil og kollektiv. Dernæst benyttes differencen mellem rejseservice for bil og kollektiv og differencen mellem omkostninger for bil og kollektiv som de egentlige modelvariable.

Derved udgår man korrelationen mellem f.eks. bilrejsetid og kollektivrejsetid. Til gengæld mangler der en målestok for rejse længde, hvorfor luftlinieafstanden indføres i modellen.

Ved estimation af modellen parametre får luftlinieafstanden imidlertid en meget stor elasticitet, mens de modelvariable, som det kunne være interessant at ændre på, f.eks. bilrejseomkostninger får meget små elasticiteter.

Modellsystemet består af to selvstændige modeller - en for bilrejser og en for kollektivrejser - det eneste der binder dem sammen er de fælles variable.

Ved sammenstillingen af pseudovariablene er det imidlertid tilstræbt at de resulterende elasticiteter for bil- og kollektivrejser passer sammen via

krydselasticiteterne således at de enkelte variables elasticitet overfor det totale antal rejser får det korrekte fortegn - således at f.eks. en øgning af benzinprisen giver færre bilrejser og flere kollektivrejser, men at væksten i kollektivrejser er mindre end eller lig med faldet i antallet af bilrejser.

Dette er tilstræbt på det totale matrix-niveau, men det er ikke muligt at sikre denne egenskab overholdt for alle zone-relationer, uanset forholdet mellem antal bil- og kollektivrejser.

På grund af ovennævnte problemer er der søgt benyttet en lidt anderledes modelstruktur, som tager hensyn til nogle af disse svagheder. I stedet for at lade $BILVAR$ repræsentere bilrejserne i kollektivmodellen og derved nærmest opbygge to modeller for bilrejser baseret på bilrejsevariablene, er det valgt at koble de to modeludtryk direkte sammen:

$$T_{ij}^{bil} = \exp(a^{bil} + b^{bil} * GENZONE_i + c^{bil} * ATTZONE_{i+} + d^{bil} * BILVAR_{ij} + e^{bil} * T_{ij}^{koll})$$

$$T_{ij}^{koll} = \exp(a^{koll} + b^{koll} * GENZONE_i + c^{koll} * ATTZONE_{i+} + d^{koll} * KOLLVAR_{ij} + e^{koll} * T_{ij}^{bil}),$$

hvor e^{bil} og e^{koll} forventes at være negative.

Det modellerede antal kollektivrejser indgår altså nu direkte i udtrykket for det modellerede antal bilrejser og vice versa.

Estimationen af parametrene bliver nu en iterativ proces, hvor bilmodellen først estimeres og resultaterne heraf benyttes ved estimationen af kollektivmodellen, hvis resultater igen benyttes ved næste trin af estimationen af bilmodellen etc.

Når parametrene er fundet og modellen skal benyttes vil en modelkørsel herefter bestå i at løse $280^2 = 78.400$ sæt af 2 ulineære ligninger med 2 ubekendte - godt vi har en computer!

Ved estimationen af f.eks. bilmodellen indgår der nu kun bilmodellens variable og problemerne med korrelation mellem f.eks. bil- og kollektivrejsetid forsvinder. Da bilrejsevariablene nu indgår direkte, behøves luftlinieafstanden ikke længere til at beskrive afstandsafhængigheden og modellernes elasticiteter overfor rejse- og omkostningsvariablene bliver større.

Da de 2 ulineære ligninger ikke lader sig løse eksplicit, men må løses numerisk, er det besværligt at bevise matematisk, at det ønskede forhold vedrørende elasticiteter overfor det totale antal rejser automatisk er opfyldt ved denne modelformulering, men det viser sig i praksis og gennem numeriske kontrolberegninger.

Træerne vokser imidlertid ikke ind i himlen - i praksis er antallet af bilrejser og antallet af kollektivrejser indbyrdes positivt korrelerede, således at konstanterne e^{bil} og e^{kol} har en tendens til at blive positive.

Den eneste mulige løsning har været manuelt at fastsætte e^{bil} og e^{kol} til passende negative værdier og på den måde manuelt fastlægge krydselasticiteterne i forhold til de direkte elasticiteter. Dette er naturligvis igen at "øve vold" på statistikken, men i bund og grund nok mere acceptabelt end manuelt at fastlægge de direkte elasticiteter som i formeludtrykket for pseudovariablen KSERVICE som vist ovenfor.

Myldretidsmodellens formåen

Det har aldrig været intentionen at modellen på baggrund af de relativt simple ind-data skulle kunne beskrive rejser mellem alle zonepar "korrekt". Modellen er en generalisering og har en tendens til at udglatte rejsemønstret mere end det ses i de observerede matricer, hvor faktorer som ikke forklares med modellens variable kan betyde helt forskellige rejseantal til to nabozoner.

Imidlertid formår modellen på rimelig vis at modellere det totale antal rejser ind og ud af zonerne, hvilket viser sig at være tilstrækkeligt til at sikre en rimelig tilpasning til den observerede trafik på strækninger.

Det har dog vist sig, at selv indenfor Hovedstadsområdet er der for store regionale forskelle til at det kan rummes i ét modeludtryk. Det har derfor været nødvendigt at inddele området i 5 regioner, som har fået hvert sit konstantled a^{bil} og a^{kol} (jvnf. modeludtryk på forrige side), mens det øvrige modeludtryk har kunnet fastholdes for alle 5 regioner.