



Modellering av järnvägstransporter

En översikt kring datakällor

Kristofer Odolinski
Hanna Lindgren

Modellering av järnvägstransporter

En översikt kring datakällor

Kristofer Odolinski

Hanna Lindgren

Författare: Kristofer Odolinski, VTI, Hanna Lindgren, VTI
Diarienummer: 2017/0535-7.4
Omslagsbilder: Mostphotos
Tryck: VTI, Linköping 2017

Förord

Trafikanalys har fått ett regeringsuppdrag (N2017/03480/TS) om att utreda hur kunskapen om järnvägstransporter kan förbättras. Som en del i detta har Statens väg- och transportforskningsinstitut (VTI) fått i uppdrag av Trafikanalys att genomföra en studie om datakällor för modellering av järnvägstransporter. Uppdraget är att inventera forskningsstudier av relevans på området, dokumentera transportmodeller, datakällor, metoder och slutsatser, samt sammanställa resultatet i en rapport. Kristofer Odolinski och Hanna Lindgren har skrivit notatet. Magnus Johansson har granskat ett utkast av notatet och bidragit med värdefulla synpunkter.

Stockholm, november 2017

Kristofer Odolinski

Kvalitetsgranskning

Intern peer review har genomförts 24 november 2017 av Magnus Johansson. Kristofer Odolinski och Hanna Lindgren har genomfört justeringar av slutligt rapportmanus. Forskningschef Mattias Haraldsson har därefter granskat och godkänt publikationen för publicering 1 december 2017. De slutsatser och rekommendationer som uttrycks är författarnas egna och speglar inte nödvändigtvis myndigheten VTI:s uppfattning.

Quality review

Internal peer review was performed on 24 November 2017 by Magnus Johansson. Kristofer Odolinski and Hanna Lindgren has made alterations to the final manuscript of the report. The research director Mattias Haraldsson examined and approved the report for publication on 1 December 2017. The conclusions and recommendations expressed are the authors' and do not necessarily reflect VTI's opinion as an authority.

Innehållsförteckning

Sammanfattning	7
Summary	9
1. Inledning	11
1.1. Syfte	11
1.2. Metod	11
1.3. Avgränsningar	11
1.4. Disposition	12
2. Nationella transportmodeller	13
2.1. Sampers	13
2.2. Samgods	17
2.3. (SAM)VIPS	20
2.4. Transtools och transportmodeller i Storbritannien och Norge	22
3. Forskning kring metoder och nya indata	25
3.1. Utveckling av transportmodeller	25
3.2. Nya indata	26
3.2.1. Satellitnavigationssystem	27
3.2.2. Mobilnätdata	28
3.2.3. Radio Frequency Identification (RFID), vägning, och Automatic Passenger Counting (APC). 29	
3.2.4. Kombination av datakällor	30
4. Diskussion och slutsatser	32
Referenser	35

Sammanfattning

Modellering av järnvägstransporter – en översikt kring datakällor

av Kristofer Odolinski (VTI) och Hanna Lindgren (VTI)

Information om flöden av gods och passagerare är viktig för att utveckla kunskapsunderlaget om järnvägstransporter. Förutom att ge en nulägesbeskrivning kan informationen användas i transportmodeller för att genomföra prognoser av flöden, bland annat i syfte att planera och utvärdera åtgärder såsom infrastrukturinvesteringar, införande av olika begränsningar eller förändrade styrmedel. I Sverige består transportstatistiken främst av aggregerade data, vilket innebär att prognoser av transportflöden och andra typer av skattningar och analyser blir osäkrare. Det finns därför anledning att beskriva hur modelleringen av järnvägstransporter ser ut idag, samt dess utveckling och möjligheter att kombinera olika (och nya) datakällor för att ta fram den information om järnvägstransporter som efterfrågas.

Sampers och Samgods är de svenska nationella trafikslagsövergripande modellerna för person- respektive godstransporter. Tre olika typer av indata används i Sampers: data som beskriver resebeteenden, data som beskriver trafiknätet och utbudet av kollektivtrafik samt data för hur attraktivt det är att resa till de områden (eller zoner) som finns i modellen. Resultatet från Sampers efterfrågemodell är passagerarflöden per färdmedel och destination. I fördelningen av tågpassagerare används nät- och linjekodning som motsvarar trafikeringen för nulägesåret (till prognoserna används prognostidtabeller). Detta ger exempelvis information om antalet resor per länk under ett år (det vill säga mellan stationer eller noder på linjer) samt antal på- och avstigande på de olika stationerna. Dessa resultat behöver valideras, vilket bland annat Trafikverkets regioner gör med hjälp av resandestatistik.

Indata i Samgods är information om transportnätverket, kostnadsdata, regler, uppgifter om produktion och konsumtion samt varuflöden. Information om varuflöden kommer från utrikeshandelsstatistiken och varuflödesundersökningen. Uppgifterna om produktion och konsumtion används tillsammans med uppgifter om varuflöden för att ta fram en beräknad godsvolym för 34 varugrupper som ska transporteras mellan 464 zoner (varav 290 är kommuner i Sverige). För att ta fram transportdata för järnvägstrafiken i form av godsflöden och antal tåg fördelat på olika bandelar på järnvägsnätet, används modellverktyget Bangods. Detta sker bland annat via information om trafikering, som dock inte innehåller fullständig information om typ av godståg eller om typ av gods och godsmängd per tåg. I nuläget består utdata av godsflöden under ett helt år uppdelat på antal fjärr-, lokal-, system-, malm-, och kombitåg, 12 varugrupper och järnvägsnätets olika bandelar (cirka 300 stycken).

Modelleringen av järnvägstransporter (som oftast ingår i trafikslagsövergripande modeller) går mot mer disaggregerade ansatser. På persontransportsidan går utvecklingen mot aktivitetsbaserade modeller, medan modellering av logistikbeslut på företagsnivå blir vanligare i godsmodeller. Denna utveckling innebär att det finns ett behov av disaggregerade data, med en finare detaljnivå i både tid och rum. Samtidigt har den snabba teknikutvecklingen inneburit nya möjligheter för insamling av disaggregerade data, framförallt genom låga priser och bättre funktion hos de tekniska enheter som behövs för att registrera och kommunicera information om positionering. Exempel på datakällor är satellitnavigationssystem (Global Positioning System, GPS), mobilnätdata, Radio Frequency Identification (RFID) eller andra typer av sensorer (för exempelvis tågets vikt eller antal på- och avstigande).

De olika datakällorna till transportmodelleringen har sina för- och nackdelar; en och samma datakälla kan inte täcka det behov av information som behövs. En kombination av de olika datakällorna kan därför vara en lösning. Exempelvis kan GPS-data och mobilnätdata vara bra komplement då den förra har en hög detaljnivå, men större vita fläckar jämfört med den senare datakällan. Ett generellt problem

med den automatiserade datainsamlingen är dock att den oftast fångar en del av resorna; även om den ger betydligt mer detaljerad information om resornas position och tid jämfört med traditionella undersökningar, så finns det ett snedvridet urval. Därutöver saknas det ofta socioekonomisk information samt information om syftet med resan, även om det via rörelsemönster från GPS- eller mobilnätdata är möjligt att dra slutsatser om jobbets och hemmets position samt koppla denna information till (åtminstone övergripande) socioekonomiska data. Oavsett detta kan information om passagerarflöden och godsflöden vara användbara i kalibrering och validering av resultat. Detektorer för tågens vikt kan exempelvis användas för att ta fram information om antal passagerare på varje tåg (potentiellt mellan alla stationer om detektorn finns på tåget, eller om tillräckligt många stationära detektorer finns utplacerade). Även om godstågets vikt inte ger information om godstyp, så kan den informationen vara användbar i kalibreringen av Samgods och ge en tillförlitligare beskrivning av godsflöden fördelat på olika tåg och bandelar.

Det finns med andra ord en stor potential att ta fram mer disaggregerade data över transportflöden med hjälp av ny teknik och datainsamlingsmetoder.

Summary

Modelling railway transport – an overview of data sources

by Kristofer Odolinski (VTI) and Hanna Lindgren (VTI)

Information on freight and passenger flows is vital for advancing knowledge on railway transport. In addition to providing a description of the current use of railway transportation, the flows of freight and passengers can be used in transport models to make transportation forecasts. Such forecasts can, for example, be used to evaluate infrastructure investments, or to analyze the impact of new restrictions or new policies. The transport statistics in Sweden are, however, quite aggregate, which makes forecasts and other types of estimations rather uncertain. There is therefore reason to describe the current modelling of railway transport, as well as its development and potential in combining different (and new) data sources to provide the information on railway transport that is requested.

Sampers and Samgods are the national models in Sweden for freight and passenger transports, respectively. Three types of data are used in Sampers: data describing travel behavior, information about the transport network and public transport services, as well as attraction measures for trips to the different geographical zones in the model. The output from the demand model are passenger flows for each transport mode and destination. Train passengers are distributed on the railway network based on current traffic flows, while anticipated timetables are used in the forecasts. This generates information on the number of trips for each railway link during a year (that is, between stations or nodes on railway lines), as well as the number of embarking and disembarking passengers on the railway stations. These results need validation, which *inter alia* is performed by regions within the Swedish Transport Administration, using passenger statistics.

Data inputs for Samgods are information on the transport network, cost data, regulations, production and consumption data and goods flows. The data on goods flows are gathered from foreign trade statistics and the national commodity flow survey. The production and consumption data are used together with the goods flows data to calculate the transported volume of 34 different commodity types between 464 zones (of which 290 are municipalities in Sweden). A separate tool called Bangods is used to generate rail freight transport on the different track sections comprising the railway network. This is based on rail traffic data, which however lacks information on the type of freight trains as well as commodity type and volume of goods per train. At present, the output for rail freight transport consists of goods flows during a year on the different track sections (about 300), segmented into different types of freight trains and 12 groups of commodities.

Transport modelling approaches are becoming more disaggregate. More specifically, there is a development towards activity-based models for passenger transport, while modelling of logistic choices at company level is becoming more common for freight transport. This development requires more disaggregate data, in both time and space. The rapid technological development has provided new opportunities in collecting disaggregate data, notably through low prices and improved performance of the technical units required to record and communicate information on the positioning of people, goods or vehicles. Examples on data sources are data from radionavigation-satellite services (Global Positioning System, GPS), cell phone data, Radio Frequency Identification (RFID) or other types of sensors (for train weight or number of embarking and disembarking passengers).

There are pros and cons with these different data sources in transport modelling; one type of data source cannot cover the information needs. A combination of data can therefore be a solution. For example, GPS data and cell phone data can be good complements, as the former have a higher level of detail while the latter have larger blind spots. Still, a general problem with the automatic collection of data is that only a sub-sample of the trips are captured; even though it generates more detailed

information on the position and timing of trips compared to traditional travel surveys, it is still based on a sub-sample of the trips made, which can create a selection bias. Moreover, this type of data lacks information on socio-economic aspects and the purpose of the trip. However, information on work and home location can be derived from mobility patterns based on GPS or cell phone data, which in turn can be linked to socio-economic information (at least at an aggregate level). Nevertheless, information on passenger and freight flows can be useful in the calibration and validation of the transport models. For instance, train payload measurements can be used to derive information on the number of passengers on each train (potentially between all stations if the train has a weighing capability, or if enough stationary weight detectors are installed on the infrastructure). Moreover, even though the weight of a freight train does not reveal the commodity type, it can be used in the calibration of Samgods and provide a more accurate flow of goods on different trains and track sections. In other words, new technology and data collection methods provide a significant potential for generating more disaggregate data on transport flows.

1. Inledning

Statistik om järnvägstransporter utgör en central del i kunskapen om transportsystemet och dess funktion. I Sverige består denna transportstatistik främst av aggregerade data. Företagens behov av sekretess anges ofta vara en anledning till varför en mer detaljerad nivå på statistiken inte kan redovisas. Det innebär emellertid att prognoser av transportflöden och andra typer av skattningar och analyser blir osäkrare. Mer detaljerad information om järnvägstransporter skulle exempelvis kunna användas för att göra säkrare utvärderingar (ex ante och ex post) av olika åtgärder i transportsystemet, eller kombineras med trafikdata för att bättre kunna analysera effekter av transportsystemets tillförlitlighet.¹ Information om flöden av gods och passagerare på en mer disaggregerad nivå i både tid och rum är med andra ord nödvändigt för att utveckla kunskapsunderlaget om järnvägstransporter.

När det saknas statistik finns det möjlighet att imputera värden, dvs. fylla i de luckor där det saknas värden. En nära besläktad metod är prognoser (kan även ses som specialfall av imputation), vilket är en förutsägelse om framtida värden. Båda dessa metoder faller inom det vidare begreppet matematisk modellering, vilket är en representation av verkligheten (och/eller av en teori) som delvis består av en eller flera ekvationer och används för att lära oss något om världen (Frigg & Hartman 2012). En transportmodell är en sådan representation av verkligheten, och används exempelvis för att planera och utvärdera olika typer av åtgärder såsom infrastrukturinvesteringar, införande av begränsningar, förändringar av styrmedel, eller för att analysera effekter av störningar. Det finns en rad olika modeller som varierar med avseende på geografisk omfattning, typ av transport (gods- eller persontransport), samt vilka datakällor och metoder som används. Modellerna är dessutom under ständig utveckling.

Då transportmodeller är viktiga för att ta fram kunskapsunderlag om transportsystemet - särskilt i frånvaron av detaljerad statistik över järnvägstransporter - finns det anledning att närmare beskriva modelleringen av järnvägstransporter, dess utveckling och möjligheter att kombinera olika (och nya) datakällor för att få fram den information som av olika anledningar efterfrågas.

1.1. Syfte

Studiens syfte är att inventera forskningsstudier av relevans på området, dokumentera transportmodeller, datakällor, metoder och slutsatser. I synnerhet beskrivs forskning om modellering av järnvägstransporter där andra typer av data än transportdata används, till exempel trafikdata, demografiska och ekonomiska data.

1.2. Metod

Metoden som har använts är en litteraturstudie. Litteratursökningen inkluderar trafikslagsövergripande transportmodeller då modellering av järnvägstransporter (i syfte att generera transportdata) ofta är integrerad i nationella och internationella transportmodeller som beskriver transportflöden för flera trafikslag. Studier som enbart behandlar ett annat trafikslag än järnvägstransporter ingår också i sökningen eftersom metoder för att ta fram transportdata får anses vara relativt överförbara. Sökningen har främst utförts i databaser för akademiska artiklar i tidskrifter och konferensartiklar, men även rapporter och promemorior är en del av underlagsmaterialet.

1.3. Avgränsningar

Det finns ett otal transportmodeller som har utvecklats runtom i världen. I beskrivningen av nationella och internationella transportmodeller görs därför en avgränsning till de modeller som används i Sverige och till viss del i Europa. Dock utförs en bredare sökning efter forskning kring aktuella

¹ Notera att transporter avser förflyttning av gods och passagerare, medan förflyttning av fordon benämns som trafik.

metoder och datakällor som används för att ta fram transportdata för ett eller flera trafikslag i mer avgränsade områden, såsom städer eller regioner i ett land.

1.4. Disposition

I avsnitt 2 beskrivs modeller för person- och godstransporter som används för att analysera järnvägstransporter och transporter med övriga trafikslag, i Sverige och till viss del Europa. Forskning om nya datakällor och metoder för att ta fram transportdata behandlas i avsnitt 3. Notatet avslutas med en diskussion och slutsatser i avsnitt 4.

2. Nationella transportmodeller

I Sverige används idag i huvudsak två olika modellsystem för analyser av transporter; Sampers för persontransporter och Samgoods för godstransporter. Många länder i Europa har sina egna transportmodeller, samtidigt som det har utvecklats en europeisk transportmodell som täcker 42 länder. Nedan ges en översiktlig beskrivning av modeller i Sverige (och i viss mån modeller i andra europeiska länder), deras användningsområden, samt vilka metoder och indata som används.

2.1. Sampers

Sampers-systemet används för att beräkna de kort- och långsiktiga effekterna av kostnadsförändringar och förändringar i trafikutbudet på personresor i Sverige och mellan Sverige och Danmark (Själland). En översiktlig bild av de olika modellerna i Sampers-systemet och den omvärldsindata som används visas i Figur 1. Sampers är uppdelad i fem regionala modeller för resor under 100 km, en nationell modell för resor över 100 km och en internationell modell där resan startar eller slutar i Sverige (Algers et al. 2009). De regionala modellerna har sammanlagt ca 10 000 zoner som utgör start- och målpunkter, medan den nationella har ca 700 områden. Den internationella modellen inkluderar ca 200 områden utanför Sverige. I de regionala modellerna skattas sex olika slag av ärenden (arbetsresor, tjänsteresor, resor till skola, besöksresor, fritidsresor och övriga resor) och för varje ärende modelleras sex olika färdstätt: bil som förare, bil som passagerare, buss, tåg, cykel och gång. I den nationella modellen finns två ärenden (privat resa och tjänsteresa), och fyra färdstätt: bil, buss, tåg och flyg (Algers et al. 2009).

Sampersmodellen är en variant av en så kallad fyrstegsmodell som behöver tre typer av data: data som beskriver resebeteenden, data som beskriver trafiknätet och utbudet av kollektivtrafik samt data för hur attraktivt det är att resa till de områden (eller zoner) som finns i modellen (Trafikanalys, 2016a). I det första steget beräknas antalet resor som förväntas ske under en dag, i det andra steget beräknas mellan vilka punkter resorna sker och i det tredje steget beräknas vilka färdstätt som används. Dessa tre steg genererar start- och målpunktsmatriser (OD-matriser) som innehåller information om antal resor fördelat på olika färdstätt. I Sampers beräknas emellertid steg 1 till 3 simultant och inte sekventiellt. I det fjärde steget avgörs ruttvalet för resorna. Mer specifikt används ett nätutläggningsprogram för att fördela ut trafiken från OD-matriserna på transportnätverket utifrån principen om att minimera de generaliserade kostnaderna, dvs. transportkostnader inklusive tidskostnader (Trafikanalys, 2016). Resultaten från Sampers utgör också indata till Samkalk person, som är en samhällsekonomisk kalkylmodell integrerad med Sampers. Vi fokuserar här på indata till Sampers (se Figur 1). För den nuvarande bastrafikprognosen (publicerad 2016-04-01) har 2014 använts som basår, det vill säga att nulägesberäkningen gör med priser, kostnader, trafikutbud, etc. för år 2014. Till prognoserna för år 2040 och 2060 används bland annat bedömningar av ekonomisk utveckling från den senaste Långtidsutredningen och prognoser för befolkningsutveckling från SCB. Nedan presenteras ett urval av indata.

Resvanedata

Den indata som beskriver resebeteende kommer från resvaneundersökningar. Dessa undersökningar ger information om genomförda resor och socioekonomisk information om resenärerna. I den långväga (nationella) modellen används data från resvaneundersökningen från 2005/2006, medan de regionala modellerna använder data från resvaneundersökningar för perioden 1994 till 2000 (Trafikanalys 2016a). Resvaneundersökningar har även utförts åren 2011–2015, men dessa har ansetts ha för låg svarsfrekvens för att kunna användas i Sampers; som argument anges att många korta resor saknades i svaren. I de nya skattningarna av Sampers efterfrågemodeller används resvaneundersökningen från 2005/2006 även till de regionala modellerna (Trafikanalys 2016a).

Sams-data

Socioekonomiska data presenteras på Sams-områdesnivå (delkommunnivå), vilka motsvaras av Sampers prognosområden (Trafikverket, 2016b). Den information som behövs på Sams-områdesnivå sammanställs, för varje scenario som testas, i en Sams-databas som modellen arbetar emot. Sams-databasen innehåller all indata som beskriver markanvändning i form av områdesindelningar, befolkning, arbetsplatser, bilinnehav etc. och är därmed viktig för flera steg i Sampers modellsystem, se Figur 1. Exempel på socioekonomiska data är befolkning efter kön, ålder och sysselsättning; förvärvsarbetande efter näringsgren; sammanräknade inkomster; bostadsyta; fastighetsareal och bebyggd areal samt befolkningstyngdpunkt för Sams-områden. Dessa data kommer från SCB (Trafikverket, 2016a).

Bilnehav inkluderar indata som antal bilar, bilägare, bildisponerare, antal leasingbilar och antal körkort per Sams-område. För basåret 2014 har information tagits fram med hjälp av SCB och kompletterats med information om fordonsflottan och körkortsfördelning. Bilnehav för prognosåret 2040 skulle från början beräknas med hjälp av Bilnehavsmodellen, men resultaten av körningarna blev orimliga. En anledning till detta är att modellen bygger på empiriska underlag från 70- och 80-talen. Bilnehavet är istället beräknat manuellt med befolkningsförändringen per Sams-område. Metoden innebär att antalet bilar och körkort per person hålls konstant över tiden (Trafikverket 2016b).

Sams-databasen *Attraktion* innehåller attraktionsvariabler som t.ex. sjukhus, universitet, köpcentra och turistattraktioner som kan öka sannolikheten att resor görs. Dessa har inte uppdaterats sedan 2001 (Trafikverket 2016a).

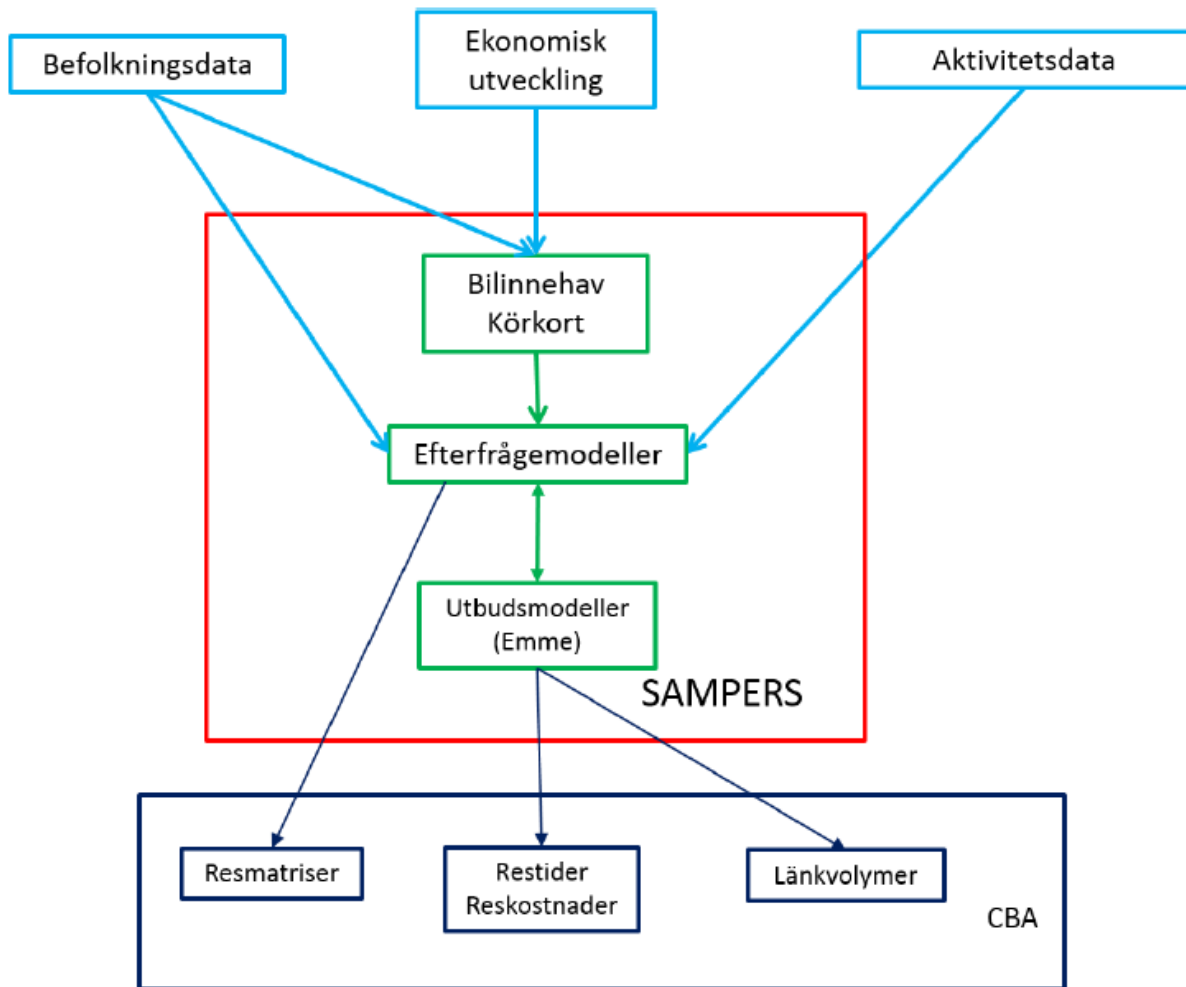
Indata till nätutläggningsmodell

Exempel på indata till nätutläggningsmodellen (Emme) är biljettpriser, väg- och linjenät samt trafikering. Järnvägsnätets nuläge, alltså 2014, är inkodat så nära verkligheten som möjligt och infrastrukturen för 2040 och 2060 motsvarar det som är planerat i nationella och regionala infrastrukturplaner för 2014–2025. Järnvägstrafikeringen antas utökas till 2040, men utbudet av regional busstrafik, som automatkodas från Samtrafikens databas och är 2012 års trafikering, antas vara detsamma för basprognoserna 2014, 2040 och 2060 (Trafikverket 2016a). Undantaget är Stockholms län där det finns beslut om utbyggnad av den regionala kollektivtrafiken (Trafikverket 2016e).

Metoden för att ta fram långväga tågtaxor, till den nationella modellen, är avståndsbaserad och bygger delvis på antaganden baserade på hur man reste 2001 utifrån de taxor som gällde då (M4traffic 2016). Taxor för kollektivtrafik i de regionala modellerna har räknats upp med KPI från 2010 till 2014 och antas vara reellt oförändrade till 2040 och 2060. Taxorna utgörs av taxematriser med priser för respektive färdmedel och ärende i reserelationer. För regionala resor finns endast ett färdmedel vilket är både buss och tåg och en gemensam taxa för dessa. Detta beskrivs i Trafikverket (2012) tillsammans med redogörelse för vilka antaganden som ligger till grund för att taxorna antas vara reellt oförändrade till 2030 och 2050 (men ligger även till grund för åren 2040 och 2060). Hur de ursprungliga taxematriser tagits fram till de regionala modellerna i Sampers har vi inte haft möjlighet att utläsa.

Trafikverket har tagit fram en speciell tidtabellsmodell (i Excel) för att ta fram indata i form av prognostidtabeller för järnväg till Sampers basprognoser. Den hanterar samband mellan trafikmängd, kapacitetsberäkning och tidtabellstider. Till Samgods tar den fram indata i form av kapacitetsutnyttjande (Trafikverket 2016b).²

² Beskrivs mer ingående i Trafikverket (2016j).



Figur 1. Översiktsbild av omvärldsindata till Sampers (Algers 2014, s.4).

Utdata, järnväg

I Trafikverkets rapport Prognos för persontrafiken 2040 (2016e) listas några exempel på resultat från basprognoserna som är relaterade till järnväg: trafik tillväxt för persontrafikanalyser av järnvägstrafiken, linjelänk-sammanställning för samtliga järnvägslinjer, linjetabell för samtliga kollektivtrafiklinjer samt av- och påstigande på järnvägsstationer. Informationen är ganska kortfattad.

Trafik tillväxt för persontrafikanalyser av järnvägstrafiken baseras på fyra grupper av indata/prognosförutsättningar: omvärld, infrastruktur, reskostnader och utbud av kollektivtrafik. Beroende på hur dessa varierar erhålls olika mått på tillväxttal avseende förändring i resandevolymer (transportarbete) mellan 2014, 2040 och 2060 (Trafikverket 2016g).

I linjelänk-sammanställningen för samtliga järnvägslinjer redovisas antalet resor per länk³ för samtliga järnvägslinjer som ingår i trafikutbudet. Sammanställningarna för 2014 respektive 2040 finns redovisade på Trafikverkets hemsida i form av Exceltabeller. I dessa anges resor per år i tusental per länk (dvs. mellan noder på linjer). Även länkavståndet framgår. Tabellerna är dock okommenterade (Trafikverket 2016h).

Linjetabellen visar transportarbete, medelreslängd, emissioner etc. per linje för samtliga kollektivtrafiklinjer, vilket enligt Trafikverket (2016e) ger underlag för beräkningar av

³ En länk avser sträckan mellan två trafikplatser/noder för på- och avstigande.

producentöverskott och externa effekter i Samkalk. I Exceltabellerna (Trafikverket 2016i) går det att utläsa att informationen täcker både lokal, regional och nationell kollektivtrafik med alla färdmedel.

Vad gäller av- och påstigande på järnvägsstationer finns det estimat, men Trafikverket betonar att de bör beaktas med försiktighet då de baseras på prognosresultat (Trafikverket 2016e).

Kalibrering och validering

Valideringsarbetet innebär en kvalitetssäkring av prognoserna, vilket sker genom en jämförelse mellan modellresultaten och observerade resevägsträngder. Vid valideringen kan det framkomma att skillnaderna mellan modellresultat och resestatistik är så stora att det inte kan förklaras av statistisk osäkerhet i undersökningarna. Detta kan bland annat bero på att man antar samma värderingar i restid i modellen för alla individer, till exempel oberoende var i landet man bor. För att kompensera för detta införs påslag på resekostnader i efterfrågemodellen som påverkar attraktionsvärdet för olika kombinationer av ärenden, färdmedel och reseavstånd. Dessa påslag kallas kalibreringsparametrar och kalibrering innebär att man justerar dessa så att modellresultaten stämmer bättre överens med antalet resor i de olika kombinationerna enligt statistiken från resevägundersökningarna. De regionala modellerna är skattade på resevägundersökningarna från 1994–2000. Till Sampers 3.3 - som är den modellversion som använts för att ta fram Trafikverkets Basprognoser med publiceringsdatum 1 april 2016 – togs nya kalibreringsparametrar fram för de regionala modellerna som justerats mot RVU 2005/2006 (Nilsson et al. 2015). Den ursprungliga planen, enligt Persson et al. (2015), var att använda resevägundersökningarna från 2011–2013 för att kalibrera de regionala modellerna, men då det inte gick att säkerställa kvaliteten från dessa användes istället RVU från 2005/2006.

För att kunna validera modellresultaten behövs valideringsunderlag, dvs. resestatistik. Valideringsarbetet beskrivs i Trafikverkets PM (2016f) där det framkommer att nulägesprognosen för 2014 på nationell nivå validerats mot tillgänglig resestatistik (från Trafikanalys). I den tabell där jämförelsen mellan modellresultat och statistik redovisas, finns statistikuppgifter främst för totalt antal resor med olika färdmedel; de är inte uppdelade mellan långväga och regionala resor. För järnväg presenteras statistik enbart för ”regional övrig spår” och totalt antal resor med tåg (se s.10, Trafikverket 2016f).

Ett mer detaljerat valideringsarbete för de regionala modellerna har gjorts på Trafikverkets regioner. En del av syftet med regionernas valideringsarbete är att avgöra huruvida de framtagna kalibreringsparametrarna fungerar bättre än de gamla på en mer disaggregerad nivå. Valideringsarbetet resulterade i att den nya kalibreringsuppsättningen valdes för två regionala modeller (Palt och Sydost), medan den gamla uppsättningen valdes för de övriga tre (Samm, Väst och Skåne). Ett annat syfte med den regionala valideringen, enligt Trafikverket (2016f), är att dokumentera identifierade avvikelser mellan modellresultat och verkligt resande som inte går att åtgärda eller förklara som fel i indata. Då kan dokumentationen användas till att avgöra framtida modellutvecklingsinsatser samt ge kunskap till användare om möjliga modellbrister.

I regionernas valideringsdokumentation framkommer det olika tydligt vilket valideringsunderlag de har haft för resande på järnväg.⁴ I den PM som beskriver vilken validering som gjorts för kollektivtrafikprognosen för region MITT (**Palt Syd**) (Trupina 2015) framgår det tydligt att det för Region Gävleborg har funnits resestatistik per vardagsmedeldygn för två regionala tåglinjer (Gävle–Sundsvall samt Gävle–Ljusdal) samt för alla busslinjer, men till exempel inte för de linjer som SJ AB trafikerar. För bägge tåglinjerna visade jämförelsen att den nya kalibreringsuppsättningen gav bättre modellresultat än den gamla uppsättningen.

⁴ Alla valideringsunderlagsrapporter vi hänvisar till går att finna på <https://www.trafikverket.se/for-dig-i-branschen/Planera-och-utreda/Planerings-och-analysmetoder/Samhallsekonomisk-analys-och-trafikanalys/Kortom-trafikprognoser/> (2017-11-10).

I valideringen av modellresultaten i region **Palt Nord** i Sampers Basprognos framkommer det inte om någon av de kollektivtrafiklinjer som valideras gäller tåg eller buss, eller vilken statistik som använts (Trafikverket 2015e).

I Stockholms län (region **Samm Sthlm**) (Canella 2015) finns jämförande statistik för transportarbete från SL för all spårtrafik i länet. För påstigande i spårtrafiken finns snitt för vardagsmedeldygn samt morgonrusningens maxtimme. Spårtrafiken delas då upp i de olika lokala banorna (totalt), tunnelbanans linjer (uppdelat i färg) samt alla pendeltåg. Det finns även jämförande statistik för påstigande vid större bytespunkter (dygn respektive morgonrusningens maxtimme).

I valideringsarbetet av region **Samm (exklusive Stockholm)** drogs slutsatsen att det var små skillnader mellan den gamla och nya uppsättningen kalibreringsparametrar, men att det saknas valideringsdata för att avgöra vilken som bäst speglar verkligheten. Större skillnader mellan parameteruppsättningarna uppstod för Sydost (Östergötland), men även där saknades valideringsdata för att kunna avgöra vilka modellresultat som är att föredra (M4Traffic 2015).

För region **Sydost (exklusive Östergötland)** finns statistik på sex tåglinjer som inhämtats från de regionala kollektivtrafikmyndigheterna (Validering Sampers – Sydost, version 2016-04). Det är oklart huruvida författarnas beskrivning gäller buss- eller tåglinjer, men de påpekar att de endast har erhållit information om antal påstigande per linje, och inte om var på linjerna påstigningarna skett. De menar vidare att det ofta råder osäkerhet om vilken kvalitet man kan förvänta sig av statistiken. Bägge parameteruppsättningarna överskattar kollektivtrafikresandet i modellen, men de kan inte avgöra om det är statistiken eller modellresultaten som är problemet.

I rapporten om validering i Region Syd (**Skåne**) framkommer att då Skånetrafiken tillhandahåller statistik för på- och avstigande per hållplats/station och per linje, så skulle en mycket mer detaljerad validering kunna ha genomförts (Mcdaniel u.å.) om författarna hade haft mer resurser. De valde istället ut 13 viktiga punkter vid vilka de jämförde passagerarflöden måndag till fredag med Sampers. Dock utan de mer långväga tågens statistik (t.ex. SJ AB).

I region **Väst** jämförs Västrafiks stämplingsstatistik uppdelat på tåg och spårvagn med Sampers på total nivå, samt mer disaggreerat genom påstigande per tåglinje och spårvagnslinje (Sweco och M4Traffic, 2015). Det är emellertid oklart om de räknar med resor per dygn eller vilken tidsrymd de använder. Då resvaneundersökningen 2014 genomfördes med ett utökat underlag för Göteborgs-regionen kunde även dessa resultat användas som valideringsunderlag. Dock slogs all kollektivtrafik ihop till ett färdmedel, vilket innebar att järnvägsresor inte går att utläsa i valideringsrapporten.

2.2. Samgods

Samgods modellerar godstrafik på nationell nivå genom att för en given efterfrågan på godstransporter – i ton per år och varugrupp mellan avsändare och mottagare – söka efter en så kostnadseffektiv transportlösning som möjligt genom att optimera sändningsfrekvenser och sändningsstorlekar samt användningen av olika transportalternativ. Utifrån ett antal fördefinierade transportkedjealternativ genererar modellen först en valmängd av lösningar som kan bestå av både direkttransporter, med exempelvis lastbil, och kombinationer av fordon i transportkedjor. Sedan söker modellen efter den kombination av lösningar som minimerar hela transportsystemets logistikkostnad. Modellen beaktar möjligheten till samlastning inom de varugrupper som används (Trafikverket 2015a). I Figur 2 visas en översikt av hur Samgods är uppbyggt. Den efterfrågan på godstransporter som har beskrivits ovan representeras i figuren av det som benämns basmatriser och det är logistikmodulen som ska lösa det transportproblem som basmatriserna representerar. För att lösa problemet behövs indata i form av nätverksinformation (väglängder, hastigheter etc.), logistiska data (varuvärden, lagerhållningskostnader och liknande), kostnadsdata (exempelvis kostnader för att framföra olika fordon, både per kilometer och per timme) samt regler, som exempelvis att lastbilar över 40 ton totalvikt endast får köra

i Sverige och Finland (Trafikverket 2016a). Den senaste prognosuppdateringen släpptes i april 2016 med basår 2012 och prognosår 2040.

Basmatriserna

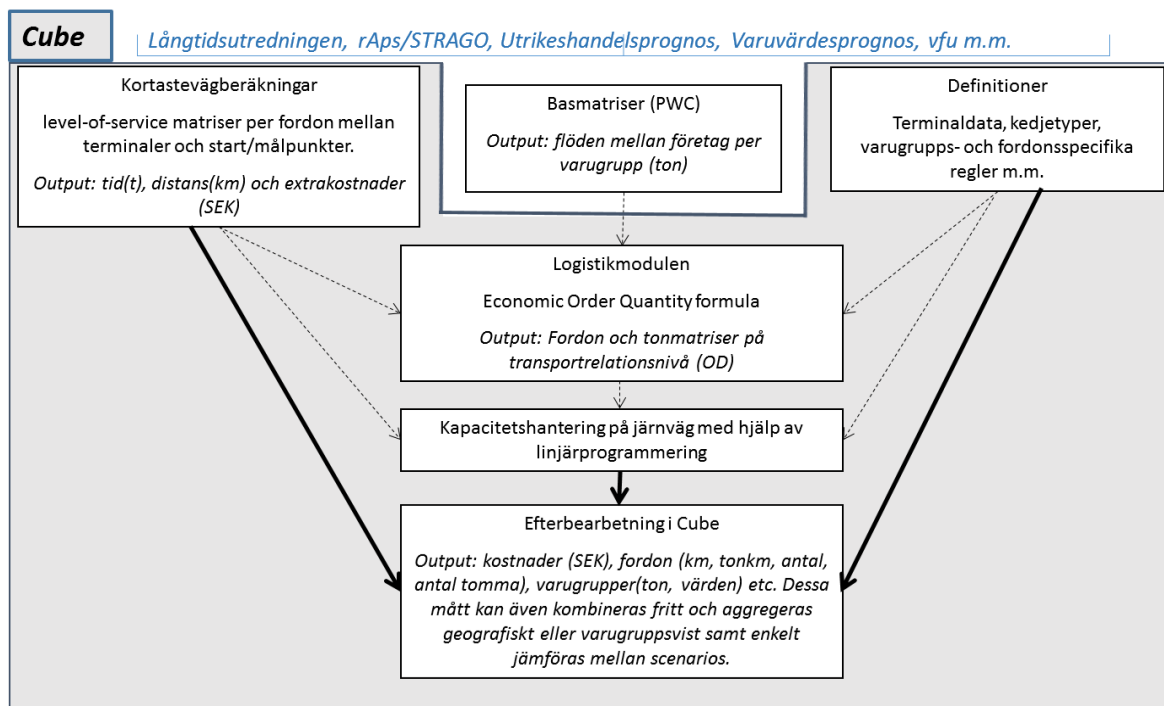
De viktigaste indata till basmatriserna kommer från nationalräkenskaperna, varuflödesundersökningen och utrikeshandelsstatistiken. De matriser som genereras innehåller transportefterfrågan i ton uppdelat på 34 varugrupper och 464 zoner: mellan 290 kommuner i Sverige samt export, import och transit där 174 zoner utgör större regioner i utlandet (Trafikverket 2015a). De varuflödesundersökningar som använts är från 2001 och 2004/2005, men senare undersökningar har gjorts (Trafikanalys ansvar). Utöver varuflödesundersökningen används datakällor som avser de branscher (SNI-koder) som genererar produktion, förbrukning och handel för de olika varugrupperna. Datakällorna kommer främst från Statistiska centralbyrån, vilket också gäller export och import (Anderstig et al. 2015).

Kostnader

Kostnader är olika tids- och avståndsberoende kostnader, kostnader för fordon, skatter och avgifter m.m. När nya trafikprognoser tas fram uppdateras kostnaderna utifrån ny kunskap. Gällande kostnader är presenterade i trafikverket (2016c) tillsammans med avvikelser jämfört med rekommenderade ASEK-värden.

Övriga

Utöver varuflöden och kostnader behövs, precis som i Sampers, indata om noder och länkar i infrastrukturnätverket. De så kallade LOS-matriserna (level-of-service) är aggregerade beskrivningar av nätverket i form av beräknade uppgifter om tid och sträcka mellan terminaler och handelsrelationernas start- och målpunkter. Även logistiska data om terminaler, varugrupper och fordon behövs samt regler som säger vad det är för maxvikter, kapacitetsbegränsningar och tillåten trafikering som gäller.



Figur 2. Översiktsbild över indata till Samgods (Trafikverket, 2016, s.11).

Utdata, järnväg

Samgods kan ta fram transportdata i form av godsflöden och antal tåg för olika bandelar på järnvägsnätet. Utdata på en sådan disaggregerad nivå anses dock vara osäkra. Trafikverket använder därför verktyget Bangods som har utvecklats för att ta fram uppskattningar om antal tåg och godsflöden på bandelnivå fördelat på 12 varugrupper, dels för ett basår, dels för olika prognosår. För att göra denna uppskattning använder sig Bangods av information om trafikering under basåret. Det finns emellertid inte fullständig information om vilken typ av godståg som trafikerar banorna, dvs. om det är vagnslasttåg, kombitåg eller systemtåg. Det saknas också information om typ av gods och godsmängd på varje tåg. En del uppskattningar måste därför göras, exempelvis baserat på antaganden om hur transporterade varor i de olika godstågtyperna fördelas över de tolv varugrupperna (bl.a. baserat på statistik över transportarbete på järnväg för de olika varugrupperna på nationell nivå, se Trafikanalys 2016b). En del information om trafikering, typ av gods och godsmängd inhämtas via direktkontakt med företag (Berglund och Edwards 2016). Vidare påpekar Berglund och Edwards (2016) att nu gällande metodik baseras på information om trafikering enligt tidtabell och att det skulle innebära förbättringar om metoden istället utnyttjar information om den genomförda trafikeringen enligt uppföljningssystemet Lupp. Genom att använda Lupp blir det också möjligt att nyttja information om inrapporterade tågvikter för att ta fram estimat över godsmängd på varje tåg. Olika förslag på hur Bangods kan integreras i Samgodssystemet (eller hur kopplingen mellan dessa kan öka) finns beskrivna i Berglund och Edwards (2017).

Indata om trafikeringen till Bangods finns med andra ord på en disaggregerad nivå i både tid och rum och utdata kan i princip genereras på samma detaljerade nivå. I nuläget består utdata från Bangods av godsflöden under hela året uppdelat på antal fjärr-, lokal-, system-, malm-, och kombitåg, 12 varugrupper och järnvägsnätets olika bandelar (ca 300 st). Varje tågtyp antas vara lika tungt lastat (Aronsson et al. 2015). Osäkerheten i utdata för ett basår orsakas främst av att godsflöden från Samgods för de olika varuslagen fördelas ut på tågen baserat delvis på antaganden och uppskattningar, delvis på information inhämtat via direktkontakt med större industriföretag. Se Berglund och Edwards (2016) för detaljer kring hur fördelningen går till.

Validering av prognoser

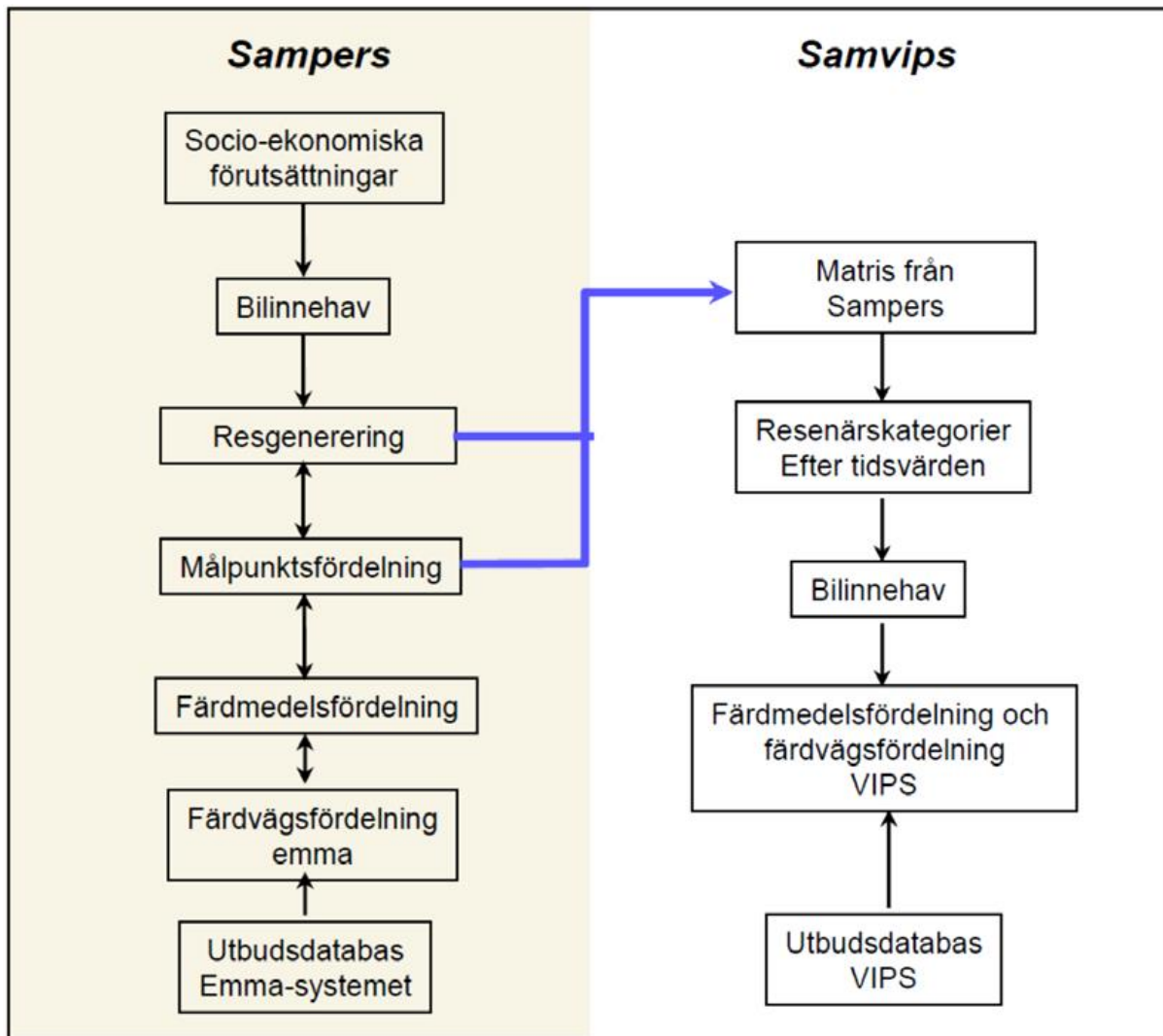
Kval är ett projekt med det övergripande syftet att ”förbättra den svenska nationella godstrafikmodellen Samgods för att tillhandahålla pålitliga prognoser av godsflöden på väg, järnväg och sjö” (Trafikverket 2016d, sid. 5). Projektet drivs av en arbetsgrupp bestående av forskare från KTH, VTI, Sweco och Trafikverket. En förstudie har genomförts och dokumenterats och för tillfället pågår ett fortsättningsprojekt. Förstudien undersökte möjligheter att validera resultat från Samgods med tillgängliga data. Med validering menar författarna hur väl modellen fångar ett verkligt utfall, vilket benämns som den *externa validiteten*.⁵ För att kunna avgöra den externa validiteten bör modellresultaten inte valideras mot data som använts för att estimeras eller kalibrera modellen (eller åtminstone inte på data med samma aggregeringsnivå som använts i modellen). Eftersom varuflödesundersökningen för 2009 inte användes för att skatta nuvarande efterfrågematriser (för 2012) kommer rapporten fram till att resultaten från VFU 2009 skulle kunna användas som valideringsdata gentemot ett Samgodsscenario för år 2009, givet att basmatriser kan tas fram för 2009 (utan att använda data om godstransporter från 2009). Varuflödesundersökningarna erbjuder emellertid endast valideringsmöjligheter på en aggregerad nivå. Vid finare geografisk upplösningen och i synnerhet om resultaten delas upp i varugrupper, blir punktskattningarna behäftade med stor osäkerhet.

⁵ Se Vierth et al. (2016) för en longitudinell uppföljning av basprognoser för både person- och godstransportprognoser som publicerats mellan 1975 och 2009.

2.3. (SAM)VIPS

Vips är en modell för prognoser av persontransporter som utvecklades av VIPS AB. Modellen tillhör numera det tyska bolaget PTV som har integrerat programmet i VISUM, vilket också används för att beräkna trafikflöden. I Vips ingår ingen modell som genererar efterfrågematriser, utan Vips består endast av en nätverksmodell som använder en given efterfrågan på resor och fördelar ut den på olika färdvägar och sträckor. En viktig skillnad mellan Sampers och Vips är att den senare modellen kan skilja mellan operatörer och därmed variera restid och pris för ett visst färdmedel mellan två platser, samt att den kan kombinera olika färdmedel och operatörer för en vald sträcka. Möjligheten att kombinera färdmedel kan vara särskilt viktig för resor över längre avstånd. På så sätt har Vips en mer realistisk modellering av kollektivtrafikens olika utbud och hur detta påverkar fördelningen av resor (Johansen och Lindberg 2016). Därutöver beskriver Algers et al. (2013) hur Sampers och Vips skiljer sig åt när det gäller trafikanternas förhållningssätt till tidtabeller, där Vips lägger en större vikt vid önskemål om tid för avresa och ankomst medan Sampers lägger en större vikt hos individernas preferenser i andra avseenden. Mer specifikt tillåter Sampers en variation i individernas val p.g.a. smak, måttfel, saknade variabler etc., medan Vips antar en uniform fördelning hos individerna vad gäller ideal avgångs- och ankomsttid samt val av rutt. Vips ignorerar därmed den stokastiska delen som rör smak, måttfel, saknade variabler etc. För mer information om skillnader mellan Sampers och Vips, se exempelvis Sika (2009), Algers et al. (2013) och Johansen och Lindberg (2016).

Då Vips inte genererar efterfrågematriser används den ofta i kombination med Sampers och får då benämningen Samvips. Figur 3 är en illustration över hur dessa modellsystem har kombinerats i exempelvis Nelldal et al. (2009) och Wajzman och Nelldal (2012).



Figur 3. Översiktsskild över indata till Samgods (Wajzman och Nelldal, 2012, s.25).

Ett exempel där Samvips används för prognoser av transporter på en nationell nivå är Wajzman och Nelldal (2012) som är en underlagsrapport till den så kallade kapacitetsutredningen⁶. I studien utförs prognoser över transportarbetet i Sverige från år 2010 till 2030. Utöver Samvips används en omarbetad version av Transportrådets godsmodell (Transportrådet 1983). Wajzman och Nelldal nämner ett antal delmodeller som används för att förfinas den prognos som kan tas fram med Transportrådets modell. Exempelvis har de tagit fram en modell för att beräkna effekter av olika banavgifter, en modell för lättkombitransporter och en modell som beaktar förändringar av infrastruktur, trafikering och priser. Prognoserna utgår från en databas över transportflöden inom Sverige samt flöden till och från utlandet, samt information om ekonomi, befolkning, sysselsättning, bilnehav och regional fördelning. Även information om transportnätet ingår, dvs. beskrivning av linjenät, restider och priser.

I Wajzman och Nelldal anges att informationen om godstransporter på järnväg har hämtats från Sveriges Järnvägars redovisningar i Sveriges Officiella Statistik (SOS) och från Green Cargo AB. Information om transportarbetet från övriga trafikslag kommer från flera olika datakällor, exempelvis dåvarande Vägverkets trafikmätningar, körsträckor enligt Bilprovningens redovisningar, SCBs och SIKAs beräkningar av körsträckor etc. De anger att det i de olika datakällorna saknas uppgifter för en

⁶ Benämningen på det uppdrag Trafikverket fick i mars 2011 av regeringen (N2011/1933/TE) att utreda behovet av ökad kapacitet i järnvägssystemet och lämna förslag till effektivitets- och kapacitetshöjande åtgärder.

del år och att de därför har utfört olika typer av beräkningar för att få fram en beskrivning över godstransporterna. Vidare kan nämnas att Långtidsutredningen 2000 användes som underlag för prognosen över den ekonomiska utvecklingen.

2.4. Transtools⁷ och transportmodeller i Storbritannien och Norge

EU-kommissionen har finansierat flera projekt som lett till utvecklandet av modellsystemet Transtools för person- och godstransporter, och som kan appliceras på 42 länder. Modellen täcker Europa men även andra delar av världen för vissa färdstätt. Syftet med modellen är främst att jämföra olika större europeiska projekt, snarare än att användas som en prognosmodell.

Totalt finns det 1 525 zoner i modellen, vilka är något modifierade NUTS 3-zoner. Output i Transtools är därmed transporter mellan dessa zoner. Resorna fördelas mellan tjänste-, privat- och pendlingsresor, medan godstransporter fördelas mellan skåpbil (LGV), lastbilar (HGV) och så kallade gicaliners (GGV). De olika färdstättarna är gång, cykel, bil som förare, bil som passagerare, kollektivtrafik (buss, spårväg/tunnelbana, tåg och flyg). Det finns elva olika varugrupper i modellen.

Transtools består av en passagerarmodell, en godsmodell, och en modell för att fördela trafik på transportnätverket (dvs. nätutläggning). Passagerarmodellen består i sin tur av två delmodeller: en för resor kortare än 100 km och en för resor längre än 100 km. Även godsmodellen består av två delmodeller: en modell för handel och en logistikmodell. Modellen för handel skapar olika scenarion för en framtida exogen tillväxt för produktions- och konsumtionsmatriser (PC-matriser; i Samgods benämns dessa basmatriser PWC-matriser), medan logistikmodellen genererar OD-matriser med godsflöden för olika trafikslag och varuslag. Tillsammans kan dessa modeller användas för att prognosticera flöden. Här kan det noteras att godsmodellen i Transtools är lik Samgods och godsmodellerna i Norge och Danmark på så sätt att den är en "aggregate-disaggregate-aggregate"-modell. Det innebär att genereringen av godsflöden och nätutläggningen sker på en aggregerad nivå (p.g.a. datatillgång), medan logistikbesluten (storlek på transport och val av transportkedja) modelleras på företagsnivå (Ben-Akiva och de Jong 2013).

I ett första steg genererar passagerar- och godsmodellerna transportefterfrågan som är en input i nätutläggningsmodellen. Där skapas variabler för trafikvolym fördelade på olika delar av transportnätverket och mellan olika trafikslag, vilket används som input i passagerar- och godsmodellerna för en nästa iteration. Dessa beräkningar fortsätter tills det uppnås en jämvikt mellan efterfrågan och utbud.

När passagerare ska fördelas ut till spårtrafiken används inte disaggregerad information om tidtabeller för de olika linjerna. Istället används information om hastigheter för olika delar av järnvägsnätet för att beskriva trafiken (vissa restriktioner används för att inte skapa orealistiska trafiknivåer och val av rutter, se sid. 88-90 i Nielsen et al. 2017). Här ingår även information om byten vid stationer. För den del av godsflödena som tilldelas spårtrafik görs en fördelning mellan container- och ej container-trafik, samt huruvida tåget körs på el eller diesel.

Data till Transtools kommer från en rad olika källor, bl.a. annat från TenConnect 2-projektet, ETISplus-projektet, Världsbanken etc.⁸ Därutöver använder godsmodellen data från den svenska varuflödesundersökningen från 2009 och den franska ECHO-undersökningen (Envoi-CHargeurs-Opérateurs). Indata i modellen består bl.a. av information om infrastrukturen (inklusive

⁷ Beskrivningen av Transtools bygger främst på Nielsen et al. (2017).

⁸ Etis är en förkortning för European Transport policy Information System, vars mål är att skapa en länk mellan officiell statistik och tillämpningar inom transportpolicy-området (se Szimba et al. 2012 för mer information). TenConnect 2 var ett projekt med huvudmålet att kalibrera en tidigare version av Transtools, vilket krävde nya indata från medlemsstater (se <https://www.re2.uni-kiel.de/en/research-1/research/ten-connect-2> för mer information).

kapacitetsbegränsningar), transportkostnader (biljettpriser, skatter och avgifter), fordonskostnader (skatter och avgifter), innovationer (fordonstyper, bränsletyp, nya färdmedel inom kollektivtrafiken), restidvärden och diskonteringsräntor. För att beräkna efterfrågan på resor inom passagerarmodellen är data om populationen viktig information, i synnerhet demografiska data med information om ålder, kön, anställning, inkomst, och utbildning. Även information om bilägandet används. Information om observerat antal resor är också en central del i beräkningen av efterfrågan. Denna information kommer från Dateline-undersökningen från år 2001. För nätutläggningsmodellen används data om olika egenskaper hos färdmedlet för olika reseärenden. För bil ingår exempelvis data om avstånd, restid vid fritt flöde och vid trängsel, vägavgifter, bränslekostnader etc. För spårtrafik ingår avstånd mellan hållplatser, tid för på- och avstigning, avstånd och tid ombord spårtrafiken, etc.

I skattningen av logistikmodellen, som är en delmodell till godsmodellen, används data om varuflöden (disaggregerade data från den svenska varuflödesundersökningen 2009 och den franska ECHO-undersökningen), avstånd och transporttid mellan olika zoner, samt information om olika restriktioner (gränsövergångar, begränsningar för olika fordon). Efterfrågedata till godsmodellen består av tillväxttal och socioekonomiska data (inkl. utveckling av arbetsplatser och population). Delmodellen för handel använder en så kallad produktions-konsumtions-matris (från Etisplus-projektet) med information om godsflöden på nationell nivå.

I **Storbritannien** tillhandahåller järnvägsoperatörer information om antal sålda biljetter mellan olika avgångs- och ankomststationer, vilket ger en relativt bra utgångspunkt för att beräkna antal tågpassagerare för olika reserelationer. Detta sker via tågoperatörernas databas Lennon (Latest Earnings Networked Nationally Overnight). Dock saknas det oftast information om den exakta färdvägen. För sålda rese- eller zon-kort är det svårt att beräkna hur många gånger en person har rest och var.⁹ Därför används hushållsundersökningar för att göra uppskattningar av antal resande mellan olika platser. Ett problem med dessa undersökningar är att de är relativt få, vilket gör det svårt att få en fullständig geografisk täckning av resor med olika färdmedel, särskilt tågresor över längre avstånd (Worsley och McNamara 2010, Worsley 2012). Dessa data används i den järnvägsspecifika modellen Network Modelling Framework (NMF) som är en del av den trafikslagsövergripande modellen National Transport Model (NTM) som används av Department for Transport. Totalt finns det information om sålda biljetter för ca 2 650 stationer, vilka aggregeras till ca 570 efterfrågezoner i NMF som ligger till grund för att generera tillväxtfaktorer för antal resenärer mellan dessa zoner. (Department for Transport 2009 och 2015). En ny version (v.5) av NTM är under utveckling.

Vid sidan av den nationella transportmodellen förvaltar järnvägsindustrin i Storbritannien (tillsammans med olika myndigheter) en handbok för prognoser av efterfrågan på tågresor (Passenger Demand Forecasting Handbook, PDFH). Modellsystemet finns översiktligt beskrivet i Worsley (2012). PDHF kan användas för prognoser på en mer disaggregerad nivå jämfört med NMF och NTM. Dock har prognosernas överrensstämmelse med verkligheten dalat för åren efter 2005. Ett arbete för att komma till rätta med detta har initierats, där förstudier pekar på behovet av bättre kunskap om hur olika externa faktorer påverkar efterfrågan på tågresor, dvs. faktorer som inte kontrolleras av olika järnvägsaktörer (se ITS et al. 2016).

En annan delmodell till NTM är **godsmodellen** GBFM (Great Britain Freight Model) som har utvecklats av MDS Transmodal på uppdrag av Department for Transportation (se MDS Transmodal 2008). Modellen har uppdaterats ett antal gånger; version 5 användes i den senaste prognosen som gjordes för det statsägda företaget Network Rail, som ansvarar för underhåll och utveckling av majoriteten av järnvägsinfrastrukturen i Storbritannien (MDS Transmodal 2013). GBFM har 15 varugrupper, och modellerar dessa för väg och järnvägstransporter med en ansats likt

⁹ I Sverige har kollektivtrafikföretagen och de regionala kollektivtrafikmyndigheterna information om antal påstigande. Periodkort gör det svårt att få fram fullständig information om antal på- och avstigande på olika stationer. (SOU 2015)

fyrstegsmodeller. I modellen ingår totalt 2 650 zoner i Storbritannien, samt 350 zoner i utlandet. Indata består bl.a. av trafikmätningar för att få fram antal tunga lastbilar (HGV) på varje länk och i båda riktningar, data från undersökningen CSRGT (Continuing Survey of Roads Goods Transport) som ger information om tonflöden på vägsidan, faktureringsdata från Network Rail (infrastrukturförvaltaren), sjöfartsstatistik, transportkostnadsdata, handelsdata, information om olika industriernas golvutrymmen etc. Modellsystemet består av olika moduler; en basmatris som beskriver flöden av olika varugrupper mellan olika zoner, en kostnadsmodell som beskriver val av trafikslag, en modell för val av fordonstyp, en nätutläggningsmodell för väg och en för järnväg. Modellen saknar alltså en specifik logistikmodellering. Godsflödet fördelas ut på tåg via beräkningar av genomsnittliga nettovikter för tåg lastade med vart och ett av de olika varugrupperna. På så sätt översätts godsflöden till trafikering (MDS Transmodal 2013).

Den **norska godsmodellen** är snarlik, men inte identisk med, den svenska Samgodsmodellen. Den har 39 varugrupper och 554 zoner varav 76 är regioner utomlands. Zonerna inom Norge är på kommunnivå men där de sex största städerna är uppdelade i mellan 5 och 12 zoner. Den norska varuflödesundersökningen för år 2008 har använts i utvecklingen av nuvarande modellversion. Varuflödesundersökningen genomfördes 2009 och är den enda som gjorts i Norge. Dock görs kompletteringar med annan data då varuflödesundersökningen endast täcker inhemska transporter från tillverkningsindustrin och grossisthandeln. Inte heller ger undersökningen information om vilken vara som har transporterats. Istället antas att det undersökta företaget har transporterat den vara som är vanligast att producera inom den näringslivskod företaget tillhör (Vierth et al. 2017). Precis som Samgods modellerar den norska godsmodellen logistikbeslut på en disaggregerad nivå. Enligt de Jong et al. (2016) har detta gett goda resultat då de överensstämmer väl med tillgänglig transportstatistik, även på en detaljerad nivå som terminaler och hamnar.

Även **persontransportmodellen i Norge** är lik sin svenska motsvarighet, med fem regionala modeller för resor kortare än tio mil, en nationell modell för resor över tio mil samt en modell för internationella resor. Precis som i Sverige utgör resvaneundersökningar viktiga indata. När Trafikanalys (2016a) gjorde en jämförelse av synen på och problemen med resvaneundersökningar mellan Sverige och Norge, framkom det att Norge ska skatta om de regionala modellerna mot data från den senaste resvaneundersökningen från 2013/2014, trots att svarsfrekvensen endast var ca 20 procent. Med andra ord anses användningen av gamla undersökningar vara ett större problem i Norge än i Sverige. (Trafikanalys 2016a).

3. Forskning kring metoder och nya indata

För att öka kunskapen om transportsystemets funktion, och effekter av olika åtgärder, genomförs forskning och utveckling kring nya datakällor och metoder. Dels genomförs forskning kring specifikationer av transportmodeller, dels kring nya indata och metoder för att få fram nödvändig information från olika datakällor. Forskningen kring nya indata och metoder kan ofta behandla ett annat trafikslag än järnvägstransporter, men har ändå inkluderats i avsnittet eftersom metoderna får anses vara relativt överförbara till mätning och studier av järnvägstransporter. Utvecklingen av transportmodellerna beskrivs översiktligt i avsnitt 3.1 nedan. I avsnitt 3.2 presenteras forskning kring nya indata.

3.1. Utveckling av transportmodeller

En genomgång av nationella och internationella **godsmodeller** görs av de Jong et al. (2013), som även redovisar utvecklingen kring dessa modeller sedan 2004. Den tydligaste trenden är att modellering av logistiska beteenden hos företag har inkluderats i godsmodellerna, något den holländska modellen Smile var ensam om innan 2005. I utvecklingen mot modellering av logistikbeslut har de svenska, norska och danska godsmodellerna antagit en så kallad ADA-ansats, där ADA står för ”aggregate-disaggregate-aggregate”. Även den europeiska modellen Transtools (se avsnitt 3.3.1) har använt denna ansats, dvs. genereringen av godsflöden och nätutläggningen sker på en aggregerad nivå, medan logistikbesluten modelleras på företagsnivå. Modelleringen av logistikbesluten i Samgods är idag deterministisk. Det som styr företagets val av transportkedja och storlek på sändningen är endast generaliserad transportkostnad och till viss del tillgänglighet till vägar och järnvägar samt varuslagets värde i förhållande till dess vikt (Abate et al. 2014). I realiteten finns det flera andra – ofta samverkande – faktorer som påverkar val av trafikslag och transportkedja, som exempelvis tillförlitlighet och flexibilitet (se exempelvis Lindgren och Vierth 2017); faktorer som till viss del ignoreras i den deterministiska modellen. Det har därför gjorts försök med att utveckla en så kallad stokastisk logistikmodell. I Abate et al. (2016) används information från den svenska varuflödesundersökningen för att estimeras multinomiala logit-modeller där valet av transportlösning förklaras av transportkostnad, transporttid och varuslagets värde i förhållande till dess vikt samt en slumpterm. Genom att fånga upp den inverkan på besluten som ligger utanför de normala förklarande variablerna kan dessa behandlas och tillåtas inverka på modellutfallet. Förhoppningen är att detta på sikt ska ge bättre utfall på en mer disaggregerad nivå av modellresultat. Angreppssättet innebär också att modellen får en bättre teoretisk förankring i producenters, köparens och transportörers beslutsprocesser.

Den stokastiska logistikmodellen som utvecklas till Samgods ligger i linje med den generella utvecklingen på godsmodellsidan, där modelleringen går mot mer disaggregerade ansatser och där det görs avsteg från mer traditionella fyrstegsansatser. De nya metoderna ställer emellertid nya krav på indata, som exempelvis en detaljerad varuflödesundersökning likt den svenska (de Jong et al. 2016). En annan utveckling är att inkludera en dynamisk efterfrågan på transporter; Samgods är en statisk modell i detta avseende, där exempelvis en prisförändring för ett trafikslag inte påverkar efterfrågan på den totala transportvolymen utan endast fördelar om transportvolymerna mellan trafikslagen. Därutöver utförs det ett arbete med olika efterarbetsprogram som ska kunna disaggregera data från Samgods med bättre kvalitet. På järnvägssidan sker detta idag via verktyget Bangods, vilket i nuläget kräver en del manuella bearbetningar samt uppskattningar och antaganden om hur olika godsflöden fördelas på olika tågtyper som trafikerat järnvägsnätets olika bandelar. Förslag på hur Bangods kan förbättras och eventuellt integreras i Samgods (eller hur kopplingen mellan dessa kan öka) ges av Berglund och Edwards (2017).

Bristande tillgång på disaggregerade data över godstransporter har länge varit ett hinder för att ta fram modeller eller göra analyser med mer tillförlitliga resultat. I synnerhet saknas det ofta information om hela transportkedjor samt tids- och säsongsmässiga effekter. Positioneringsdata från satellitnavigationsystem är en möjlig källa till att fånga denna typ av information. (NCFRP 2010). Även Radio

Frequency Identification (RFID) kan vara användbart för att ta fram nödvändig information om godsets rörelser. Nya indata till godstransportmodeller beskrivs i avsnitt 3.2.

Utveckling av **persontransportmodellerna** tycks idag gå mot så kallade aktivitetsbaserade modeller. Efterfrågan på aktiviteter skulle därmed ersätta de steg som rör frekvenser av resor och destinationsval i fyrstegsmodellen (se avsnitt 2.1). Istället modelleras individernas vilja att delta i en viss aktivitet på en viss plats. (Trafikanalys 2016a) I dessa modeller ligger fokus på individernas beteenden och deras motivation bakom en viss resa, medan den traditionella modelleringen fokuserar på själva resan. En annan viktig skillnad rör aggregeringsnivån, där den traditionella modelleringen innebär analyser av resor på en relativt aggregerad nivå (antal resor mellan två zoner med ett visst färdstätt under en viss tid på dygnet). Detta innebär att effekter av exempelvis socioekonomiska faktorer hos individer eller hushåll skattas på en aggregerad nivå, medan aktivitets-baserade modeller kan studera hur en individ (eller hushåll) beter sig med avseende på sin egen socioekonomiska egenskap. Utvecklingen mot aktivitetsbaserade efterfrågemodeller har bl.a. drivits av intresset för att analysera beteenden på individnivå efter förändringar i trafikstyrning såsom trängselavgifter och samåkningsinitiativ. Ökade krav på prognosernas precision har också drivit på utvecklingen, där även bättre möjligheter att samla in data och ökad datorkraft och -kapacitet har underlättat. (Pinjari och Bhat 2011) Ett exempel där en aktivitetsbaserad modell används är Sacramento (se SACOG 2015). Trafikverket har en långsiktig vision att Sampers ska bli en aktivitetsbaserad modell. För övriga visioner och utvecklingsplaner kring Sampers, se Trafikverket (2015c och 2017).

Aktivitetsbaserade modeller ställer nya krav på indata. Exempelvis behövs detaljerade data över ruttval och vilka aktiviteter individerna reser till. Forskning kring nya indata beskrivs i nedanstående avsnitt.

3.2. Nya indata

I Sverige och många andra länder är resvaneundersökningar den traditionella metoden för att ta fram transportstatistik, dels för att ge en nulägesbeskrivning av transporter och dels för att beteendebeskrivningen gör det möjligt att genomföra prognoser. Det finns emellertid en fallande svarsfrekvens hos dessa undersökningar och en risk för ett snedvridet urval. Dessutom leder modellutvecklingen till ett behov av allt mer disaggregerade data. Tillsammans har detta skapat ett behov av forskning kring nya datakällor och metoder för att hantera gällande datatillgång inom transportmodellering. Samtidigt har det skett en teknikutveckling som har inneburit nya möjligheter för datainsamling inom transportområdet, framförallt genom låga priser och bättre funktion hos de tekniska enheter som behövs för att registrera och kommunicera information om positionering.

Satellitnavigationssystem som det amerikanska Global Positioning System (GPS) kan exempelvis användas för att följa både fordon, gods och människor som förflyttar sig, med en ökad detaljnivå i både tid och rum jämfört med traditionella undersökningar. Denna datainsamling innebär möjligheter att vidareutveckla transportmodelleringen. Förutom satellitnavigationssystem och den traditionella enkätbaserade metoden, finns det andra datakällor som kan användas för att ta fram transportdata; exempel är automatisk registrering via Radio Frequency Identification (RFID) eller andra typer av sensorer (tågets vikt, antal på- och avstigande), mobilnätdata, och kommersiell aktivitet (även sociala medier har lyfts fram som en potentiell datakälla med information om positionering och kopplingar till socio-ekonomiska faktorer; se exempelvis Rashidi et al. 2017). Dessa indata skapas ofta genom en passiv datainsamling och består av en stor mängd data, något som har krävt en utveckling av metoder för att få fram den information som behövs i modellering av transporter. Därtill pågår det forskning där olika datakällor kombineras för att få fram relevant transportdata. Nedan görs en kortare genomgång av forskning kring de nya datakällor och metoder som kan användas (och har använts) i transportmodellering. Vi hänvisar till Clark et al. (2017) för en detaljerad kartläggning av verktyg och applikationer för resvaneundersökningsdata, samt hur pass väl olika datainsamlingsmetoder kan svara på de frågor som är viktiga inom resvaneundersökningar.

3.2.1. Satellitnavigationssystem

Behovet av nya datakällor och metoder i transportmodelleringen har gjort att forskare under en tid har intresserat sig för data inhämtat via GPS, ett arbete som har pågått sedan mitten av 1990-talet. Datamängderna som behöver analyseras kan vara enorma, men i takt med utvecklingen av kraftfullare datorer och ökad lagringskapacitet har forskningen kring GPS-data inom transportområdet ökat (NCHRP 2014a). Till en början användes GPS-tekniken för att passivt inhämta information utan någon input från användarna, främst för att korrigera antal resor som självrapporterats. Ett antal studier lät användare aktivt ge information via en GPS-enhet för att få fram mer information om resorna, och lät dem även fylla i en traditionell resvaneundersökning för att kunna jämföra de olika metoderna. (NCHRP 2014b) Dessa typer av jämförelser har sedan fortsatt genom åren, men har främst använts för att korrigera antal resor som rapporteras (se exempelvis Cambridge Systematic 2007 och Nustats 2010) eller utföra djupare analyser kring underrapportering av resor; se exempelvis Bricka och Bhat (2006) som bl.a. analyserade hur olika demografiska faktorer påverkade graden av underrapportering.

Ett mål har länge varit att GPS-data helt ska kunna ersätta traditionella resvaneundersökningar. Exempelvis undersökte Wolf detta i sin avhandling (2000), och visade att det fanns en potential att härleda information om exempelvis syftet med resan. Stopher and Shen (2011) visade dock på blandade resultat i en jämförelse mellan de olika metoderna för datainsamling, vilken bl.a. berodde på att GPS-enheten inaktivt togs med under en del av resorna, ett problem som möjligen kan hanteras med hjälp av exempelvis smarta telefoner som användarna oftast har med sig när de lämnar hemmet.

En jämförelse mellan data från traditionella resvaneundersökningar och data inhämtad med hjälp av smarta telefoner, har nyligen gjorts i Sverige av Allström et al. (2016). De använde en applikation för smarta telefoner med GPS och jämför detta med en resvaneundersökning som var web-baserad. Syftet med studien var att undersöka om en applikation till en smarttelefon kan ersätta eller vara ett komplement till den traditionella undersökningen. De kriterier som användes för jämförelsen mellan insamlingsmetoderna var att mäta dess precision i beskrivningen av resor, vilket inkluderade antal glömda resor, beräkning av avstånd och restid, geografisk kodning av start- och slutpunkter för resor, samt hur stor bördan var för respondenter och hur mycket administration som krävdes. Målet med applikationen var inte att lösa problemet med en låg svarsfrekvens vid traditionella undersökningar, och precis som för traditionella resvaneundersökningar var det svårt att få personer att delta. En inbjudan att delta gick ut till 130 000 slumpmässigt utvalda, men endast 495 registrerade sig till fältförsöket. Under fältförsöket deltog 171 användare som kompletterade och korregerade resor, vilket gav totalt ca 1 miljon GPS-punkter från 2 142 resor. Resultaten visade bl.a. att den automatiska klassificeringen ledde till att en resekedja kunde detekteras med 79 procents noggrannhet. Vilket färdmedel som användes fick rätt klassificering i 54 procent av fallen. En slutsats i rapporten är att i nuläget kan en applikation likt den som utvecklades inom projektet endast vara ett komplement till traditionella resvaneundersökningar, men att potentialen skulle kunna höjas med hjälp av en utveckling av metoderna för att automatiskt detektera färdmedel, målpunkt och ärende.

En liknande studie har utförts av Safi et al. (2017) som jämför olika GPS-lösningar med en web-baserad resvaneundersökning. Dels användes en handburen GPS-enhet, dels en smart telefon med GPS, och dels en applikation (Atlas II) som utvecklats till smarta telefoner. Svarsfrekvensen och kvaliteten på data var bäst för applikationen, som dessutom fångade resebeteendet med en relativt hög upplösning (10 m). Individer med applikationen fick själva ange färd sätt och syfte med resan.

VTI, Trafikanalys och Trivektor arbetar för närvarande med en studie om framtidens resvaneundersökningar. En delstudie som pågår (november 2017) ska jämföra insamlingsmetoder för reseinformation från deltagare i Västra Götalandsregionen. Tre metoder testas: en webbenkät, en mobilapplikation och en traditionell resvaneundersökning som genomförs av Trafikkontoret i Göteborg. Mobilapplikationen samlar in information om hur deltagaren rör på sig och deltagaren granskar sedan resultaten och justerar vid behov (VTI 2017).

GPS-information har även använts i forskning eller utredning kring godstransporter. Enligt Giuliano et al. (2015) är det vanligt att fordonsparksförvaltare utrustar sina fordon (även containrar och lastpallar) med GPS eller aktiva RFID transponders för att effektivisera sin produktion. McCormack et al. (2010) samlade in GPS-data från lastbilar för att bl.a. kunna följa godsets transportmönster över vägnätet i delstaten Washington i USA. Med hjälp av en algoritm skapades OD-matriser. Dessutom analyserades lastbilarnas hastigheter, något som kan vara användbart för att kvantifiera effekter av olika vägarbeten och hitta återkommande flaskhalsar. Ett liknande projekt i mycket större skala har skapats av American Transportation Research Institute (ATRI). Insamlingen av GPS-data från ett stort urval av lastbilar har pågått sedan 2002. Fler än 100 miljoner GPS-positioner samlas in varje dag. Se exempelvis Liao (2014) och Johnson et al. (2014) för en beskrivning av dessa data.

För tågtrafiken i Sverige finns det uppgifter om passerade driftsplatser och signaler som hämtas från systemen TPOS och ATL. Vissa tåg är utrustade med GPS (Joborn 2015). Med andra ord finns det god kunskap om tågens position. Det är snarare informationen om godset och passagerarnas positioner som saknas, vilket skulle kunna erhållas – som beskrivits ovan - via exempelvis GPS-data från mobiler eller via GPS-sändare på varor eller lastbärare. En annan datakälla som kan vara användbar för att få fram information om antal passagerare är mobilnätdata, vilket beskrivs i nästa avsnitt.

3.2.2. Mobilnätdata

Information om positionering går även att erhålla via de signaler som sänds ut mellan mobiltelefoner och mobilnätet. En aktiv användning av telefonen ger detaljerad information om position i tid och rum, men även en telefon som inte används aktivt kan ge positioneringsdata, om än på en mindre detaljerad nivå (en beskrivning av tekniken finns i Allström et al. 2015).

Även om mobilnätdata ger bra information om användarnas positioner, behöver datamaterialet bearbetas och analyseras för att några slutsatser ska kunna dras kring rörelser, som sedan kan användas i exempelvis transportmodellering. Vilka platser som användarna har besökt och vilken plats som kan karakteriseras som ett hem eller arbetsplats är en viktig del. Halepovic och Williamson (2005) var en tidig studie inom området som bl.a. noterade att de flesta mobiltelefonanvändarna genomför merparten av sina telefonsamtal från ett fåtal positioner, och antog att en av dessa platser var hemmet när samtalsaktivitetens maxnivå var under kvällarna. En rad olika studier har sedan utvecklat metoder och modeller för att definiera mobilanvändares positioner som hem eller arbetsplats (se exempelvis Ahas et al. 2010, Phithakkitnukoon et al. 2010, Isaacman et al. 2011 och Calabrese et al. 2012).

Med mobilnätdata finns det möjlighet att följa individer under en längre period, vilket ger goda förutsättningar att prediktera deras framtida rörelser, särskilt då studier har visat att människors rörelser har en hög grad av regularitet i både tid och rum (se González et al. 2008 och Song et al. 2010a). Song et al. (2010b) analyserade mobilnätdata från en tre månader lång period där de kunde följa rörelserna hos 50 000 individer. Studien visade att det finns en teoretisk potential att prediktera deras framtida rörelser med upp till 93 procent noggrannhet, samtidigt som inga individer hade en lägre förutsägbarhet än 80 procent. De undersökte även om det fanns en variation i förutsägbarheten med avseende på olika demografiska faktorer såsom ålder och kön, men fann inga statistiskt signifikanta skillnader. Några sådana skillnader fanns inte heller med avseende på befolkningstäthet, urbana eller icke-urbana miljöer eller tillhörighet till en viss språkgrupp. En liknande studie har utförts av Lu et al. (2012) där 1,9 miljoner användare följdes under ett års tid i anslutning till jordbävningen på Haiti år 2010. Trots den stora ökningen i individernas rörelser och reseavstånd efter jordbävningen, gick det att förutsäga deras rörelser med upp till 85 procent. Ett annat exempel är Lu et al. (2013) som kommer fram till en förutsägbarhet på upp till 88 procent baserat på mobilnätdata från en halv miljon användare i Elfenbenskusten under en period på 5 månader.

Denna typ av mobilnätdata kan användas för att skapa OD-matriser som innehåller information om antal resor som utförs under en given period mellan olika start- och slutpunkter; en information som utgör en viktig del i modellering av transporter. Exempel på studier är Caseres et al. (2007) som

studerar flödet av mobiltelefoner och skapar OD-matriser baserat på antagandet att en mobil som rör sig längs en sträcka alltid byter basstation vid samma position. Iqbal et al. (2014) använder mobildata tillsammans med en begränsad mängd trafikdata för att skapa OD-matriser med trafikflöden för olika tider på dygnet. Baserat på hur mycket mobilen används, under vilka dagar i veckan och tider på dygnet, skapar Alexander et al. (2015) OD-matriser för resor med olika syften och under olika tider på dygnet. De jämför resultaten med traditionella resvaneundersökningar och andra datakällor, vilka överensstämde väl. En slutsats var att mobilnätsdata och den presenterade metoden kan vara ett bra komplement till övriga datakällor som används för transportmodellering. Ytterligare ett exempel är Frias-Martinez et al. (2012) som skapade OD-matriser för pendlare baserat på mobilnätsdata. De använde så kallade temporala associationsregler för att identifiera pendling från hemmet till jobbet och vice versa. De jämförde sedan resultaten med matriser som skapats via data från traditionella resvaneundersökningar. En slutsats var att deras metod ger en hög precision och kan användas i samma syfte som matriser genererade via traditionella undersökningar.

Från mobilnätsdata finns det även möjlighet ta fram information om vilket färdstätt som används vid olika resor. I Wang et al. (2010) användes information från mobilnätsdata om resor som har en viss start- och slutpunkt, samt en viss restid. Baserat på denna information skapades olika ej överlappande grupper, där restiden för varje medlem i gruppen är så lik de andra medlemmarnas restider som möjligt, och så olik de andra gruppernas restider som möjligt. Grupperna definierades som antingen gång, bil eller kollektivtrafik. Resultaten jämfördes sedan med restider för dessa färdstätt enligt Google maps, och andelen resor med de olika färdstätt jämfördes med andelar enligt traditionella resvaneundersökningar.

I Aguiléra et al. (2014) används mobilnätsdata för att ta fram information om beläggningsgrad, restider, och passagerarflöden till OD-matriser för tunnelbanetrafiken i Paris på en detaljerad nivå. En jämförelse av resultaten med data från resekort, visar att mobilnätsdata ger konsistent information, även om det krävs en del bearbetning för att skapa passagerarflöden med mobilnätsdata. Ett problem som tas upp är att då data kommer från endast en operatör krävs information om dess marknadsandel, något som kan variera med avseende på faktorer som är svåra att kontrollera för. Se litteraturöversikten i Wang et al. (2017) för mer information om användningen av mobilnätsdata i analyser av resor.

3.2.3. Radio Frequency Identification (RFID), vägning, och Automatic Passenger Counting (APC).

Radio Frequency Identification (RFID) är en teknik som kan användas för att automatiskt registrera transporter. Tekniken bygger på att mottagarenheter monteras på fasta positioner. Resenärer och/eller gods har i sin tur taggar som sänder ut ett unikt id-nummer när de passerar en RFID-mottagare. Denna avläsning kan ske vid snabba hastigheter, vilket har gjort att tekniken lämpas sig väl för järnvägstrafik. Under 2017 introducerade Trafikverket en RFID-tjänst för tågoperatörer som möjliggör en identifiering och positionering av enskilda järnvägsfordon. Till en början placerades 180 RFID-läsare ut vid olika delar av järnvägsnätet. Dessa läsare är placerade i anslutning till andra detektorer (varmgång-/tjuvbromsdetektorer, hjulskadedetektorer, strömavtagardetektorer), och är en frivillig tjänst som erbjuds till operatörer för att de ska få information om sina fordon och öka möjligheter till proaktivt underhåll av dessa (se Trafikverket 2015b). Detta visar att det redan idag finns installerad teknik för att inhämta information om enskilda järnvägsfordon som inte enbart bygger på självrapportering. Detta genererar emellertid främst trafikdata som behöver kombineras med annan information för att få fram disaggregerade transportdata.

Vägning av tågen är ett sätt att få fram information om det som transporteras. Ca 190 stationära detektorer finns idag installerade på det statliga järnvägsnätet (varav 180 har RFID-läsare). Förutom fordonsvikten ger dessa detektorer information om exempelvis hastighet, tåglängd, och axelantal (det är hjulskadedetektorerna som ger information om fordonsvikt). Informationen lagras i en detektor-databas (se Trafikverket 2015d); en information som kan användas för att beräkna antal passagerare på

ett tåg. Ett exempel på en sådan analys är Nielsen et al. (2014) som beräknar antal passagerare med hjälp av de viktdektorer som finns på tåget och som vanligen används för att kontrollera inbromsning. Resultaten från vägningar av tågen jämfördes med information från manuell räkning av passagerare och med resultat från sensorer som installerades vid tågdörrarna (benämns ofta Automatic Passenger Counting, se nedan). Analyserna visade att tågvikten gav bättre information jämfört med sensorerna. Frumin (2010) är ett annat exempel på en studie där tågens vikt användes för att ta fram passagerarflöden. Med hjälp av information om antal passagerare på tunnelbanetåg från manuella räkningar, skattades parametrar som kan användas för att dra slutsatser om hur många passagerare en tåg vikt motsvarar (på så sätt kan mätfel hanteras). Slutsatsen är att tågens vikt är användbar och har en stor potential när det gäller att ta fram information om passagerarflöden.

Inom persontransportmarknaden används ofta RFID-tekniken i form av resekort, vilket ger möjligheter att följa passagerarflöden på en disaggregerad nivå. Zhu et al. (2014) är ett exempel på en studie som använder information från resekort för att kalibrera nätutläggningen av tågresor i transportmodeller, medan Sun et al. (2015) använder samma typ av data för att lägga ut passagerarflöden i ett tunnelbanesystem. Ett problem med resekortdata är att det saknas information om ruttval. Zhou och Xu (2012) försöker lösa detta problem med hjälp av en modell som lägger ut passagerarflöden baserat på tider när passagerare har påbörjat och avslutat sin tågresor, vilket ger färre möjliga ruttval för modellen (som väljer den mest troliga av dessa). Modellen verifierades mot data från tunnelbanesystemet i Peking. Det finns många fler exempel på studier av resekort för att ta fram transportdata (se Pelletier et al. 2011 för en litteraturoversikt). Dessa kombinerar ofta resekortdata med annan typ av data, något som behandlas i avsnitt 3.2.4.

En annan typ av automatisk registrering kan ske via så kallade Automatic Passenger Counting (APC), en teknik som exempelvis finns på en del nya tåg i Norge (se nästa avsnitt). I Ji et al. (2015) används data från APC på bussar som trafikerar campusområdet på Ohio State University för att skapa OD-matriser med passagerarflöden. Resultaten jämförs med OD-matriser som skapats via reseundersökningar som utförs på bussar. Resultaten visar att indata från APC kan ge OD-matriser med minst lika bra kvalitet, även när ett ”stort” antal reseundersökningar har utförts. Liknande resultat redovisas i Mishalani et al. (2011), som även visar hur kombinationen av APC och reseundersökningar påverkar estimatens kvalitet.

Registrering via videoövervakning är ytterligare en metod för att ta fram statistik över antal passagerare. Chen et al. (2008) använder information från kameror i bussar som fångar flödet av passagerare som stiger av och på vid olika hållplatser. Med hjälp av experiment visar de att metoden kan ge en noggrannhet på 92 procent.

3.2.4. Kombination av datakällor

De ovannämnda datakällorna har sina för- och nackdelar; en och samma datakälla kan inte täcka det behov av information som behövs. En kombination av de olika datakällorna kan därför vara en lösning. Exempelvis kan GPS-data och mobilnätdata vara bra komplement då den förra har en hög detaljnivå men större vita fläckar jämfört med den senare datakällan (Allström et al. 2015).

I Holleczek et al. (2014) kombineras mobilnätdata med data från kollektivtrafiken för att ta fram information om antal resande och val av färdmedel i Singapore. Positioneringsdata från 3,4 miljoner mobilanvändare samlades in från den största nätoperatören i Singapore och användes tillsammans med data från kollektivtrafiken bestående av 4,4 miljoner resor som registrerats via digitala resekort. Dessa data innehåller information om på- och avstigningar på bussar och tåg, samt var och när detta skedde. Totalt antal resor (kollektivtrafik och privata resor med bil eller taxi) beräknades genom att skala upp antal resor som generats av mobilnätdata med hänsyn till operatörens marknadsandel och invånarnas mobilanvändning. Antal privata resor (bil eller taxi) erhöles genom att subtrahera kollektivtrafikresorna som registrerats via digitala resekort. Resultaten överensstämde förvånansvärt bra med antal resor och färdmedelsval enligt genomförda resvaneundersökningar.

Bull-Berg et al. (2017) kombinerar mobilnätdata och punktlighetsdata för att få fram antal tågpassagerare. Mobilnätdata erhöles från en operatör (Telenor) och bestod av antal uppkopplingar till fem basstationer som fanns utmed en järnvägssträcka. Alla uppkopplingar till dessa basstationer skedde dock inte enbart via mobilanvändare som reste med tåg, även om studien valde basstationer som ligger utanför tätbebyggt område och en bit från större vägar. Punktlighetsdata användes sedan för att beräkna när ett tåg har passerat basstationerna. Studien hade även tillgång till passagerardata från automatisk räkning av antal på- och avstigande passagerare (APC) som installerats på ett antal nya tåg i Norge. På så sätt kunde resultaten från mobilnätdata jämföras med passagerarmätningar enligt APC och ge ett svar på huruvida det går att skala upp antal resenärer med en viss faktor (alla resenärer har inte Telenor som operatör). Förhållandet mellan resultaten från de olika datakällorna varierar över olika tider på dygnet, men uppvisar en liknande utveckling när flera dagar jämförs med varandra. En slutsats är att det behövs information under en längre tidsperiod för att kunna ta fram faktorer som kan användas för att skala upp antal resande enligt mobilnätdata, och att dessa faktorer kommer att variera för olika tider på dygnet och för olika sträckor.

Zhao et al. (2007) använder flera olika datakällor där insamling har skett automatiskt för att skapa OD-matriser för buss- och tågresor. Källorna som används är; Data från ett system för lokalisering av fordon (Automatic Vehicle Location, AVL), data från passagerarräkning (Automatic Passenger Counting APC) och ett biljettsystem (Automatic Fare Collection, AFC). Endast information om antal påstigande var tillgängligt från biljettsystemet. En brist är dock att resor där endast tåg användes inte kunde ingå i OD-matriserna. Den använda metoden lyckades endast skapa OD-matriser för 71 procent av de resor som gjordes med resekort. Då andelen resor med dessa kort var 88 procent fångade metoden ca 64 procent av alla resor.

Reddy et al. (2010) använder information från smarta telefoner som är utrustade med en GPS och accelerometer för att identifiera huruvida en individ går, springer, cyklar eller färdas i ett motorfordon. Att kunna skilja mellan olika motorfordon är centralt för transportmodellering, något som Stenneth et al. (2011) gjorde med hjälp av GPS-data och information om transportnätverket (tågstationer och busshållplatser) samt realtidsinformation om bussarnas position. Med 93,5 procents noggrannhet lyckades studien identifiera om individer färdades med bil, buss, tåg, cykel eller om de gick.

Munizaga och Palma (2012) kombinerar data från resekort och GPS, tillsammans med geokodad information om kollektivtrafiknätet, för att ta fram OD-matriser för resor med buss och tunnelbana i Santiago, Chile. Precis som i många andra liknande studier ingår inte alla resor i datamaterialet. Studien skalar därför upp antalet resenärer genom att skapa olika uppräkningsfaktorer beroende på vilken typ av resa som saknas (om slutpunkt saknas i biljettsystemet, om både start- och slutpunkt saknas i biljettsystemet, resor som inte ingår i biljettsystemet). En liknande studie har utförts av Farzin (2008) som använder data från resekort tillsammans med GPS-information för att skapa en OD-matris i Sao Paolo. Resultaten jämfördes med information inhämtat via traditionella resvaneundersökningar. En slutsats är att metoden ger ett större datamaterial som kan användas för att skapa mer omfattande och detaljerade OD-matriser som möjliggör analyser på en mer disaggregerad nivå.

4. Diskussion och slutsatser

Transportmodeller är en förenklad representation av verkligheten och innehåller inte alla detaljer som påverkar transportefterfrågan och val av färdväg. Graden av förenkling skiljer sig mellan olika modeller, bl.a. beroende på vilka frågor som de är ämnade att ge svar på. Det sker emellertid en utveckling mot alltmer detaljerade modeller, som ska vara bättre på att fånga transportbesluten hos företag och personer, något som kan ge mer komplicerade modeller med större krav på indata.¹⁰ Mer specifikt tycks modelleringen av persontransporter gå mot aktivitetsbaserade ansatser, där resebeteenden modelleras på en disaggregerad nivå, både vad gäller syfte, tid och position. Denna trend finns även inom modellering av godsflöden; den heterogena godsmarknaden och svårigheten att förutsäga godsflöden har gjort att utvecklingen även här har gått mot aktivitetsbaserade modeller eller logistikmodeller för beslut på företagsnivå. Dessa modeller behöver disaggregerade data med information om antal passagerare och flöden av olika varuslag på en finare detaljnivå i både tid och rum. Nya typer av indata och nya insamlingsmetoder har därför blivit aktuella i takt med teknikutvecklingen, något som även har drivits på av en allt sämre svarsfrekvens för de traditionella resvaneundersökningarna. De nya typerna av indata kan dock i nuläget främst fungera som komplement till övriga data som används i transportmodellering, och de är sällan helt problemfria.

För de flesta studier som använde mobilnätdata var det exempelvis ett problem att datamaterialet i de flesta fall endast kom från en av flera operatörer. Därtill finns problemet att inte alla resande har en mobiltelefon eller att mobiltelefonen behöver vara aktiv för att positionen ska registreras. Detta kräver någon form av urvalsanalys (vilket inte alltid förekommer i studierna) samt en lämplig faktor som skalar upp antalet resenärer - en faktor som kan variera för olika tidpunkter och linjer. Ett generellt problem med en del av den automatiserade datainsamlingen är med andra ord att de endast fångar en del av resorna. Även om automatiserad datainsamling ger betydligt mer information om positioner och tidpunkter jämfört med traditionella undersökningar, så har även dessa problem med att urvalen kan vara snedvridna.

För att kunna genomföra prognoser av persontransporter behövs även information om exempelvis socioekonomiska faktorer, något som GPS- eller mobilnätdata inte innehåller. Det främsta användningsområdet hos dessa nya indata har därför bedömts vara som komplement till övriga indata, samt för att kalibrera och validera modellresultaten (Trafikanalys 2016a). En slutsats i rapporten från Trafikanalys (2016a) är att resvaneundersökningar även i framtiden kommer att vara en viktig datakälla till persontransportmodeller, och att de därför behöver vidareutvecklas för att tackla de problem som finns idag, som exempelvis låg svarsfrekvens. Ordentliga bortfallsanalyser är ett sätt att få information om vilka konsekvenser den låga svarsfrekvensen innebär. Denna problematik finns även med undersökningar gjorda via exempelvis mobilapplikationer.

Automatisk insamling av resekortdata har använts i olika studier för att ta fram passagerarflöden för ett transportnätverk. Den informationen är sällan heltäckande då den inte ger information om ruttval och antal avstigande. Vidare saknas ofta detaljerad information från resor som gjorts med period- eller zonkort. Studierna från Storbritannien är exempel på att information om antal tågresenärer via biljetter och resekort behöver kompletteras med ytterligare disaggregerad information för att kunna skapa tillförlitliga estimat över flöden av passagerare mellan olika stationer, och framförallt för att kunna skapa tillförlitliga prognoser.

¹⁰ Samtidigt finns det undantag där det har skapats enklare modeller, exempelvis den svenska modellen LuTrans (LandUse Transport model) som är en förenklad version av Sampers för att minska beräkningstider och förenkla hantering av indata, samt nederländska modeller för övergripande analyser (RTK 2009 och Trafikanalys 2014).

Samtidigt är det viktigt att inte glömma potentialen hos den mängd av information som kan samlas in och bearbetas.¹¹ Studier över mobilnätdata som följer individer under en period visar att vi har ganska förutsägbara resmönster, även om dessa kan förändras p.g.a. innovationer som exempelvis nya transportmedel eller kollektivtrafiklösningar. Denna typ av data kan användas i stor skala och det finns dessutom möjligheter att koppla den till socioekonomiska data och andra faktorer som är viktiga för att förklara variationer i resmönster. Vissa studier har exempelvis visat att det är möjligt att bestämma individernas arbetsplats och hem med hjälp av detaljerad information om individernas position i tid och rum. Studier har även visat på goda resultat när det gäller att skapa OD-matriser med hjälp av mobilnätdata. En del studier har också fördelat dessa resor på olika färdsätt, även om precisionen i detta kan variera.

När det gäller godsflöden saknas det ofta information om hela transportkedjor och tids- och säsongsmässiga effekter. Positioneringsdata från satellitnavigationssystem har möjlighet att tillhandahålla denna typ av information. För tågtrafiken kan detektorer för tågens vikt vara användbara för att beskriva godsflöden. Även om godstågens vikt inte ger information om godstyp, så kan den informationen vara användbar i kalibreringen av godsmodellerna och ge en tillförlitligare beskrivning av godsflöden fördelat på olika tåg och delar av järnvägsnätet. I dagsläget sker detta med bl.a. ej fullständig information om godsmängd per tåg. Detektorer för tågens vikt kan även användas för att ta fram information om antal passagerare per tåg (potentiellt mellan alla stationer om detektorn finns på tåget, eller om ett stort antal stationära detektorer placeras ut), något som skulle ge en bra nulägesbeskrivning av tågresor.

Sammanfattningsvis finns det en stor potential i att ta fram och använda mer disaggregerade data över transportflöden med hjälp av nu tillgänglig teknik och datainsamlingsmetoder.

¹¹ Ett tecken på detta är att det pågår en hel del forskning inom området, både i Sverige och övriga världen, vilket till viss del har beskrivits i föreliggande studie. Se även Clark et al. (2017) för en aktuell genomgång av olika applikationer och verktyg som är under utveckling i Sverige och i andra länder.

Referenser

Abate, M., Vierth, I., de Jong, G. (2014). Joint econometric models of freight transport chain shipment size choice. CTS Working Paper 2014:9, Centre for Transport Studies, Stockholm.

Abate, M., Vierth, I., Karlsson, R., de Jong, G., Baak, J. (2016). Estimation and implementation of joint econometric models of freight transport chain and shipment size choice. CTS Working Paper 2016:1, Centre for Transport Studies, Stockholm.

Aguiléra, V., Allio, S., Benezech, V., Combes, F., och Milion, C. (2014). Using cell phone data to measure quality of service and passenger flows of Paris transit system. *Transportation Research Part C*, 43, 198-211.

Ahas, R., Silm, S., Järv, O., Saluveer, E., och Tiru, M. (2010). Using Mobile Positioning Data to Model Location Meaningful to Users of Mobile Phones. *Journal of Urban Technology*, 17(1), 3-27.

Alexander, L., Jiang, S., Murga, M, och González, M. C. (2015). Origin-destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C*, 58, 240-250.

Algers, S. (2014). Så används nedbrutna socioekonomiska indata i Sampers. TPmod AB.

Algers, S., Mattsson, L-G., Rydergren, C. och Östlund, B. (2009). Sampers – erfarenheter och utvecklingsmöjligheter på kort och lång sikt. Norrköping: Linköpings universitet

Algers, S., Bates, J., Jansson, K., Lang, H., Larsen, O. och Swahn, H. (2013). Towards a model for long distance passenger travel in the context of infrastructure and public transport planning. KTH Arkitektur och samhällsbyggnad, Avdelningen för trafik och logistik. TRITA-TSC-RR 13-013, Report, Stockholm 2013.

Allström, A., Fransson, M., Kristofferson, I., och Gundlegård, D. (2015). Mobilnätdata som indata till prognosmodeller. Rapport, Version 0.91, Sweco, 2015-09-03.

Allström, A., Gidofalvi, G., Kristofferson, I., Prelipcean, A. C., Rydergren C, Susilo, Y. och Widell, J. (2016). Experiences from smartphone based travel data collection – System development and evaluation. Final report from the SPOT-project. Version 1.0, 2016-03-09.

Anderstig, C., Berglund, M., Edwards, H. och Sundberg, M. (2015). PWC Matrices: new method and updated Base Matrices – Final Report. Stockholm: WSP Sverige AB

Aronsson, M., Joborn, M. och Ranjbar, Z. (2015). FBI-BAS. Framtidssäkra BangårdsInvesteringar – Pilot BANGods-Sävenäs. SICS Swedish ICT.

Ben-Akiva, M., och de Jong, G. (2013). The Aggregate-Disaggregate-Aggregate (ADA) Freight Model System. I Ben-Akiva, M., Meersman, H., och Van de Voorde, E. (red.) *Freight Transport Modelling*. 69-90. Emerald Group Publishing Limited.

Berglund, M., och Edwards, H. (2016). Bangods – Beskrivning av befintligt system. WSP Analys & Strategi och Sweco.

Berglund, M., och Edwards, H. (2017). Utvecklingsförslag för Bangods-Samgods. WSP och Sweco, 2017-06-30.

- Bricka, S. och Bhat, C. R. (2006). Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1972, 9-20.
- Bull-Berg, H., Olsson, N., och Sørenssen, A. (2017). Mobiltelefondata – en potentiell datakilde i evaluering av jernbaneutbygginger? Arbeidsrapport. Concept. Maj 2017.
- Calabrese, F., Diao, M., Di Lorenze, G., Ferreira, J., Ratti, C. (2012). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C*, 303-313.
- Cambridge Systematics (2007). PSRC 2006 Household Activity Survey Analysis Report. Final Report, April 2007. Prepared for Puget Sounds Regional Council, Washington State Department of Transportation.
- Canella, O. (2015). Validering av Sampers nulägesmodell 2014 – Basprognos 2016, Stockholms län. 2015-12-17. Stockholm: WSP Sweden AB
- Caseres, N., Wideberg, J.P. och Benitez, F.G. (2007). Deriving origin-destination data from a mobile phone network. *IET Intelligent Transportsystems*, 1(1), 15-26.
- Chen, C-H., Chang, Y-C. och Chen, T-Y. (2008). People Counting System for Getting In/Out of a Bus Based on Video Processing. Proceedings of the Eight International Conference on Intelligent Systems Design and Applications, Kaohsiung, Taiwan, Nov. 26-28, 2008, 3, 565-569.
- Clark, A., Adell, E., Nilsson, A., och Indebetou, L. (2017). Detaljerad kartläggning av verktyg och applikationer för resvaneundersökningar. Trivektor Rapport 2017:32, Version 1.2.
- De Jong, G., Vierth, I., Tavasszy, L., och Ben-Akiva, M. (2013). Recent developments in national and international freight transport models within Europe. *Transportation*, 40, 347-371.
- De Jong, G., Tavasszy, L., Bates, J., Grønland, S.E., Huber, S., Kleven, O., Lange, P., Ottemöller, O., och Schmorak, N. (2016). The issues in modelling freight transport at the national level. *Case Studies on Transport Policy*, 4, 13-21.
- Department for Transport (2009). National Transport Model. High Level Overview. Version 3, September 2009.
- Department for Transport (2015). Road Traffic Forecasts 2015. March 2015.
- Farzin, J. (2008). Constructing an Automated Bus Origin-Destination Matrix Using Farecard and Global Positioning System Data in Sao Paulo, Brazil. *Transportation Research Record: Journal of the Transportation Research Board*, 2072, 30-37.
- Frias-Martinez, V., Soguero, C., och Frias-Martinez, E. (2012). Estimation of Urban Commuting Patterns Using Cellphone Network Data. Proceedings of the ACM SIGKDD International Workshop on Urban Computing, 9-16.
- Frigg, R. och Hartmann, S. (2012). Models in Science. The Stanford Encyclopedia of Philosophy. E. N. Zalta (Red.).
- Frumin, M.S. (2010). Automatic Data for Applied Railway Management: Passenger Demand, Service Quality Measurement, and Tactical Planning on the London Overground Network. Thesis, Master of

Science in Transportation and Master of Science in Operations Research, Massachusetts Institute of Technology, June 2010.

Giuliano, G., Kang, S., Yuan, Q., och Hutson, N. (2015). The Freight Landscape: Using Secondary Data Sources to Describe Metropolitan Freight Flows. METRANS UTC 1-1B, Final Report, December 2015.

Halepovic, E., och Williamson C. (2005). Characterizing and Modeling User Mobility in a Cellular Data Network. PE-WASUN '05 Proceedings of the 2nd ACM international workshop on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks, 71-78.

Holleczek, T., Yu, L., Lee, J.K., Senn, O., Ratti, C., och Jaillet, P. (2014). Detecting weak public transport connections from cellphone and public transport data. Proceedings of the 2014 International Conference on Big Data Science and Computing. Article No. 9.

ITS, Leigh Fischer, Rand Europe, Systra (2016). Rail Demand Forecasting Estimation. Final Report, Final Draft, Redacted. Prepared for Department for Transport, November 2016.

Isaacman, S., Becker, R., Cáceres, R., Kobourov, S., Martonosi, M., Rowland, J., och Varshavsky, A. (2011). Identifying Important Places in People's Lives from Cellular Network Data. I Lyons, K., Hightower, J., Huang E.M. (Red). Pervasive Computing, Pervasive 2011. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg.

Iqbal, M.S., Choudhury, C.F., Wang, P., och González, M.C. (2014). Development of origin-destination matrices using mobile phone call data. *Transportation Research Part C*, 40, 63-74.

Ji, X., Mishalani, R.G. och McCord, M.R. (2015). Transit passenger origin-destination flow estimation: Efficiently combining onboard survey and large automatic passenger count datasets. *Transportation Research Part C*, 58, 178-192.

Joborn, M. (2015). Oplanerade stopp och potential för målpunktsstyrande system. Om trafikala effekter av att införa målpunktsstyrande system för Sveriges godstrafik på järnvägen. Slutrapport från projektet PUMPS – Punktlighet genom målstyrning. SICS Technical Report T2015:01.

Johansen, K. W. och Lindberg, G. (2016). Høhastighetstog i Sverige. Beregningsverktøy og resultater. En vurdering av transport- og samfunnsøkonomiske analyser. TØI rapport 1537/2016, Oslo desember 2016.

Johnson, Z., Psarros, I., Golias, M., Mishra, S. (2014). Developing Freight Performance Measures Using GPS Truck Data. Transportation Research Board 93rd Annual Meeting, Washington DC, 2014.

Liao, C-F. (2014). Using Truck GPS Data for Freight Performance Analysis in the Twin Cities Metro Area. Final Report 2014-14. Minnesota Department of Transportation. Research Services & Library. Office of Transportation System Management.

Lindgren, S., och Vierth, I. (2017). Vad styr valet av trafikslag för godstransporter? En kunskapsöversikt. VTI notat 3-2017.

Lu, X., Bengtsson, L., och Holme, P. (2010). Predictability of population displacement after the 2010 Haiti earthquake. Proceedings of the National Academy of Sciences of the United States of America, 109(29), 11576-11581.

- Lu, X., Wetter, E., Bharti, N., Tatem, A. J., och Bengtsson, L. (2013). Approaching the Limit of Predictability in Human Mobility. *Scientific Reports*, 3, 2923.
- McCormac, E.D., Ma, X., Klocow, C., Currei, A., Wright, D. (2010). Developing a GPS-based truck freight performance measure platform. Final Research Report. April 2010. Transportation Northwest Regional Center X (TransNow), Washington State Transportation Center (TRAC).
- Mcdaniel, J. U.Å. Förvaltning av regional sampersmodell Skåne-TASS Validering av 2016-04-01 modellen. Uppdragsnummer 1320016767. Malmö: Ramböll
- MDS Transmodal (2008). GBFM Version 5.0. Report by MDS Transmodal Limited, 205030 GBFM upgrade report v2, March 2008.
- MDS Transmodal (2013). Rail Freight forecasts to 2023/4, 2033/4 and 2043/4. April 2013, Ref: 212028r9, Final Report. MDS Transmodal Limited 2013.
- Mishalani, R.G, Ji, Y. och McCord, M.R. (2011). Effect of Onboard Survey Sample Size on Estimation of Transit Bus Route Passenger Origin-Destination Flow Matrix Using Automatic Passenger Counter Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2246, 64-73.
- M4Traffic (2015). PM Validering Sampers SAMM och SYDOST.
- M4traffic. (2016). PM – Kollektivtrafiktaxor – långväga järnväg (Sampers Basprognoser ver 1604).
- Munizaga, M. A. och Palma, C. (2012). Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C*, 24, 9-18.
- Naboulsi, D., Fiore, M., Ribot, S., och Stanica, R. (2016). Large-Scale Mobile Traffic Analysis: a Survey. *Communications Surveys and Tutorials*, IEEE Communications Society. Institute of Electrical and Electronics Engineers, 18(1), 124-161.
- NCFRP (2010). Freight-Demand Modeling to Support Public-Sector Decision Making. Report 8. Cambridge Systematics Inc. and GeoStats LLP. Transportation Research Board, Washington D.C.
- NCHRP (2014a). Applying GPS Data to Understand Travel Behavior. Volume II: Guidelines. Report 775. National Cooperative Highway Research Program, Transportation Research Board of the National Academies.
- NCHRP (2014b). Applying GPS Data to Understand Travel Behavior. Volume I: Background, Methods, and Tests. Report 775. National Cooperative Highway Research Program, Transportation Research Board of the National Academies.
- Nelldal, B-O., Jansson, K., och Halldin, C. (2009). Prognoser och samhällsekonomiska kalkyler med Samvips för Götalandsbanan – Underlagsmaterial till Banverket. KTH Arkitektur och samhällsbyggnad, Avdelningen för trafik och logistik. TRITA-TSC-RR 12-002. Rapport, Stockholm 2009.
- Nielsen, B.F., Frolich, L., Nielsen O. A. och Filges, D. (2014). Estimating passenger numbers in trains using existing weighing capabilities. *Transportmetrica A: Transport Science*, 10(6), 502-517.

- Nielsen, A.O., Jensen, A.F., Rich, J., De Jong, G., Berglund, S., Algers, S., Thorhauge, M., Pedersen, T.R., och Bates, J. (2017). Research and development of the European Transport Network Model – Transtools Version 3. Scientific documentation report. Deliverable 2.3.
- Nilsson, C., Thorsson, T. och Engelson, L. (2015). Sampers 3.3 och autokalibrering. Stockholm: WSP Sverige AB.
- Pelletier, M-P., Trépanier, M., och Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C*, 19, 557-568.
- Persson, C., Jiang, S., Wang, Q., Engelson, L. och Thorsson, T. (2015). Framtagande av kalibreringsmål för Sampers regionala modeller. Stockholm: WSP Sverige AB
- Phithakkitnukoon, S., Horanon, T., Di Lorenzo, G., Shibasaki, R., och Ratti, C. (2010). Activity-Aware Map: Identifying Human Daily Activity Pattern Using Mobile Phone Data. Proceedings of the Human Behavior Understanding, First International Workshop, HBU 2010, Istanbul, Turkey, August 22, 2010, 14-25.
- Rashidi, T.H., Abbasi, A., Maghrebi, M., Hasan, S., och Waller, T.S. (2017). Exploring the capacity of social media data for modelling travel behavior: Opportunities and challenges. *Transportation Research Part C*, 75, 197-211.
- Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., och Srivastava, M. (2010). Using Mobile Phones to Determine Transportation Modes. *ACM Transaction on Sensor Networks*, 6(2), Article No. 13.
- RTK (2009). Trafikanalyser – Underlag i arbetet med ny regional utvecklingsplan RUF5 2010 - oktober 2008. Arbetsmaterial 2:2009. Regionplane- och trafikkontoret, Stockholms läns landsting.
- SACOG (2015). Sacramento Activity-Based Travel Simulation Model (SACSIM15): Draft Model Reference Report. Sacramento Area Council of Governments, SACOG. September 2015.
- Safi, H., Assemi, B., Mesbah, M., och Ferreira, L. (2017). An empirical comparison of four technology-mediated travel survey methods. *Journal of Traffic and Transportation Engineering*, 4(1), 80-87.
- SIKA (2009). Analys av kollektivtrafikåtgärder – Jämförande tester med modellsystemen Sampers och Vips. SIKA PM 2009:3, Juni 2009.
- Song, C., Koren, T., Wang, P., och Barabási A-L. (2010a). Modeling the scaling properties of human mobility. *Nature Physics*, 6, 818-823.
- Song, C., Qu, Z., Blumm, N., och Barabási, A-L. (2010b). Limits of Predictability in Human Mobility. *Science*, 327, 1018-1021.
- SOU (2015). En annan tågordning – bortom järnvägsknuten. Slutbetänkande av Utredningen om järnvägens organisation. Statens Offentliga Utredningar, SOU 2015:110. Stockholm 2015.
- Stopher, P. och Shen, L. (2011). In-Depth Comparison of Global Positioning System and Diary Records. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2246, 32-37.

Sun, L., Lu, Y., Jin, J.G, Lee, D-H., Axhausen, K.W. (2015). An integrated Bayesian approach for passenger flow assignment in metro networks. *Transportation Research Part C*, 52, 116-131.

Sweco och M4Traffic (2015). Validering Sampers 2016 – Dokumentet avser validering av regionala Västmodellen av Sampers Basprognos 2014.

Szimba, E., Kraft, M., Schimke, A., Schnell, O., Kawabata, Y., Newton, S., Breemersch, T., Newton, S., Versteegh, R., van Meijeren, J., Jin-Xue, H., de Stasio, C., Fermi, F., och Ihrig, J. (2012). D6 ETISplus Database – Content and Methodology. January 2013.

Trafikanalys (2014). Transportmodeller i ett internationell perspektiv 2014. PM 2014:2.

Trafikanalys (2016a). Resvaneundersökningar som indata till persontransportmodeller – problem, möjligheter och framtida behov i Sverige och Norge. Rapport 2016:21.

Trafikanalys (2016b). Godstransportflöden – Statistikunderlag med varugrupsindelning. PM 2016:3.

Trafikverket (2012). Utveckling av taxor för kollektivtrafik till prognosåren 2030 och 2050. PM 2012-05-03.

Trafikverket (2015a). Beräkningsmetodik för transportsektorns samhällsekonomiska analyser. Kapitel 6 Prognosmodeller och trafikprognoser.

Trafikverket (2015b). Tjänstebeskrivning – Tillgång till utökad detektorinformation 2015-12-08. Version 1.0.

Trafikverket (2015c). Sampers utvecklingsstrategi. Trafikverket, Planering, Expertcenter, Samhällsekonomi och trafikprognoser, Sampers arbetsgrupp. Version 2015-12-13.

Trafikverket (2015d). BVF 592.11 Detektorer – Hantering av larm från stationära detektorer samt åtgärder efter upptäckta skador vid manuell avsyning. TDOK 2014:0689. Version 1.0.

Trafikverket (2015e). Validering av region palt Nord i Sampers Basprognos 2014_Ver 1.

Trafikverket (2016a). Modellanpassade indata- och omvärldsförutsättningar 2016-04-01 reviderad 2017-10-12. Dnr: TRV2015/78279

Trafikverket (2016b). Metoder för framtagande av indata och förutsättningar – Sampers basprognoser 2016-04-01.

Trafikverket (2016c). Representation of the Swedish transport and logistics system in Samgods v. 1.1. Report. Borlänge: Trafikverket.

Trafikverket (2016d). KVAL – Valideringsmetoder för Samgods, Förstudie. 2016-06-21

Trafikverket (2016e). Prognos för persontrafiken 2040. Trafikverkets basprognoser 2016-04-01.

Trafikverket (2016f). PM – Validering av Sampers Basprognoser 2016 -04-01.

Trafikverket (2016g). Tillväxttal transportarbete med kollektivtrafik. 2016-03-31.

Trafikverket (2016h). linjelank_2014_160401.xlsx. <https://www.trafikverket.se/for-dig-i-branschen/Planera-och-utreda/Planerings--och-analysmetoder/Samhallsekonomisk-analys-och-trafikanalys/Kort-om-trafikprognoser/> (hämtad 2017-11-10).

Trafikverket (2016i). linjetabell_2014_160401.xlsx. <https://www.trafikverket.se/for-dig-i-branschen/Planera-och-utreda/Planerings--och-analysmetoder/Samhallsekonomisk-analys-och-trafikanalys/Kort-om-trafikprognoser/> (hämtad 2017-11-10).

Trafikverket (2016j). Tidtabellsmodell järnvägstrafik. PM 2016-02-26.

Trafikverket (2017). Trafiköverslagsgripande plan för utveckling av metoder, modeller och verktyg – för analys av samhällsekonomi, järnvägskapacitet, effektsamband och statistik för trafik- och transportprognoser. Version 1.1. Rapport 2017:087.

Transportrådet (1983). Transporter i Sverige, del III, Godstransportprognoser 1980-2000, TPR Rapport 1983:5.

Trupina, M. (2015). PM Uppdragsnummer 10198510. Göteborg: WSP Sverige AB.

Validering Sampers – Sydost, version 2016-04.
https://www.trafikverket.se/contentassets/ca21d4a110814f23a794c2a6ebb87979/pm_-_validering_sampers_nulage_2014_160401_sydostmod_-_trv_syd.pdf. (Hämtad 2017-11-29).

Vierth, I., Landergren, M., Andersson, M., Brundell-Freij, K., och Eliasson, J. (2016). Uppföljning av basprognoser för person- och godstransporter publicerade mellan 1975 och 2009. CTS Working paper 2016:16, Centre for Transport Studies, Stockholm.

Vierth, I., Lindgren, S., de Jong, G., Baak, J., Hovi, I. B., Berglund, M., och Edwards, H. (2017). Recommendations for a new commodity classification for the national freight model Samgods. CTS Working paper 2017:11. Stockholm: CTS.

VTI (2017). Bor du i Göteborgsregionen med omnejd? Vi behöver din hjälp!
<https://www.vti.se/sv/resa/>. (Läst 2017-11-03).

Wajzman, J., och Nelldal, B-L. (2012). Persontrafik och godstransporter 2010-2030 och kapacitetsanalys för järnväg. KTH Arkitektur och Samhällsbyggnad, Avdelningen för trafik och logistik. Rapport, Stockholm 2012.

Wang, H., Calabrese, F., Di Lorenzo och Ratti, C. (2010). Transportation Mode Inference from Anonymized and Aggregated Mobile Phone Call Detail Records. 13th International IEEE Conference on Intelligent Transportation Systems (September 2010).

Wang, Z., He, S.Y., och Leung, Y. (2017). Applying mobile phone data to travel behavior research: A literature review. *Travel Behaviour and Society* (In press).

Wolf, J. (2000). Using GPS Data Loggers To Replace Travel Diaries In the Collection of Travel Data. Georgia Institute of Technology, July 2000.

Worsley, T. och McNamara, K. (2010). Rail Modelling – integrating rail with other modes. Association for European Transport and contributors 2010. European Transport Conference 2010.

Worsley, T. (2012). Rail Demand Forecasting. Using the Passenger Demand Forecasting Handbook. On the Move – Supporting Paper 2. RAC Foundation, London, December 2012.

Zhao, J. Rahbee, A. och Wilson, N.H.M. (2007). Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems. *Computer-Aided Civil and Infrastructure Engineering*, 22, 376-387.

Zhou, F, och Xu, R-H. (2012). Model of Passenger Flow Assignment for Urban Rail Transit Based on Entry and Exit Time Constraints. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2284, 57-61.

Zhu, W., Hu, H., och Huang, Z. (2014). Calibrating Rail Transit Assignment Models with Genetic Algorithm and Automated Fare Collection Data. *Computer-Aided Civil and Infrastructure Engineering*, 29(7), 518-530.

VTI, Statens väg- och transportforskningsinstitut, är ett oberoende och internationellt framstående forskningsinstitut inom transportsektorn. Huvuduppgiften är att bedriva forskning och utveckling kring infrastruktur, trafik och transporter. Kvalitetssystemet och miljöledningssystemet är ISO-certifierat enligt ISO 9001 respektive 14001. Vissa provningsmetoder är dessutom ackrediterade av Swedac. VTI har omkring 200 medarbetare och finns i Linköping (huvudkontor), Stockholm, Göteborg, Borlänge och Lund.

The Swedish National Road and Transport Research Institute (VTI), is an independent and internationally prominent research institute in the transport sector. Its principal task is to conduct research and development related to infrastructure, traffic and transport. The institute holds the quality management systems certificate ISO 9001 and the environmental management systems certificate ISO 14001. Some of its test methods are also certified by Swedac. VTI has about 200 employees and is located in Linköping (head office), Stockholm, Gothenburg, Borlänge and Lund.

HEAD OFFICE
LINKÖPING
SE-581 95 LINKÖPING
PHONE +46 (0)13-20 40 00

STOCKHOLM
Box 55685
SE-102 15 STOCKHOLM
PHONE +46 (0)8-555 770 20

GOTHENBURG
Box 8072
SE-402 78 GOTHENBURG
PHONE +46 (0)31-750 26 00

BORLÄNGE
Box 920
SE-781 29 BORLÄNGE
PHONE +46 (0)243-44 68 60

LUND
Medicon Village AB
SE-223 81 LUND
PHONE +46 (0)46-540 75 00

