

Thesis 375

Validering av trafikmodeller med GPS-baserade probedata

Axel Persson

Trafik och Väg
Institutionen för Teknik och Samhälle
Lunds Tekniska Högskola
Lunds Universitet



Copyright © Axel Persson

LTH, Institutionen för Teknik och samhälle
CODEN: LUTVDG/(TVTT-5342)/1-85/2022
ISSN 1653-1922

Tryckt i Sverige av Media-Tryck, Lunds universitet
Lund 2022

Examensarbete CODEN: LUTVDG/(TVTT-5342)/1–85/2022

Thesis / Lunds Tekniska Högskola, ISSN 1653–1922
Institutionen för Teknik och samhälle,
Trafik och väg, 375

Author(s): Axel Persson

Title: GPS-baserade probedata som valideringsmetod

English title: GPS probe data as a method of validation

Language Svenska

Year: 2022

Keywords: GPS; Probedata; Trafikmodellering; Validering; Trafikplanering

Citation: Axel Persson, GPS-baserade probedata som valideringsmetod. Lund, Lunds universitet, LTH, Institutionen för Teknik och samhälle. Trafik och väg 2022. Thesis. 375

Abstract:

In traffic assignment, one of the key elements is the validation process. The validation allows the traffic planner to get an understanding of whether the model is working according to reality-based metrics, or if it must be improved. There are several different ways of collecting reality-based metrics, one these methods are called GPS probe data. Through this study, a survey as well as a case study has been performed to discover whether GPS probe data is a type of data suitable for this purpose. The study showed that using GPS probe data is a far more versatile method of collecting reality-based traffic data compared to conventional methods. This since information about a vehicle's location is collected and stored for several points during a journey. These points allow analysis of vehicle movements for the entire network since the vehicles, acting like probes, gathers data during the entire trip. This enables analyses of entire trips and the ability to analyze the entire road network. However, since the GPS probe data is a rather raw type of data set, an extensive amount of data processing is required to carry out analyses of the data. Extensive and rather complicated data processing might reduce the number of traffic planners using the method, which put development of this relatively new method at risk. Therefore, there is a need of creating standardized and more user-friendly methods of processing GPS probe data if the benefits of the data type are going to be taken advantage of in the future of traffic planning.

Trafik och väg
Institutionen för Teknik och samhälle
Lunds Tekniska Högskola, LTH
Lunds Universitet
Box 118, 221 00 LUND

Transport and Roads
Department of Technology and Society
Faculty of Engineering, LTH
Lund University
Box 118, SE-221 00 Lund, Sweden

Förord

Arbetet har genomförts på avdelningen Trafik och väg tillhörande Institutionen för Teknik och samhälle på Lunds Tekniska Högskola. Trots att jag står som ensam författare på detta arbete har jag inte arbetat utan hjälp. Att ta sig igenom detta examensarbete hade inte kunnat gå utan konstant hjälp från mina handledare, Susanna Sjöstrand och Anders Sjöholm på Ramboll och Ulrik Berggren på LTH. Att dessutom få möjligheten att sitta på Rambolls kontor i Malmö och ha förmånen att använda sig av deras nätverk och medarbetare har varit viktigt och bidragande för mitt arbete. Jag vill därmed rikta ett tack till alla inom Rambolls organisation som jag stött på och som hjälpt mig. Ett tack vill jag dessutom rikta till David Gundlegård och Richard Rek på Linköpings universitet som hjälpte mig med både bollande av idéer och Dynameq-licens.

Med detta examensarbete rundar jag av mina fem år som student på Väg- och Vattenbyggnadsprogrammet på LTH. Fem år som har varit oerhört lärorika och roliga, främst på grund av alla fantastiska människor som jag fått lära känna. Till alla dessa vill jag rikta ett extra stort tack, utan er hade dessa år inte varit i närheten av så fantastiska som de blev. Ett minst lika stort tack ska även min familj ha, som under alla åren inklusive dessa fem i Lund har stöttat och hjälpt mig på min resa.

Lund, juni 2022.

Axel Persson

Sammanfattning

För att säkerställa att kvaliteten hos en trafikmodell är tillräcklig för att i största möjliga mån återspegla verkligheten krävs det att utdata från modellen stämmer överens med jämförbara värden som är verklighetsförankrade. En jämförelse mellan modelldata och uppmätt data kallas för validering, vilket är en process som är essentiell vid framställandet av trafikmodeller, liksom alla typer av modeller. En typ av verklighetsbaserade data som kan användas som valideringskälla är så kallad GPS-baserade probedata. GPS-baserade probedata är data som via GPS-mottagare i fordon samlas in under en resa. Datan som samlas in redovisar för varje registrerad punkt bland annat fordonets position, klockslaget för registrering samt fordonets hastighet. Sådan information kan användas för att analysera hela resors egenskaper samt att vissa parametrar såsom hastigheter, ruttval och restider vid utvalda platser, sträckor och områden kan analyseras. En av skillnaderna mellan GPS-baserade probedata och konventionella mätmetoder är att probedatan representerar hela vägnätet medan de konventionella metoderna är stationära vid utvalda placeringar. Däremot representerar inte den GPS-baserade probedatan samtliga trafikanter då insamling endast sker för enstaka procentandelar av resorna. Eftersom GPS-signaler är känsliga mot störningar såsom fysiska hinder och mätfel finns det dessutom en påtaglig risk att andelen resor som representeras utan några felmätningar blir än lägre.

Att använda sig av GPS-baserade probedata i valideringen av trafikmodeller är en metod som jämfört med konventionella metoder är relativt ny. Denna studies syfte var således att undersöka huruvida det är en metod som är lämplig att använda sig av vid validering av trafikmodeller tillägnade biltrafik. Studien bestod av två huvuddelar: en enkätstudie och en fallstudie för Region Stockholm. I enkäten fick personer som arbetar med trafikrelaterade frågor på konsultfirman Ramboll svara på frågor angående erfarenheter samt eventuellt potential med att använda GPS-baserade probedata i validering av trafikmodeller. Fallstudien bestod av att använda inköpt probedata från företaget INRIX för att validera en befintlig trafikmodell i det dynamiska trafikmodelleringsprogrammet Dynameq för Region Stockholm. Då jämfördes utdata från modellen med registrerad probedata för parametrarna hastigheter, restider och ruttval.

Resultaten från de två delarna av studien indikerade att det finns goda möjligheter att använda probedata för att jämföra värden på utdata från trafikmodellen med den registrerade datan från verkliga fordonsrörelser. Flera av de tillfrågade svarade även i enkäten att de ansåg GPS-baserade probedata vara en datakälla som det finns god potential för att användas mer. Arbetsprocessen i fallstudien påvisade dessutom metodens medförande flexibilitet tack vare möjligheten att analysera nästintill alla delar av vägnätet. Däremot noterades det i studien att det på grund av datans storlek och format krävs omfattande arbete i exempelvis programmeringsverktyg för att hantera datan och möjliggöra analysering, då datasetet bestod av cirka 24,5 miljoner punkter. Förmågan att både kunna genomföra komplext arbete i programmeringsverktyg samt besitta goda kunskaper inom trafikplanering är omfattande. Detta kan begränsa omfattningen av arbetet med metoden och därmed även möjligheterna för utveckling.

Genom att väga fördelar och nackdelar enligt både enkät- och fallstudie blev slutsatsen att GPS-baserade probedata kan ge jämförbara och lämpliga värden för att användas inom validering av trafikmodeller. Däremot finns det behov av att införa standardiserade metoder för databearbetningen som krävs, dels för att ge möjligheten till mer konsekventa och säkra resultat, dels att göra metoden mer användarvänlig och därmed effektivare.

Summary

In traffic assignment, a vital part is to ensure that the quality of the traffic model is sufficient. To achieve this, a validation needs to be implemented where results and values from the traffic model are compared with corresponding data gathered from reality-based traffic metrics. One type of reality-based data which can be applied for traffic model validation is so called GPS probe data. GPS probe data is based on information from separate vehicles equipped with GPS-receivers. For each point where the GPS-receiver register its location, information such as position, speed and time stamp is accounted for and all these points can be put together to create complete journeys, which can be analyzed and compared to values from the traffic model. Parameters such as speed, travel times and route choices inferred from probe data are suitable for comparison with the traffic model. One of the main differences between using GPS probe data and conventional traffic measurements is that GPS probe data is collected throughout the entire road network, while more conventional methods are stationary and therefore merely collects data on single points in the network. However, GPS probe data only represent a small share of the total amount of vehicles, often a few percent of the total fleet. Since GPS-signals are sensitive to noise and disruption such as physic obstacles or instrumental errors, the risk of a decreasing number of trips without errors together with the already low penetration rate may significantly impede representativity of the data.

Using GPS probe data as a source for validation is a method that, compared to conventional methods, is relatively new. Thus, the purpose of this study was to evaluate whether GPS probe data is suitable to use as empirical basis for validating traffic models. The study consisted of two main parts, a survey, and a case study including validation of a dynamic traffic model in the dynamic traffic assignment program Dynameq using GPS probe data. For the survey, specialists working with traffic related assignments at Ramboll took part by sharing their experiences from working with GPS probe data as a source of validation as well as their thoughts on the methods potential. For the case study, GPS probe data collected by INRIX and purchased by Ramboll was used for comparison with the output from the Dynameq model. The parameters analyzed in this study was speed, travel times and route choices.

The results from the two main parts indicated good possibilities of retrieving data suitable for comparison between the GPS probe data set with outputs from a traffic model. In the survey, several of the responders considered GPS probe data to have a potential of being used more extensively compared to today. The process of working with the GPS probe data and comparing them to the output from the Dynameq model indicated advantages including being able to analyze essentially the entire road network. However, the process of working with the GPS probe data was also perceived as rather extensive. Because of its sheer size, in the case study a data set containing around 24,5 million points, the data had to be processed in a programming tool to enable the analysis of the results. The ability of both being able to do complex work in programming tools as well as possess great knowledge in traffic planning is highly comprehensive. This is a competence which a great number of

people do not possess, which could prevent the method from further development and reaching its full potential.

After considering the advantages together with the disadvantages, the conclusion was that GPS probe data is a data source which can provide suitable and comparable values to be used for validating traffic models. However, there is a great need of establishing standardized methods for the data processing needed to be able to analyze the results. This could improve the possibility of getting consequent and reliable results. In addition, it would make the data processing more user friendly, increasing the efficiency of working with GPS probe data.

Innehållsförteckning

Förord	II
Sammanfattning	III
Summary	V
1 Inledning	1
1.1 Bakgrund	1
1.2 Syfte och mål	2
1.3 Avgränsning	2
1.4 Rapportens disposition	3
1.4.1 Teori	3
1.4.2 Metod	3
1.4.3 Resultat	3
1.4.4 Diskussioner	3
1.4.5 Slutsatser	3
1.4.6 Rekommendationer	4
2 Teori	5
2.1 GPS-baserade probedata	5
2.1.1 Positionering med hjälp av GPS	5
2.1.2 GPS-baserade probedata	6
2.1.3 Databeskrivningar	7
2.1.4 Penetrationsgrad	8
2.1.5 Samplingsintervall	9
2.1.6 Felkällor	9
2.2 Trafikmodellering	10
2.2.1 Fyrstegsmodeller	10
2.2.2 Validering av trafikmodeller	11
2.2.3 Olika nivåer och typer av trafikmodeller	11
2.2.4 Dynamisk trafikmodellering	12
2.3 Trafikverkets modell för Region Stockholm i INRO Dynameq	15

3	Metodik och metoder	17
3.1	Metodik	17
3.2	Metod för enkätstudie	17
3.2.1	Utformning av enkät	17
3.2.2	Frågor	18
3.2.3	Tillfrågade personer	20
3.3	Metod för fallstudie Region Stockholm	20
3.3.1	Arbetsgång i INRO Dynameq	20
3.3.2	Hantering av data	20
3.3.3	Validerande parametrar	21
3.3.4	Analys	27
4	Resultat	29
4.1	Enkätstudien	29
4.2	Fallstudie Region Stockholm	31
4.2.1	Hastigheter	31
4.2.2	Restider	38
4.2.3	Ruttval	43
5	Diskussion	49
5.1	Resultatdiskussion	49
5.1.1	Enkätstudie	49
5.1.2	Fallstudie Region Stockholm	49
5.2	Metoddiskussion	53
5.2.1	Enkätstudien	53
5.2.2	Fallstudie Region Stockholm	53
5.3	Sammanfattande diskussion	54
6	Slutsatser	57
7	Rekommendationer	59
	Referenser	61
	Bilagor	64

1 Inledning

1.1 Bakgrund

Vid trafikplanering baserad på kvantitativa trafikdata från simulerings- och prognosmodeller är ett av de viktigaste momenten att säkerställa kvaliteten hos använda trafikmodeller. En trafikmodells uppgift är att återspegla hur trafiken beter sig och således är det viktigt att modellen som används vid trafikplanering stämmer överens med verkligheten på bästa sätt. En utmaning i detta är att trafikens verkliga förutsättningar kan variera beroende på vecka, veckodag, tid på dygnet etcetera. Hur väl en trafikmodell fungerar beror således till stor del på kvaliteten på den indata som modellen bygger på, samt den empiriska data som modellen valideras mot.

Empiriska data som är av relevans för validering av trafikmodeller visar exempelvis vilka ruttval trafikanter gör, medelhastigheter, restider, flöden på olika länkar och vilken efterfrågan som finns hos trafikanterna (Olstam, Köhler, Tapani, Perez & Wahlstedt 2014). För att samla in sådan data finns det idag flertalet olika metoder. En metod som är väl beprövad är enligt Wärnhjelm (2005) resvaneundersökningar där enskilda trafikanter bland annat får besvara hur de reser gällande rutt, reslängd, val av färdmedel och anledning till resa. Svaren från dessa undersökningar ger tydliga och säkra svar, men det är enligt Wärnhjelm (2005) tidskrävande och dyrt att genomföra sådana undersökningar, och således begränsas ofta dess omfattning. Enligt Eriksson et al. (2018), har resvaneundersökningar på senare år dessutom fått allt större problem med sjunkande svarsfrekvens, exempelvis från 68% till 32% för Trafikverkets nationella resvaneundersökning mellan 2005 och 2016. På grund av dessa anledningar blir resvaneundersökningar en metod som är relativt ineffektiv. I Trafikverkets (2021a) Forsknings och innovationsplan beskrivs just behovet av att utveckla effektiva och flexibla metoder och således är det av vikt att hitta andra metoder för att samla in data.

Enligt Wärnhjelm (2005) brukar resvaneundersökningar ofta kompletteras med data från olika typer av trafikmätningar, vilket är mätningar som sker vid fasta punkter i vägnätet. Där mäts antalet passager förbi punkten och vid registreringen av de passerande fordonen mäts även hastigheten. Trafikmätningar kan därmed användas för att få svar på hur flödena och hastigheterna ser ut förbi olika punkter, men eftersom det enbart sker mätningar på enskilda platser går det inte att få svar på hur fordonets kompletta förflytningsmönster ser ut.

En typ av data som på senare år har börjat användas i allt större omfattning inom trafikmodellering är GPS-baserade probedata (GPS: Global Positioning System). Data samlas in från enskilda fordon genom att deras ofta inbyggda GPS-mottagare registrerar var fordonet befinner sig under resans gång. Genom att analysera den registrerade probedatan går det enligt Ahlberg et al. (2021) att bland annat se start- och slutpunkter, ruttval, restider och hastigheter. Eftersom data samlas in under hela resans gång täcks därmed större delar av nätverket och inte enbart vid några utvalda punkter (Mitrovich, Montesano, Pugliese, Mancini & Valenti, 2010). Registreringen av data sker under dygnets alla timmar och

därmed går det att använda probedata för att analysera trafikens situation vid olika tidpunkter på dygnet, och varje förflyttning över tid. Probedata har således potentialen att användas inom många områden av trafikmodellering.

Enligt Toledo och Koutsopoulos (2004) är en av de viktigaste delarna inom trafikmodellering valideringen av den färdiga modellen, vilket tillsammans med kalibrering utgör det sista steget vid framställningen av en trafikmodell. Vid valideringen av en trafikmodell jämförs modellens utdata med värden som är verklighetsförankrade. Parametrar som kan jämföras är restider, hastigheter, ruttval och flöden etcetera. Värdena som används vid valideringen baseras traditionellt på mätningar genom exempelvis slangmätningar eller radarsensorer såsom MCS (Motorway Control System). Dessa mätningar sker enbart vid fasta punkter och således ger det data för flöden och hastigheter vid dessa punkter. Eftersom valideringen är en av de vitala delarna inom trafikmodellering är det fördelaktigt om datan som ligger till grund för valideringen går att tillämpa för olika parametrar, olika platser och tidpunkter för att undvika att modellen enbart fungerar lokalt. Probedata bör ha potentialen att användas i valideringssyfte just eftersom datan är verklighetsbaserad samt att det med probedata går att analysera flertalet olika parametrar vid olika förhållanden, platser och tidpunkter.

GPS-baserade probedata har potentialen att användas i ökad omfattning då antalet fordon som har GPS-mottagare installerade ökar, enligt Mitrovich et al. (2010), vilket kan ge större möjlighet till fler användningsområden. Eftersom metoden är relativt ny finns det intresse från branschen att sammanställa erfarenheterna av att arbeta med probedata inom trafikmodellering. En sådan sammanställning bör innehålla vilka erfarenheter branschen har av GPS-baserade probedata i valideringssyfte, hur probedata utnyttjas i dagsläget samt vilken kunskap som behövs för att kunna utveckla metoden.

1.2 Syfte och mål

Syftet med denna studie har varit att undersöka huruvida GPS-baserade probedata är en datakälla som kan anses vara lämplig att använda vid validering av dynamiska trafikmodeller. Målet med arbetet var dels att genom en enkätstudie och en fallstudie få en bättre bild över vad som fungerar bra och mindre bra vid validering av dynamiska trafikmodeller med GPS-baserade probedata. Genom att belysa styrkor och svagheter med den undersökta metoden var målet dels att bidra till effektiviserad och förbättrad trafikmodellering. Genom den fallstudie som genomförts var målet dessutom att ge ett utvärderande underlag för hur väl Dynameq som verktyg, samt den befintliga modellen, fungerar i jämförelse med den verklighetsbaserade probedatan.

1.3 Avgränsning

Skapandet och valideringen av en väl fungerande trafikmodell består av flertalet steg och omfattar arbete i flera olika steg. Ett exempel på en process för att framställa modeller och dess steg redovisas i Fyrstegsmodeller. Således var det nödvändigt att begränsa vilka delar av trafikmodelleringens arbete som beaktades i studien. För denna studie var det GPS-baserade probedatas tillämpbarhet i valideringsprocessen som fokuset lades på. Eftersom det endast

var valideringsprocessen som var av relevans för studien bestämdes det även att enkätens omfattning skulle begränsas. Enkäten omfattade ursprungligen samtliga delar av trafikmodellering där det sista avsnittet var fokuserat på validering av trafikmodeller med GPS-baserade probedata. Därmed var det endast detta senare avsnitt som var av intresse för denna studie. I fallstudien av Region Stockholm fokuserades det även där endast på att validera Trafikverkets trafikmodell för Region Stockholm i Dynameq.

Vid validering av trafikmodeller finns det flera olika parametrar som kan jämföras och analyseras. För att begränsa arbetet bestämdes det att tre olika parametrar skulle användas i fallstudien, dels för att ha lämplig storlek på arbetet, dels för att kunna fokusera på de delar som enligt teorin och enkäten anses mest lämpliga för GPS-baserade probedata. De tre parametrar som analyserades var hastigheter, restider och ruttval.

1.4 Rapportens disposition

1.4.1 Teori

Här beskrivs den teoretiska ram som studien bygger på. Teorin som redovisas fokuseras på GPS, trafikmodellering, tidigare studier rörande trafikmodellering och den dynamiska trafikmodell som användes i studien.

1.4.2 Metod

I detta kapitel beskrivs de i studien ingående metoderna. Här beskrivs hur undersökningarna och arbetsgångarna för enkätstudien respektive fallstudien såg ut, samt hur analysen av dess resultat genomfördes.

1.4.3 Resultat

Här redovisas och tolkas svaren som enkätens deltagare framförde i enkätstudien. Resultaten och analys av resultaten från fallstudien redovisas även i detta kapitel.

1.4.4 Diskussioner

I detta kapitel diskuteras resultaten som de två studierna påvisade utifrån den teori som arbetet baserades på. I kapitlet redovisas även slutsatserna baserade utifrån en diskussion av resultaten kopplade till syftet och målet med studien. Det finns i kapitlet även förslag på vidare arbeten och studier.

1.4.5 Slutsatser

Slutsatserna baserade utifrån diskussion av resultaten kopplade till syftet och målet med studien redovisas i detta kapitel.

1.4.6 Rekommendationer

I detta kapitel redovisas förslag på vidare arbeten och studier.

2 Teori

2.1 GPS-baserade probedata

2.1.1 Positionering med hjälp av GPS

Enligt Geotab (2020) är GPS en förkortning av det engelska uttrycket Global Positioning System. GPS används för att få information om positionering, hastighet samt att det kan ge möjlighet till att synkronisera tiden för olika enheter. För att använda sig av GPS behövs tre sorters resurser, den första är satelliter i omlopp runt jorden som sänder ut signaler. Sedan behövs det även GPS-mottagare som exempelvis telefoner och navigatorer, vilka tar emot, hanterar och tolkar GPS-signalerna. Till sist behövs även stationer på jorden som övervakar, följer och kontrollerar satelliterna (Geotab, 2020).

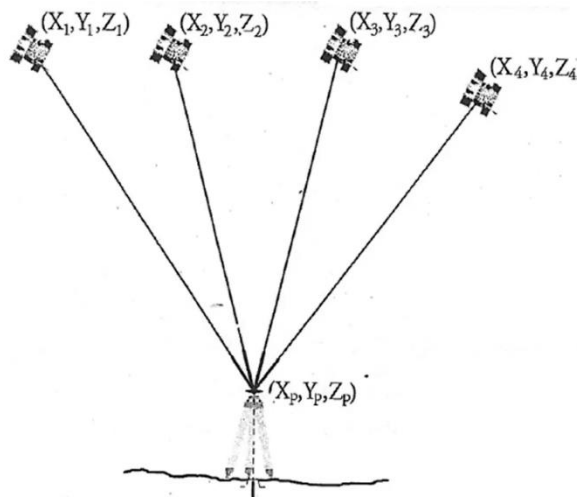
För att en mottagare för GPS-signaler ska kunna bestämma mottagarens position används, enligt Jamali (2021) avståndet från satelliten till mottagaren. Satellitens position är känd och det finns två metoder att ta reda på avståndet mellan satelliten och mottagaren. Den ena metoden är mätning av kodtyper. Dessa baseras på att mottagaren analyserar GPS-signalens egenskaper däribland tidpunkterna för utsänd och mottagen signal. Med det får mottagaren fram den totala tiden som signalen färdades mellan satelliten och mottagaren. Med ekvation 1 där d betecknar avståndet, c är ljusets hastighet och Δt är tiden det tar för signalen att färdas kan avståndet mellan mottagare och satellit bestämmas (Jamali, 2021).

$$d = c * \Delta t \quad (\text{Ekvation 1})$$

En annan metod för att bestämma avståndet mellan mottagare och satellit är med bärvågsmätning. För bärvågsmätning är det våglängdernas längd och antal som bestämmer avståndet mellan mottagaren och satelliten. GPS-signalens våglängd är konstant och om mottagaren registrerar antalet våglängder som signalen mellan de två punkterna består av går det således att få fram avståndet. Avståndet bestäms genom ekvation 2 där d betecknar avståndet, λ betecknar våglängden och φ är storleken på den sista fasen (Jamali, 2021).

$$d = n * \lambda + \varphi * \lambda \quad (\text{Ekvation 2})$$

Eftersom det endast är avståndet mellan satellit och mottagare som går att urskilja krävs det således att mottagaren tar emot signaler från flera satelliter samtidigt eftersom positionering utgörs av koordinater för longitud och latitud samt höjddled. Eftersom det är tre variabler som behöver bestämmas för att få korrekt positionering krävs det således minst tre ekvationer för att kunna bestämma samtliga variabler. Därmed krävs det att mottagaren tar emot signaler från minst tre satelliter. Ofta används fyra satelliter där den fjärde satelliten nyttjas för att synkronisera mottagarens klocka. Figur 1 illustrerar förhållandet mellan mottagaren och satelliterna (Jamali, 2021).



Figur 1. Illustration av förhållandet mellan mottagare och satelliter. Bildkälla: Jamali (2021). Beviljat tillstånd från källa.

Med Pythagoras sats går det att skapa ett ekvationssystem med tre okända variabler som utgör mottagarens positionering, och där satelliternas positionering och avstånd från mottagaren är kända. Utifrån detta ekvationssystem går det att bestämma de okända variablerna och därmed mottagarens position (Jamali, 2021). Ekvationssystemet redovisas i ekvation 3.

$$\begin{cases} d_1 = \sqrt{(x_1 - x_p)^2 + (y_1 - y_p)^2 + (z_1 - z_p)^2} \\ d_2 = \sqrt{(x_2 - x_p)^2 + (y_2 - y_p)^2 + (z_2 - z_p)^2} \\ d_3 = \sqrt{(x_3 - x_p)^2 + (y_3 - y_p)^2 + (z_3 - z_p)^2} \end{cases} \quad (\text{Ekvation 3})$$

I ekvationssystemet är det variablerna x_p , y_p och z_p som beskriver mottagarens koordinater.

2.1.2 GPS-baserade probedata

I trafiken blir det allt vanligare att fordon är uppkopplade via smarta enheter såsom navigatorer. Dessa smarta enheter har ofta GPS-mottagare som registrerar platsinformation såsom koordinater och hastigheter. Dessa GPS-mottagare används för flera olika syften såsom styrning och kontroll av fordonsflottan (Ahlberg et al. 2021) och enligt Mitrovich et al. (2010) av försäkringsföretag och i ett säkerhetssyfte. GPS-data insamlat från enskilda fordon kallas för GPS-baserade probedata, där varje enskilt fordon agerar som en sond (probe, på engelska, vilket är den term som kommer användas i denna studie).

När ett fordon agerar som en sond registrerar dess GPS-mottagare flera punkter under en resa. På grund av att fordonen registrerar flera punkter under en och samma resa möjliggör probedata att användas inom trafikplanering och trafikmodellering. Med hjälp av spridningen av mätpunkter går det därmed att analysera hur trafikanter väljer att resa på samtliga delar av vägnätet där registrerade fordon färdats. Användning av probedata är till skillnad från mätningar som sker vid fasta punkter såsom radarsensorer och kameror, ett flexibelt sätt att samla in data från trafiken (Mitrovich et al. 2010). Den flexibla inmätningen

bidrar till en större täckningsgrad för trafiknätet och därmed blir utökad möjligheten att skapa trafikmodeller som fungerar väl vid fler platser.

2.1.3 Databeskrivningar

För att kunna analysera och använda GPS-baserade probedata krävs det att den insamlade datan bryts ner i olika delar. I rapporten *Probedata, förstudie kring användning av GPS-baserade probedata för skattning av hastigheter, länkflöden och ruttval* använder sig Ahlberg et al. (2021) använder sig av data inköpt från INRIX¹ och är den data som används i fallstudien i detta arbete. Datan beskrivs med fyra olika databeskrivningar: ”Waypoints”, ”Trips”, ”Trajectories” och ”Through” (svenska för ”vägpunkter”, ”resor”, ”resvägar” och ”genomgående”, dessa kommer fortsatt att benämnas enligt den engelska nomenklaturen).

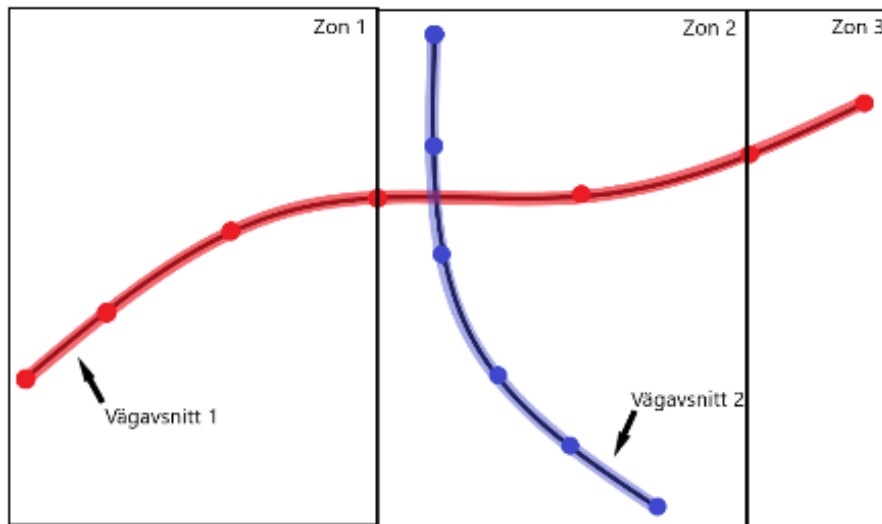
Waypoints är all samlade rådata från gällande dataset. När ett fordon färdas kommer det kontinuerligt registreras information såsom enhets-ID, fordonets lokalisering, tidpunkt och hastighet. Genom att slå ihop de registrerade mätpunkterna tillhörande samma enhets-ID går det att få fram hur enskilda resor ser ut (Trips). För att olika resor genomförda av samma fordon inte ska slås ihop till samma resa kan en resa definieras som avslutad när fordonet inte rört sig mer än en bestämd sträcka under en bestämd tid. Ahlberg et al. (2021) definierar resan som avslutad då fordonet som mest färdats 200 meter under tio minuters tid. För att få information om vilka länkar (vägavsnitt) som resorna genomförs via (Trajectories) utförs så kallad kartmatchning. Eftersom det vid varje waypoint registreras platsinformation kan dessa matchas mot befintliga nätverk, för probedatan från INRIX matchas punkterna med kartor från Open Street Map². Med liknande metod kan waypoints användas för att analysera genom vilka zoner i modellen som fordonen färdas för de enskilda resorna (Through), exempelvis mindre områden i en stad. Probedatan från INRIX som analyserats i fallstudien använder sig av Sampers³ zonindelning.

Hur de olika databeskrivningarna beror på varandra redovisas i Figur 2. De markerade prickarna är registrerade waypoints, färgerna markerar olika resor (trips) där även kartmatchning (trajectories) markeras. Illustreringen är indelad i tre zoner.

¹ INRIX är en leverantör av trafikdata. De levererar bland annat GPS-baserade probedata (INRIX, 2022).

² Open Street Map är en leverantör av kartdata (Open Street Map, 2022).

³ Se avsnitt 2.3 Trafikverkets modell för Region Stockholm, för beskrivning av Sampers.



Figur 2. Illustrering av förhållandet mellan Waypoints (prickar), Trips (olika färger), Trajectories (färg utmed vägavsnitt) och Through (genomfart av olika zoner).

Utifrån ovan nämnda databeskrivningar går det enligt Ahlberg et al. (2021) bland annat att analysera hastigheter, restider, ruttval, reslängd och länkflöden.

2.1.4 Penetrationsgrad

De fordon vars resor registrerar waypoints och som används som probedata utgör endast en mindre andel av alla fordon som färdas i nätverket, vilket benämns som penetrationsgrad. Penetrationsgraden beräknas som kvoten av antalet registrerade resor dividerat med totala antalet resor.

Värden på det totala resandet brukar erhållas från statistiska trafikmätningar, såsom MCS. MCS-mätningarna jämförs med antalet resor registrerade hos probedatan för samma snitt i nätverket. Genom att upprepa processen för flertalet mätpunkter kan en genomsnittlig penetrationsgrad beräknas. Den beräknade penetrationsgraden kan i sin tur användas tillsammans med probedatan för att beräkna flöden på länkar där det inte finns några statistiska trafikmätningar (Markovic, Sekula, Laan, Andrienko, G., Andrienko, N., 2018). Att beräkna en penetrationsgrad som väl avspeglar verkligheten är således av stor vikt för att kunna få fram korrekta länkflöden med hjälp av probedata. Det är dock ofta svårt att beräkna en exakt penetrationsgrad då det inte är konsekvent lika stor andel av fordonen som registrerar och samlar in GPS-data. Den beräknade penetrationsgraden är dessutom endast lokal eftersom de jämförande MCS-mätningarna är stationära.

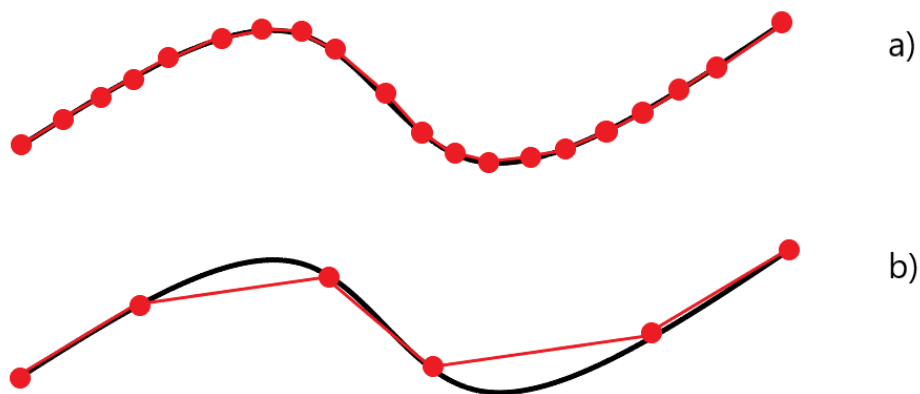
Enligt Ahsani, Amin-Naseri, Knickerbocker och Sharma (2018) är en hög penetrationsgrad att föredra sett utifrån hur verklighetstroga värdena hos probedatan är. Fler fordon medför att resandet som avspeglas i probedatan blir mer representativt för samtliga resor och därmed säkrare.

Enligt Ahlberg et al. (2021) har den data som analyseras deras förstudie och således även i detta arbete en genomsnittlig penetrationsgrad mellan cirka 1–1,5 %. Penetrationsgraden kan

däremot variera beroende på tid på dygnet samt att det finns skillnader beroende på vilka vägar som analyseras.

2.1.5 Samplingsintervall

Varje enskild resa består av ett antal waypoints som registrerats mellan start och mål. Hur många waypoints som en resa består av beror på hur ofta fordonet registrerar sin position. Intervallet mellan varje registrering kallas för samplingsintervall. Intervallet är inte konstant för en resa utan samplingsintervallet beräknas som ett medelvärde för varje enskild resa. Kortare samplingsintervall är att föredra över längre samplingsintervall då noggrannheten blir högre desto fler punkter som registreras. Exempelvis blir medelhastigheten för hela resan mer representativ i förhållande till den verkliga resan samt att kvaliteten vid kartmatchningen blir bättre (Ahlberg et al. 2021). I Figur 3 redovisas hur kartmatchningen beror på samplingsintervallet.



Figur 3. Illustration över olika samplingsintervall. De röda markeringarna är registrerade waypoints. a) har ett kortare samplingsintervall jämfört med b).

2.1.6 Felkällor

Instrument som vid positionsmätning använder sig av GPS-signaler utsätts konstant för olika typer av störningar. Sådana störningar kan exempelvis vara att klockslagen då signalen sänds ut och tas emot av mottagaren eller satelliterna inte är korrekta, att hinder såsom byggnader och träd reflekterar signalen eller att laddade partiklar i jonosfären påverkar signalens frekvens (Lantmäteriet, 2022).

På grund av att GPS-signaler ofta påverkas av felkällor kommer det även att finnas waypoints i ett probedata-set där positionering, hastighet eller klockslag inte är korrekt. Det kan förekomma att mottagningen inte fungerar utmed vissa delar av resan och att mätpunkter därmed uteblir. Om det uppstår fel med positioneringen riskerar kartmatchningen att påverkas negativt, vilket kan medföra förlorad möjlighet att analysera länkflöden och ruttval på vissa länkar. Om exempelvis klockslagen som registreras är felaktiga kan medelhastigheter påvisa resultat som inte är korrekta. Ahlberg et al. (2021) noterar i studien *Probedata, förstudie kring användning av GPS-baserade probedata för skattning av hastigheter, länkflöden och ruttval* att det vanligaste felet som uppstår vid de olika punkterna är att de hamnar för långt ifrån vägnätet. Andra vanligt förekommande fel är att resan som registreras är kortare än en meter samt att den projicerade hastigheten är över 250 km/h, vilket anses vara orimligt högt. I enstaka fall registrerades det negativa hastigheter för vissa punkter.

2.2 Trafikmodellering

För att kunna analysera effekter av exempelvis infrastrukturförändringar eller hur trafiksituationen kommer förändras i framtiden behövs trafikmodeller. En trafikmodell är en förenkling av verkligheten och har som syfte att göra det möjligt att analysera trafiken. Trafikmodeller kan både användas för det befintliga nätverket samt som verktyg för analys av hur framtida åtgärder och trafikförändringar kommer påverka trafiken. Beroende på behovet av analysform och detaljnivå finns det olika typer av modeller, vilket beskrivs i delkapitel 2.2.3 Olika nivåer och typer av trafikmodeller. Det går även att använda sig av olika typer av datainsamlingsmetoder av indata till modellframtagning, kalibrering och validering av en modell. Nedan beskrivs fyrstegsmodeller, vilket är den typ av modell som använts i denna studie.

2.2.1 Fyrstegsmodeller

Vid framtagning av en ny trafikmodell används olika typer av data som tillsammans skapar en uppskattning över hur människor väljer att resa. En typ av modell som syftar till att återskapa hela processen i hur resor genereras och genomförs är så kallade fyrstegsmodeller (Lindborg, Johansson & Sundbergh, 2016). Fyrstegsmodeller består av fyra huvudsakliga delar, steg:

1. Resegenerering
2. Val av destinationer
3. Val av färdmedel
4. Ruttval

I det första steget är målet att ta fram hur många resor som dagligen genereras i det undersökta området. Detta kan exempelvis bero på befolkningens mängd och invånarnas socioekonomiska situation. Nästa steg går ut på att uppskatta mellan vilka punkter som de genererade resorna kommer att bedrivas. Hur resorna fördelas redovisas i så kallade OD-matriser (Origin Destination-matriser), vilket beskriver start och målpunkter. Hur resornas start- och målpunkter fördelas kan bland annat bero på vilka verksamheter som finns i respektive område samt hur många som arbetar eller bor där. I det tredje steget uppskattas fördelningen mellan olika färdmedel hos de resor som genererats. Valet av färdmedel är komplext och beror ofta på personliga preferenser. Generellt brukar färdmedelsvalet bero på tillgång till bil och den upplevda kostnaden för resa med ett visst trafikslag. Den upplevda kostnaden beror dels på den monetära kostnaden såsom bränslekostnader, kostnad för bussbiljetter och diskonterade investeringskostnader för eget fordon. Dels beror det även på den upplevda restiden, vilket bland annat innefattar faktisk restid (åktid i fordon), komfort, fysisk ansträngning, byte av buss eller tåg och anslutningar. Inför det sista steget har det därmed uppskattats hur många resor som genomförs, mellan vilka destinationer resorna sker samt färdmedelsfördelningen. Det sista steget är att ta fram vilka rutter som trafikanterna kommer att färdas på för att ta sig till sin målpunkt. I ett nätverk av vägar finns det många olika alternativ på rutter som tar en person från och till samma destinationer (Lindborg, Johansson & Sundbergh, 2016). Det är viktigt att uppskatta hur nätverksutläggningen ser ut

bland annat för att veta vilka länkar det kan uppstå trängsel och köer på. Dessa fyra steg brukar följas av en validering av modellen.

2.2.2 Validering av trafikmodeller

När det sista steget i modellframställandet är genomfört behöver modellen genomgå en validering. Enligt Toledo och Koutsopoulos (2004) är valideringen ett viktigt steg i framställandet av trafikmodeller eftersom det är då modellen testas utifrån hur väl den fungerar sett till verkligheten. Validering av trafikmodeller går ut på att jämföra utdata från modellen med, idealt, data från en eller flera utomstående källor som i många fall är verklighetsbaserade (Travel Forecasting Resource, 2022). Om valideringen visar på att modellens utdata inte stämmer överens med värdena som den valideras mot, kalibreras modellen och efter det följer en ny validering. Denna process upprepas tills det att modellen visar på önskade resultat jämfört med de validerande värdena. Data som används i valideringen kan dels vara samma typ av källa som använts vid framtagningen av modellen, men inte med exakt samma data, dels data från andra källor. Att enbart använda sig av samma sorts data vid valideringen bör enligt Lindborg, Johansson och Sundbergh (2016) undvikas. Om modellen är både skattad och validerad med samma data kan det bli missvisande resultat då risken är stor att den inte fungerar vid andra omständigheter. Toledo och Koutsopoulos (2004) anser det vara viktigt att valideringen sker med data från andra källor som är oberoende av modellen. De förespråkar dessutom att inte enbart validera resultaten från modellen, utan att även validera den data som modellen skattas ur.

Vilka typer av värden som används för validering beror på var datan härstammar från för typ av mätning. Om datan exempelvis kommer från statiska trafikmätningar såsom MCS eller slangmätningar, kan flöden och hastigheter utgöra parametrar som jämförs mellan modell och mätvärden. Dessa typer av mätningar ger god täckning för de platser som mätinstrumenten är installerade vid. Däremot finns det många parametrar som dessa inte tar hänsyn till. Eftersom mätningarna är stationära kan det enligt Jelenius, Kristoffersson och Fransson (2017) uppstå osäkerhet kring hur väl valideringen fungerar för de delar av trafiksystemet som inte täcks av mätpunkterna.

En av skillnaderna mellan GPS-baserade probedata och statiska trafikmätningar såsom MCS och slangmätningar är att det för probedatan är de enskilda fordonen som agerar mätpunkter. Således kommer spridningen av mätpunkter att bli betydligt större och täckningen av trafiknätet blir bättre (Jelenius, Kristoffersson & Fransson, 2017). Eftersom fordonen konstant rör på sig och samlar in information kontinuerligt kan probedata som samlas i sin tur användas för att validera modellens ruttval, restider, hastigheter, trängsel och köer, beroende på vilken typ av modell som ska valideras.

2.2.3 Olika nivåer och typer av trafikmodeller

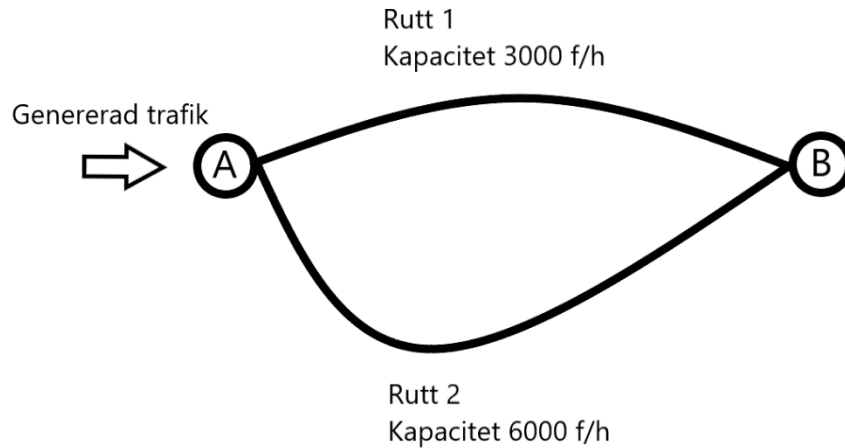
Beroende på vad trafikmodellen har för syfte och behov av detaljnivå finns det enligt Söder (2020) olika nivåer att tillgå. De tre huvudsakliga nivåerna av trafikmodellering är mikro-, meso-, och makronivå. Om det exempelvis behövs en trafikmodell som ska ge svar på hur en specifik del av trafiknätet, såsom en viss cirkulationsplats eller korsning, påverkar köbildning och enskilda trafikanter är det enligt Söder (2020) lämpligt med en modell på mikronivå. Vidare om modellen ska användas för att ta reda på hur pendlingstrafiken

regionalt ser ut, för generella flöden och efterfrågan för en stad, används modeller på makronivå. Mellan mikro- och makronivå finns den så kallade mesonivån (Söder, 2020), där skillnaden är detaljnivån. Mikroskopiska modeller används för att analysera enskilda lösningar i vägnätet medan de mesoskopiska modellerna används för att analysera hur trafiken ser ut i något större skala och hur olika delar av det gällande området påverkar varandra. Enligt Caas, Perarnau och Torday (2011) går det i en mesoskopisk modell se hur ruttval och hastigheter förändras över på dagen och om det till exempel är trängsel och köer på den ”snabbaste” sträckan vid olika tidpunkter på dagen. Mesoskopiska modeller kan även användas för att analysera hur olika åtgärder påverkar trafikanters val och beteende i ett större perspektiv än för mikroskopiska modeller (Söder, 2020). Exempelvis hur kapacitetshöjande åtgärder kommer att påverka resmönstret hos trafikanterna, om ruttval förändras och om hastigheter och restider påverkas.

2.2.4 Dynamisk trafikmodellering

Modeller på mikro- och mesoskopisk nivå är två olika typer av så kallade dynamiska trafikmodeller, på engelska kallat ”Dynamic traffic assignment”, förkortat DTA. Dynamiska trafikmodeller används på en detaljerad nivå där enskilda fordon (mikro) eller grupper av fordon (meso) går att följa. För dessa typer av modeller går det enligt Chiu et al. (2011) att identifiera hur olika omständigheter påverkar val och beteende hos trafikanterna. I en dynamisk modell går det till exempel vid trafikplanering se hur resandet skiljer sig över dygnets timmar, identifiera flaskhalsar i systemet, hur hastigheter och restider beror på tiden på dygnet och hur trafikanter förhåller sig till varandra. Eftersom det går att analysera hur trängsel, köer och flaskhalsar beror på tidpunkt har dynamiska trafikmodeller även enligt Szeto och Wong (2011) potentialen att användas som underlag vid trafikstyrning. Sådana styrmedel kan exempelvis vara att kontrollera trafiksignaler och implementera olika nivåer av trängselskatt beroende på tidpunkt.

Enligt Szeto och Wong (2011) är grunden för dynamiska trafikmodeller att trafiken skiljer sig vid olika tidpunkter beroende på vilka förhållanden som råder i trafiken. Exempelvis vid lågtrafik kommer det förmodligen inte uppstå köer och trängsel i vägnätet, således brukar de sträckor som är kortast alternativt erbjuder högst nominell hastighet vara de snabbaste. Men köer och trängsel påverkar kapaciteten, vilket innebär att en rutt som vid lågtrafik är långsammare kan visa sig vara den snabbare rutten vid högtrafik då den, egentligen snabbaste rutten, då påverkas av trängsel. Detta fenomen exemplifieras i Figur 4 för resor mellan punkterna A och B. I detta fall har rutt 1 en lägsta kapacitet på 3 000 fordon per timme vid en flaskhals. Rutt 2 har däremot som lägst en kapacitet på 6 000 fordon per timme. Rutt 1 är kortare än rutt 2 och i detta fall är rutten vid fritt flöde den snabbaste, och förmodligen den rutt flest trafikanter väljer. Däremot när belastningen mellan punkt A och B överstiger 3 000 fordon per timme kommer det att uppstå trängsel och köer på rutt 1 och därmed längre restider. När restiderna för rutt 1 blir längre kommer fler trafikanter välja att färdas via rutt 2.



Figur 4. Exempel på rutter mellan två punkter A och B

Hur fördelningen sker mellan olika rutter efter det att trafiken överstiger den snabbaste ruttens kapacitet är svårt att förutspå. Ju större ett nätverk är, desto mer komplext blir det att beräkna och förutspå fördelningen eftersom antalet rutter och länkar som trafikanterna kan välja ökar (Szeto & Wong, 2011).

Enligt Szeto och Wong (2011) baseras trafikanters val i dynamiska trafikmodeller på Wardrops första och andra princip för statiska trafikmodeller. Wardrops första princip, vilket förklaras av Ekvation 4–6 och som kallas principen för användarjämvikt (*user equilibrium principle* på engelska), innebär att vid jämvikt kommer samtliga resor mellan ett visst OD-par att ha lika lång restid oberoende rutt. Wardrops andra princip innefattar att restiderna vid jämvikt motsvarar den kortast möjliga restiden. Det vill säga att så länge kapaciteten för den snabbaste ruten ej överskrids kommer samtliga trafikanter välja denna, men när kapaciteten överskrids kommer även andra rutter att väljas. För Wardrops första princip antas det att samtliga trafikanter är medvetna om vilken rutt de ska ta för att behålla jämvikt i trafiken och att alla är identiska i sitt resonemang. Istället för restid går det även att använda sig av den totala kostnaden för resan som utgångspunkt för jämvikten, det vill säga att bränslekostnader, parkeringsavgifter, trängselskatt etcetera räknas in. För att förhållandena vid jämvikt ska kunna förklaras med dessa principer krävs det att följande tre förutsättningar uppfylls:

$$f_p^{rs} \geq 0, \forall p, rs, \quad (\text{Ekvation 4})$$

$$n_p^{rs} - u^{rs} \geq 0, \forall p, rs, \quad (\text{Ekvation 5})$$

$$f_p^{rs} (n_p^{rs} - u^{rs}) = 0, \forall p, rs \quad (\text{Ekvation 6})$$

I dessa ekvationer står rs för OD-paret med start i zonen r och slut i zonen s , p är en slumpmässig rutt innefattande i OD-paret rs . f_p^{rs} för flödet på rutt p för OD-paret rs , n_p^{rs} är restid, alternativt total kostnad för rutt p för OD-paret rs och u^{rs} är minimal restid, alternativt minimal total kostnad för rutt p för OD-paret rs .

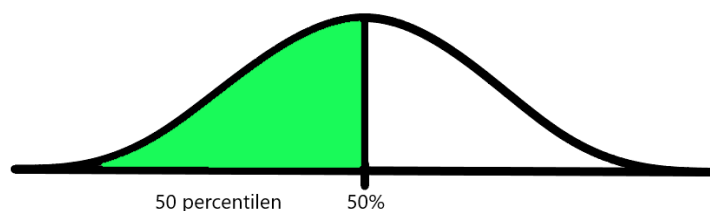
På grund av att Wardrops första princip grundar sig i att samtliga trafikanter för samma resonemang och är fullt medvetna om hur deras val påverkar restiden adderas det enligt Szeto och Wong (2011) i regel två ytterligare dimensioner som tar hänsyn till slumpmässiga

händelser. Dessa dimensioner innebär att jämvikten tar hänsyn till trafikanters uppfattning av restid och kostnader samt att det finns osäkerheter vid bestämning av restider. Eftersom olika trafikanter uppfattar en viss rutt olika kommer den upplevda kostnaden att differentiera beroende på vilken trafikant det gäller. När det är uppfattad istället för faktisk kostnad eller restid som påverkar ruttval, går det att anta att trafikanterna strävar efter stokastisk användarjämvikt (*Stochastic user equilibrium på engelska*). Stokastisk användarjämvikt och deterministisk användarjämvikt följer samma princip, men stokastisk användarjämvikt uppnås när trafikanterna inte längre kan minska sin *upplevda* restid eller kostnad genom att göra ett ruttbyte, till skillnad från deterministisk användarjämvikt som tar hänsyn till *faktisk* restid eller kostnad och som utgår från att samtliga trafikanter agerar likformigt.

Osäkerheter som kan påverka restider är enligt Szeto och Wong (2011) säkerhetsmarginaler, pålitlighet hos restiderna, restidspercentiler, restid vid sämsta möjliga förhållanden och trafikanters förväntade värde på resan. Anledningen till att säkerhetsmarginaler bidrar till mer osäkra restider grundar sig i att olika trafikanter är olika villiga att riskera sen ankomst på grund av längre restid än förväntat. En trafikant kan därmed välja att resa tidigare än nödvändigt för att komma till sin destination i tid. När hänsyn tas till säkerhetsmarginaler vid jämvikt kommer restiderna att vara summan av den faktiska restiden, samt den extra tid som trafikanten väljer att lägga för att inte komma försent.

För att ta hänsyn till pålitligheten i att restiderna är konstanta kan ett annat typ av mått än faktiska restider användas. Exempelvis kan viktade summor för de olika rutterna skapas. Summorna består av två typer av index, ett restidsindex samt ett pålitlighetsindex. Summan av dessa två resulterar i ett normaliserat värde för de olika ruttvalen och ju mindre värdet är, desto mer benägna är trafikanterna att riskera försenad ankomst i utbyte mot kortare restid (Szeto & Wong, 2011).

En metod som kan ersätta användning av medelrestid är att dela upp restiderna i percentiler beroende på hur villiga trafikanterna är att riskera förseningar på grund av för långa restider. Percentilerna visar i detta fall således på de restider som inte är långsammare än en viss procentandel. Om trafikanterna kan leva med risken att komma försent i utbyte mot att ha möjlighet att uppnå minimal restid är det den 50e percentilen som är gränsen. Men om trafikanterna är mer försiktiga med att bli försenade blir percentilen större, det vill säga att fler och långsammare restider används i jämviktsförhållandet. Vid jämvikt har samtliga rutter för ett OD-par samma restidspercentil (Szeto & Wong, 2011). Förklarande illustration för den 50e percentilen redovisas i Figur 5.



Figur 5. Illustrering av den 50e percentilen. Det grönmärkade området markerar värden innefattande i den 50e percentilen.

Vid uppskattning av percentiljämvikt är det enligt Szeto och Wong (2011) viktigt att även ta hänsyn till restider vid sämsta möjliga förhållanden. I detta fall anser författarna att det är

sannolikt att en resenär vid val av rutt kommer ta hänsyn till vad restiden för de olika rutterna är då det råder sämsta möjliga förhållanden. Vid dessa förhållanden antas maximal restid att råda. Jämvikten är beroende på hur mycket längre restiderna förväntas bli samt hur benägna trafikanterna är att riskera att hamna i en sådan situation i utbyte mot att resa på en rutt som annars är snabb.

Det förväntade värdet för en resa delas enligt Szeto och Wong (2011) i regel in i tre principer av beteende hos trafikanterna, men som hör ihop med varandra. Det första beteendet som tas till hänsyn är att trafikanterna kommer att väga in vad de tjänar respektive förlorar för respektive alternativ. Detta hör i sin tur ihop med det andra beteendet, vilket är att människor tenderar att vara mer vaksamma för risker och förluster än vad de anser sig få ut av vinster. Detta beteende går att koppla till den tredje principen vilket lyder att människor generellt anser att sannolikheten att det sker mycket ovanliga händelser, såsom stora olyckor eller ovanligt mycket trängsel, är större än vad den egentligen är.

2.3 Trafikverkets modell för Region Stockholm i INRO Dynameq

Trafikmodellen som används i fallstudien är en dynamisk trafikmodell över Region Stockholm, framtagen av Trafikverket, version 0.8c. Modellen är framställd genom det dynamiska modelleringsverktyget INRO Dynameq (Trafikverket 2021b), vilket är ett modelleringsverktyg som används för att skapa trafikmodeller på meso- och mikronivå. Modeller skapta i Dynameq kan användas för att analysera köer och trängsel, egenskaper såsom restid, hastigheter och ruttval, påverkan på trafikanter av styrmedel såsom signaler och avgifter samt hur olika infrastrukturlösningar påverkar biltrafiken. Dynameq kan även användas för att analysera dessa parametrar för en framtida och förändrad trafiksituation (INRO, 2022). Trafikmodellen som framställts i Dynameq och som användes i fallstudien var en trafikmodell på mesonivå, se delkapitlet om Olika nivåer och typer av trafikmodeller.

Efterfrågan hos trafikanterna som representeras i trafikmodellen tas från den nationella prognosmodellen Sampers, version 3.4.4, vilket enligt Trafikverket (2022a) är den senaste versionen av Sampers. Sampers är ett modellsystem ägt av Trafikverket, som bland annat används för att beräkna trafikvolymerna. Volymerna beror på olika parametrar som även kan varieras i syfte att kunna analysera olika scenarier. Grundparametrar i modellen är bland annat BNP, befolkningstillväxt, sysselsättning och bränslepris (Trafikverket, 2022a). Sampers-modellen är en fyrstegsmodell där efterfrågan och trafikflöden beräknas integrerat. Sampers-modellen är uppdelad i zoner som till exempel motsvarar olika områden inom städer (Trafikverket, 2020). Zonerna innehåller information om bland annat boende, arbetande och socioekonomiska situationer, vilket motsvarar det första steget i en fyrstegsmodell. Som indata till Dynameq-modellen är det trafikflöden vid förmidagens högtrafiktimmar, vilket är mellan 06.00 och 10.00 på vardagar, som används. Således är det under dessa timmar på en slumpmässigt vald vardag som är det scenario som modellen ska efterlikna.

En del av framställandet av en dynamisk trafikmodell är arbetet med att få trafikmiljöerna att stämma överens med verkligheten. Trafikmiljöer som för den aktuella trafikmodellen ansetts viktiga för modellens funktion är hastighets- och flödessamband, signalregleringar, cirkulationsplatser, påfarter och växlingssträckor. Samtliga av dessa aspekter är viktiga för

modellens verklighetsförankring och många hör ihop med varandra där främst hastighets- och flödessamband påverkas av varje aspekt. Korrekt kodad signalreglering i modellen medför att trafikanterna beter sig på liknande sätt i modellen som i verkligheten när de når en signalreglerad korsning. Om en signal exempelvis är grön för länge kommer både hastighet och flöde bli för stort. Cirkulationsplatser, påfarter och växlingssträckor liknar varandra ur den aspekten att det som påverkar hastighet och flöden vid dessa är hur trafikanter förhåller sig sinsemellan och väjer för varandra vid exempelvis väjningsplikter och filbyten.

3 Metodik och metoder

3.1 Metodik

Metoden för insamling och analys av data har bestått av två olika arbetsmoment. Det första momentet har varit att genomföra en enkätstudie där erfarenheter och kunskap från sakkunniga personer som arbetar med trafikplanering deltagit. Det andra momentet har varit en fallstudie där en befintlig trafikmodell för Region Stockholms i det dynamiska trafikmodelleringsverktyget INRO Dynameq har genomgått en validering med hjälp av GPS-baserade probedata.

3.2 Metod för enkätstudie

Den ena metoden för datainsamling har skett genom en enkätstudie. Syftet med enkätstudien var tudelat. Dels var syftet med enkäten att kunna skapa en kunskapsbas där erfarenheter och kunskap om användning av GPS-baserad probedata i branschen sammanställs för att förenkla framtida arbete och utveckling inom området. Denna kunskapsbas var ämnad för användning inom Rambolls organisation. Genom att samla erfarenheter och metoder i en sådan kunskapsbas var förhoppningen att kunna förkorta tiden det tar för att finna nyttig information. Formuläret skulle även visa på vilka personer som har vilka kunskaper, vilket förkortar kommunikationsvägarna. Syftet med enkäten var dels att den skulle kunna användas som både ett underlag och ett komplement till den efterföljande fallstudien för Region Stockholm i INRO Dynameq, som fokuserar på att validera trafikmodeller genom att använda GPS-baserade probedata. Enkäten bestod av fyra delar, varav den ena behandlade validering av trafikmodeller med GPS-baserad probedata. Enligt avgränsningar är det främst detta syfte som är relevant för examensarbetet, och således är det resultaten från detta avsnitt som har analyserats och redovisats i denna rapport. Förhoppningen var att genom enkätens svar kunna finna intressanta parametrar att undersöka i fallstudien. Förhoppningen var att enkäten skulle ge en ökad förståelse om validering av trafikmodeller med hjälp av GPS-baserade probedata utöver den kunskap som fallstudien belyste. Enkätstudien lades upp semistrukturerat. En semistrukturerad enkät innebär att enkäten generellt består av fördefinierade frågor, och i detta fall med förbestämda svar i form av flersvarsfrågor. Däremot ges det i flera fall möjligheten till att utveckla vidare sina svar, vilket ingår i den semistrukturerade uppbyggnaden (Johansson, 2021). Enkäten analyserades genom sina flersvarsfrågor delvis kvantitativt då svaren på dessa frågor kunde räknas på och sedan jämföras statistiskt.

3.2.1 Utformning av enkät

Enkäten har utformats digitalt via programmet Microsoft Forms och distribuerades till deltagarna via e-mail. Eftersom deltagarna arbetar i olika delar av Norden var språket i

enkäten engelska. Enkäten bestod av fyra olika huvuddelar med totalt 23 frågor, varav nio har givit den tillfrågade olika svarsalternativ och resterande 14 möjlighet att svara med fri text. Frågorna i de tre första delarna har framtagits för att kunna användas i syftet att skapa en sammanställning över erfarenheter, metoder och kunskaper som fanns kopplat till projekt, data, metoder och verktyg vid arbete med GPS-baserade probedata. Men eftersom dessa tre delar är utanför studiens avgränsningar bortses dessa ifrån vid analys av resultat. Den fjärde delen har fokuserat på användning av GPS-baserade probedata vid validering av trafikmodeller. Enkätens fyra huvuddelar har kategoriserats enligt följande:

1. Användning av GPS-baserade probedata i trafikplaneringsprojekt
2. Information om använd data
3. Metoder och verktyg för att hantera data
4. Validering av trafikmodeller med hjälp av GPS-baserade probedata

Den sista huvuddelen har utformats för att täcka de tillfrågades erfarenheter kring hur väl de ansett det fungera att använda sig av GPS-baserade probedata för att validera trafikmodeller. Frågorna har formulerats för att ge svar på vad som fungerat bra, vad som inte fungerat bra samt att det har getts utrymme för att fundera kring potentiellt nya tillämpningsområden. Denna del bestod av sju frågor, varav tre har givit deltagarna olika svarsalternativ och resterande fyra frågor krävde svar med fri text.

3.2.2 Frågor

Enligt avgränsningarna för studien har det enbart varit den fjärde och sista delen av enkäten som hanterats. Således är det enbart frågorna tillhörande den delen som redovisas. Eftersom enkätens språk har varit engelska redovisas frågorna således på engelska. Denna del har bestått av följande frågor:

19. Have you been using GPS probe data for the purpose of validating traffic models?
20. For which parameters do you think that GPS probe data has been working successfully when validating traffic models?
 - Origin and destination (OD),
 - vehicle speed,
 - travel time,
 - entire trip trajectory,
 - queue lengths/delays,
 - traffic volumes,
 - driving behavior

- vehicle kilometers travelled,
- direction of movements,
- accidents,
- others.

21. Possibility to elaborate answers on question 20.

22. For which parameters has GPS probe data been difficult to use, or not given the desired results, when validating traffic models?

- Origin and destination (OD),
- Vehicle speed,
- Travel time,
- Entire trip trajectory
- Queue lengths/delays
- Traffic volumes,
- Driving behavior
- Vehicle kilometers travelled,
- Direction of movements,
- Accidents,
- Others

23. Possibility to elaborate answers on question 22.

24. What's the reason why you haven't worked with GPS probe data for the purpose of validating traffic models?

25. Do you believe that there are other ways to implement usage of GPS probe data in the validation process, which hasn't been tried yet? If so, please elaborate.

Fråga 19, 20 och 22 har flersvarsalternativ, och resterande frågor är utformade för att besvaras med fri text. I enkäten har det programmerats förgreningar, vilket innebär att beroende på hur en deltagare svarar på vissa frågor förändras frågeföljden. I detta fall har det skapats en förgrening efter fråga 19. Om den tillfrågade svarar ja kommer fråga 20 att följa och fråga 24 ej besvaras. Om personen däremot svarar nej på fråga 19 följer fråga 24 och fråga 20 till och med 23 besvaras ej.

Syftet med frågorna har således varit att undersöka vilka parametrar som det finns erfarenhet av i syftet att validera trafikmodeller med hjälp av GPS-baserade probedata. De tillfrågade har fått möjligheten att beskriva vilka parametrar som har varit av nytta att undersöka vid valideringen av trafikmodeller samt vilka parametrar som fungerat dåligt, eller sämre än önskat. Det har även givits möjlighet för deltagarna att ge sin syn på vilka ytterligare potentiella tillämpningar innefattande GPS-baserade probedata det kan finnas vid validering av trafikmodeller.

3.2.3 Tillfrågade personer

Personerna som har tillfrågats i enkätstudien är anställda på Ramboll och arbetar med trafikrelaterade frågor, majoriteten av dem på enheten Smart Mobility. Deltagarna är sakkunniga inom trafikplanering, vilket de arbetar med i sin vardag. Deltagarna är placerade i länderna Sverige, Norge, Danmark och Finland och samtliga har olika erfarenheter kring användning av GPS-baserade probedata vid trafikplanering. Eftersom enkätens syfte i stort var att bidra med en kunskapsbas för användning inom Rambolls organisation, tillfrågades således endast personer anställda på Ramboll. De tillfrågade personerna uppgav sina namn vid svar på enkäten, men redovisas anonymt i denna rapport.

3.3 Metod för fallstudie Region Stockholm

Under fallstudien användes kvantitativ analysmetod vid jämförelsen mellan modell och probedata. En kvantitativ metod innebär enligt Nationalencyklopedin (2022) att analysen baseras på mätbara och empiriska värden som kan sammanställas som olika sorters statistik. Denna metod ansågs vara väl lämpad för detta arbete eftersom den inköpta probedatan från INRIX som samlades in från verkliga fordon i detta fall analyserades som statistik över hur Region Stockholms biltrafikanter reser. Även utdata från trafikmodellen i Dynameq som analyserades var av statistisk karaktär, vilket möjliggjorde för en kvantitativ analys även av denna.

3.3.1 Arbetsgång i INRO Dynameq

För att extrahera data från den dynamiska trafikmodellen i INRO Dynameq krävdes det att simuleringar av modellen kördes. Simuleringarna i modellen motsvarade trafiksituationen mellan klockslagen 06.00 och 10.00 i Region Stockholm. Resultaten i Dynameq kunde både anges visuellt för varje länk, och extraheras i tabellform för egen datahantering och analysering i Microsoft Excel. Modellen krävde inget arbete med modifiering av kodning innan körning av hela modellen.

3.3.2 Hantering av data

Den inköpta probedatan från INRIX och tillhandahållna av Ramboll bestod av totalt 599 288 enskilda resor, registrerade mellan 2019-09-16 och 2019-10-21. Se delkapitel 2.1 GPS-baserade probedata för beskrivning av det använda datasetet. Under dessa resor registrerades det totalt 24 579 876 waypoints. På grund av den insamlade datans storlek krävdes det omfattande arbete med datahantering för att möjliggöra analysering. I detta fall

användes programmeringsverktyget RStudio version 2022.02.0. Programkoderna som framställts i denna studie redovisas under Bilagor. I bilagorna redovisas koden för hur hastigheter och restider togs fram för en plats respektive en sträcka. Dessa koder upprepades men med olika indata som motsvarade vardera analyserad plats och sträcka. Denna upprepning redovisas inte i rapporten.

3.3.3 Validerande parametrar

Vid validering av en trafikmodell finns det flertalet olika parametrar som kan analyseras. För detta arbete har det fokuserats på de tre parametrarna hastigheter, restider och ruttval där samtliga krävde olika arbetsprocesser, vilka redovisas nedan.

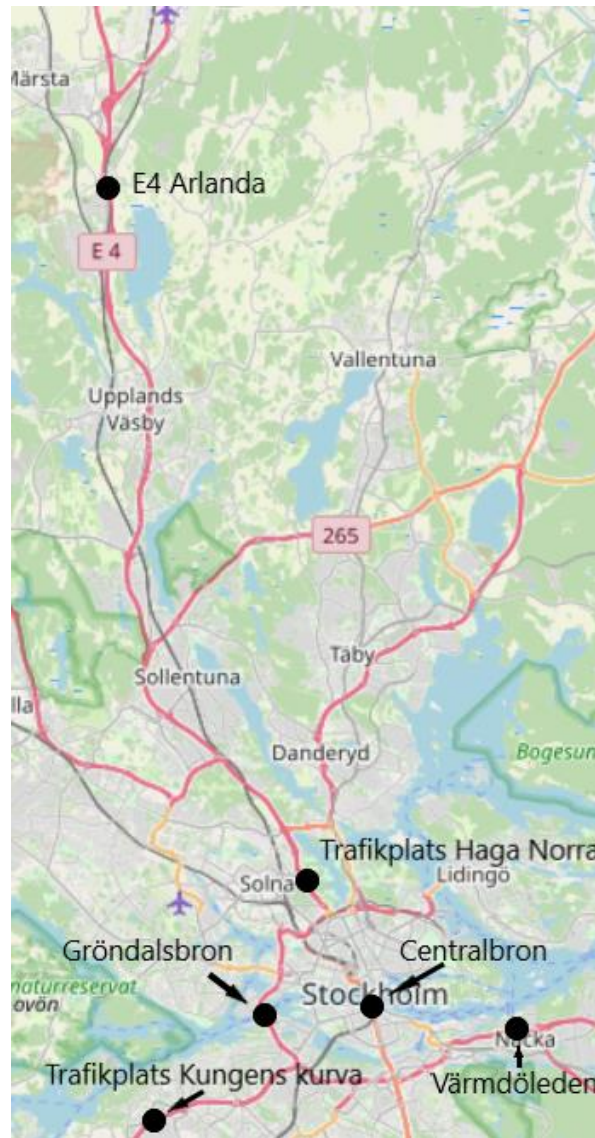
3.3.3.1 Hastigheter

För validering med hjälp av hastigheter valdes sex olika platser i analysen. I både probedatan från INRIX (trajectories) samt för trafikmodellen i Dynameq är vägnätet uppdelat i olika segment (INRIX) och länkar (Dynameq), dessa beskriver samma sak men benämns olika för de två dataseten. Vid genomförandet av valideringen valdes således segment i probedatan och länkar i Dynameq som motsvarade varandra för att göra jämförelsen pålitlig. Platserna samt deras egenskaper i form av namn, vägnummer, SegmentID (INRIX), länknummer (Dynameq) och färdriktning redovisas i Tabell 1. Samtliga av de utvalda platserna är vägar som belastas mycket under rusningstrafik, vilket antogs ge resultat med representativa värden då dessa länkar förmodades påverkas mindre av en lägre penetrationsgrad. Jämförelsen görs för samtliga körfält.

Tabell 1. Sammanställning av platserna där hastighet analyserades

Plats	Väg/gata	SegmentID	Länknummer
Gröndalsbron/Essingeleden	E4/E20	45049310 (N)	3369/16801 (N)
		205889185 (S)	1759/16802 (S)
Trafikplats Haga Norra	E4	57362330 (N)	5509 (N)
		46101899 (S)	5518 (S)
E4 söder om Arlanda flygplats	E4	4197533 (N)	16928 (N)
		4197532 (S)	16929 (S)
Centralbron	Centralbron	261818767 (N)	783 (N)
		5788351 (S)	1823 (S)
Trafikplats Kungens kurva	E4/E20	205882923 (Ö)	16795 (Ö)
		4215402 (V)	16796 (V)
Värmdöleden	222	35971216 (Ö)	153 (Ö)
		115682018 (V)	166 (V)

Var dessa platser är lokaliserade visas i Figur 6.



Figur 6. Platser för analys av hastigheter. © OpenStreetMaps bidragsgivare. Tillgänglig under licensen Open Database license

Vid hantering och analysering av hastigheter från probedatan var det databeskrivningen trajectories som användes. I trajectories fanns det information om varje enskild resas egenskaper för varje passerat segment där en eller flera waypoints registrerats. Egenskaper som redovisades var bland annat segmentets ID, koordinater, hastighet utmed segmentet samt klockslag för passage in och ut. Utifrån dessa egenskaper gick det i RStudio att applicera önskade filter för att endast få ut information som ansågs nödvändig för analysen.

Det genomfördes enskilda analyser beroende på färdriktning för samtliga av de utvalda platserna. Eftersom Dynameq-modellen utgår från högtrafik under vardagsförmiddagen bestämdes det således att alla resor som registrerades på helgdatum skulle filtreras bort. Eftersom det kan uppstå fel vid GPS-mottagning filtrerades samtliga mätningar med felkoder, såsom uteblivna mätdata, bort samt mätningar där värdet på hastigheten ansågs vara orimliga eller otroliga, i detta fall negativa hastigheter och hastigheter över 200 km/h. När hastigheten för respektive valt segment analyserades filtrerades övriga segment bort för att endast de gällande och relevanta värdena skulle kunna analyseras.

Efter att nämnda filter hade applicerats var de kvarvarande värdena endast från mätningar genomförda på vardagar under hela tidsperioden för respektive analyserat segment. Eftersom medelvärdet på penetrationsgraden för datan endast var ungefär 1–1,5% beräknades hastigheten beroende på tidpunkt utifrån samtliga vardagar under insamlingsperioden. Således minskade risken att det registrerats för få resor en slumpmässig dag, vilket kunde ge missvisande värden. Hastigheterna sorterades i RStudio på ett sådant sätt att datumet för passagen inte var relevant, utan att det var tidpunkten på dygnet som spelar roll. Sedan aggregerades de uppmätta hastigheterna i form av medelhastigheter i intervall per timme över ett helt dygn, vilket motsvarade tidsintervallet i Dynameq-modellen.

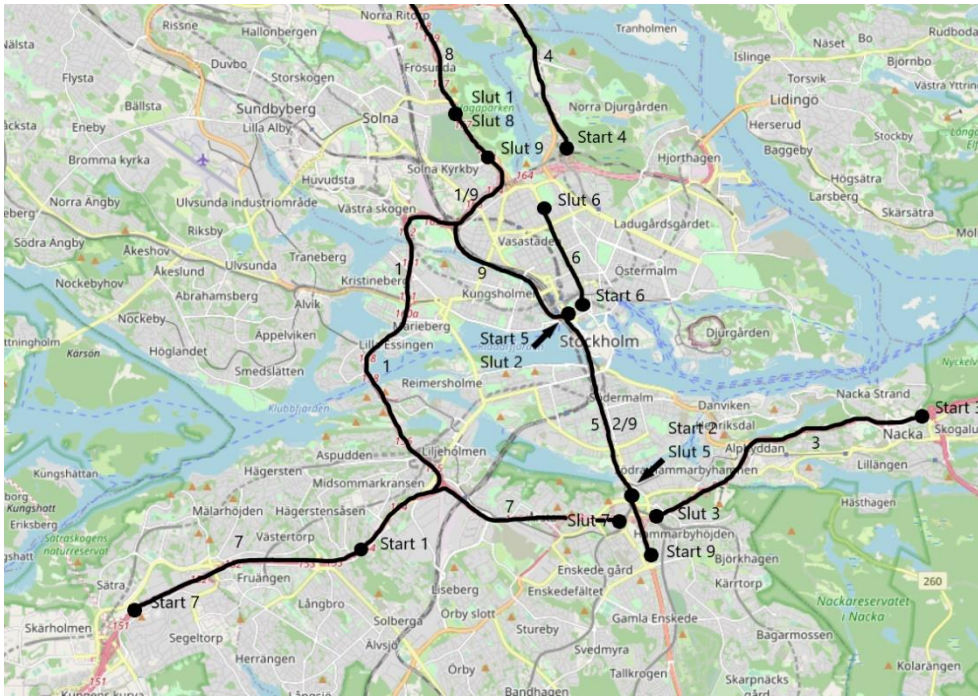
Utifrån resultat och simulering i Dynameq-modellen gick det att få fram data över hur hastigheten förändras beroende på tid på förmiddagen. Hastigheternas värde som kunde analyseras var medelhastigheten för varje enskild länk under varje hel timme vid morgonens högtrafik. Timmarna som var möjliga att analysera från Dynameq-modellen var på en slumpmässig vardag mellan klockan 06.00 och 10.00. Eftersom det endast var under dessa timmar som Dynameq-modellen kunde analyseras begränsades därmed möjligheterna att använda probedatan för övriga timmar på dygnet som en valideringskälla.

3.3.3.2 Restider

I fallstudien genomfördes arbete med att jämföra restider mellan ett antal start- och slutpunkter. För studien valdes nio olika sträckor för studien. Dessa nio var av olika karaktär, fyra sträckor hade riktning in mot centrala Stockholm under morgonens rusningstrafik, två hade riktning bort från centrala Stockholm, en var ledd kring innerstaden och två rutter genom innerstaden. Genom att välja sträckor med varierande karaktär möjliggjordes analys av delar av modellen som påverkas på olika sätt. För de sträckor med riktning in mot Stockholms innerstad gavs indikationer för hur väl anpassad modellen är för timmar med högtrafik och dess medförande trängsel och köer. De sträckor med riktning bort från innerstaden gav således antydningar på hur modellen fungerar för resor utmed lägre belastade länkar relativt till de med riktning mot centrala Stockholm, vilket antas gälla för morgontrafiken ut från den centrala delen av Stockholm. Sträckan som kodades för att gå genom centrala Stockholm förväntades ge svar på hur väl modellen är anpassad till resor som dels påverkas av trängsel och köer, dels trafiksignaler. Vilka sträckor som analyserades redovisas i Tabell 2 och illustreras i Figur 7 och Figur 8.

Tabell 2. Redovisning av sträckor där restid uppmätts.

Numrering	Sträcka
1	Hägersten-Trafikplats Haga Norra
2	Johanneshovsbron-Centralbron
3	Nacka-Johanneshov
4	Norrtull-Danderyd
5	Centralbron-Johanneshov
6	Klaratunneln-Sveavägen
7	Kungens kurva-Johanneshov
8	Sollentuna- Trafikplats Haga Norra
9	Genom innerstaden



Figur 7. Start- och slutlänkar samt rutter i Dynameq vid restidsanalysering. Startpunkter, slutpunkter och rutter är markerade och numrerade enligt Tabell 2. © OpenStreetMaps bidragsgivare. Tillgänglig under licensen Open Database license



Figur 8. Start- och slutlänkar samt rutter i Dynameq vid restidsanalysering. Startpunkter, slutpunkter och rutter är markerade och numrerade enligt Tabell 2. © OpenStreetMaps bidragsgivare. Tillgänglig under licensen Open Database license

För att kunna analysera restider för specifika sträckor i Dynameq krävdes det visst arbete med modellen. Innan simulering skapades det rutter med sammanhängande länkar. Dessa rutter hade samma start- och slutlänkar som probedatan för respektive analyserad sträcka.

Rutterna kodades i Dynameq som busslinjer med avgångar med tidsintervallet tio minuter och som mellan start- och slutlänkarna inte hade några stopp. För att fordonen skulle agera som personbilar istället för bussar krävdes viss omkodning. De parametrar som förändrades var fordonets effektiva längd, förarens reaktionstid samt till vilken procent av vägens hastighetsbegränsning fordonet färdas. Fordonens effektiva längs angavs till 6,25 meter, vilket motsvarar en personbil. Ett fordons effektiva längd i Dynameq omfattar avståndet från gällande fordons bakre stötfångare till bakomvarande bils bakre stötfångare. Reaktionstiden för trafikanten att hinna reagera och ändra riktning bestämdes till 1,25 sekunder. Det angavs även att fordonen färdas i 100% av den högsta tillåtna hastigheten, det vill säga som en personbil som till skillnad från bussar och lastbilar inte har ytterligare hastighetsbegränsningar. Eftersom rutterna kodades som busslinjer genomfördes således samtliga analyserade resor utmed samma länkar och därmed varierade inte ruttvalet mellan de simulerade trafikanterna. Efter det att samtliga rutter kodats in kunde simuleringen köras och efter det kunde resultaten extraheras och analyseras. Likt resultaten från probedatan gick det att ta fram kortast restid, längst restid och medelrestid för samtliga rutter.

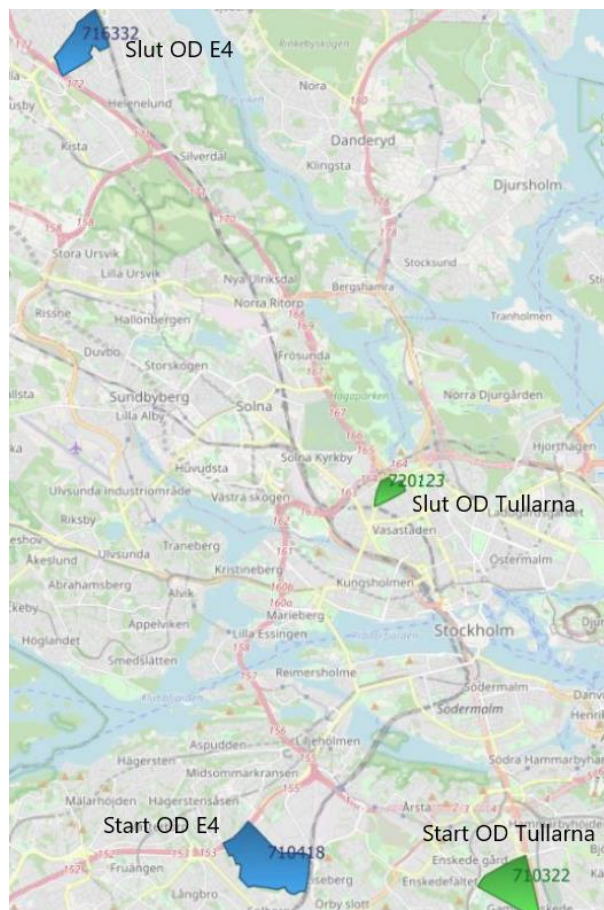
För analys av restiderna registrerade i probedatan från INRIX användes information tillhandahållen ur Trajectories. I Trajectories redovisas det för varje passerat segment bland annat klockslag för passering av början och slutet av segmentet. Dessa klockslag användes för att beräkna restid mellan det första och sista segmentet för respektive sträcka. För att möjliggöra sådana beräkningar filtrerades datan så att endast de resor där start och slutsegmentet hos respektive sträcka ingick. Genom att subtrahera klockslaget för passage ut från slutsegmentet med passage in på startsegmentet kunde restiden för varje enskild resa beräknas. Från denna subtraktion var det även möjligt att skilja resor med olika riktningar åt och således möjliggjordes det att filtrera bort resor som för analysen ansågs vara i fel riktning. Eftersom kravet endast var att start- och slutsegmentet ingick hos de analyserade sträckorna kan ruttvalen hos de olika fordonen variera. Detta utgör således en skillnad mellan den GPS-baserade probedatan och Dynameq-modellen då endast specifika rutter analyserades i Dynameq.

Eftersom Dynameq-modellen enbart simulerades under tidsintervallet 06.00-10.00 valdes endast de resor som i probedatan registrerats under samma tidsintervall. Eftersom det finns felkällor vid registrering av GPS-data såsom att mottagarens klocka går fel och att lokaliseringen inte stämmer, kan det resultera i att restider för enskilda resor mellan två segment kan anses vara orimligt korta eller långa. Således bestämdes det att endast restider inom 95-percentilen, men ej inkluderande 5-percentilen, ingick i analysen för att få bort dessa resor. Denna tillämpning gjordes för att få så många av de registrerade resorna som möjligt, men samtidigt kunna bortse från extremfallen. Utifrån det filtrerade datasetet gick det att ta ut snabbast registrerade resa, långsammast registrerade resa samt beräkna medelrestiden för respektive sträcka.

3.3.3.3 Ruttval

Ahlberg et al. (2021) genomförde i *Probedata, förstudie kring användning av GPS-baserade probedata för skattning av hastigheter, länkflöden och ruttval* arbete med att analysera vilka ruttval som trafikanterna gjorde mellan två utvalda OD-par. De resultat gällande ruttval som föreliggande studie påvisade enligt probedatan var de som även användes i fallstudien för Region Stockholm. Dessa resultat jämfördes mot motsvarande resultat erhållna ur Dynameq-modellen.

I Figur 9 går det att se vilka OD-par som analyserades. De analyserade OD-paren bestod båda av en startzon lokaliserad i den södra delen av Stockholm och med en slutzon lokaliserad längre norrut. Ur probedatan har samtliga resor som genomförts mellan dessa zoner undersökts för vilka rutter som trafikanterna använt sig av. De analyserade resorna var inte begränsade till endast resor som hade start och slut i dessa zoner. Även de resor med start och slut i andra zoner men som passerade de utvalda områdena under sin färd var med i utvärderingen (Ahlberg et al., 2021).



Figur 9. Analyserade OD-par för ruttval (Ahlberg et al., 2021). OD E4 markeras med blå färg och OD Tullarna med grön färg.

I Figur 9 illustreras de undersökta OD-paren med två olika färger, vilket innebär att zonerna med samma färg tillhör samma OD-par. Ahlberg et al. (2021) benämner de två OD-paren som *OD E4* (blå färg) och *OD Tullarna* (grön färg). OD E4 har startzon vid Solbergaskogen vilket ligger i anslutning till Trafikplats Västberga vid väg E4/E20 och slutzon vid Trafikplats Tureberg vid väg E4 beläget mellan Sollentuna och Husby. OD Tullarna har startzon vid Enskede i södra Stockholm och slutzon vid norra Vasastaden vilket ligger i norra delen av Stockholms innerstad.

Ahlberg et al. (2021) studerade resor för olika tidpunkter på dygnet samt jämförde resultat för vardagar och helger. Totalt antal resor som analyserades var 2009 för OD E4 respektive 28 för OD Tullarna. Vid insamlingen av probedatan registrerades det däremot 2375 resor för OD E4 respektive 61 för OD tullarna. Anledningarna för att vissa resor ej analyserades var dels på grund av att en andel av resorna innehöll någon form av felkod och därmed inte var

brukbara. Dels filtrerades även resor med en restid som ansågs olämplig bort. Ahlberg et al. (2021) ansåg restiderna vara orimligt långa om de översteg medelrestiden för OD-paren med 50%. Eftersom Dynameq-modellen enbart kördes för förmiddagens högtrafik under vardagar var det därmed enbart motsvarande resultat som användes från probedatan. Således blev antalet resor som analyserades i detta arbete färre. Ahlberg et al. (2021) analyserade ruttvalen bland annat under tidsintervallet 06.00 – 09.00, i detta fall var det även detta intervall som analyserades i Dynameq.

Analys av ruttval hos en modell i Dynameq har här skett med hjälp av en liknande metod som Ahlberg et al. (2021) genomförde för att analysera ruttval enligt probedata. I Dynameq valdes två länkar ut, vilket motsvarade de två punkter mellan vilka de simulerade trafikanternas ruttval undersöktes. Utifrån dessa punkter gick det att få fram egenskaper för samtliga resor som passerade dessa två länkar under färden. Således gick det att ta fram vilka länkar som använts och därmed vilka val av ruttor som de simulerade trafikanterna gjort. Ur resultatet framgick det även hur många som gjorde de olika valen.

Eftersom det var enskilda länkar som valdes som start- och slutpunkter fanns risken att resultaten på grund av detta skiljde sig från resultaten från probedatan, vilket inkluderade samtliga resor mellan två zoner istället för två länkar. Därmed var det viktigt att testa olika länkar som start och slut inom de zoner som utgjorde de två OD-paren, för att se eventuella skillnader hos ruttvalen beroende på val av länkar. Länkarna som användes var dels de stora vägarna inom eller i anslutning till zonerna, såsom E4/E20 för OD E4 och väg 73 och Sankt Eriksgatan för OD Tullarna. Dels analyserades även mindre vägar inom zonerna och avfarter från de stora vägarna, såsom Enskedevägen och Norrtullsgatan för OD Tullarna och Älvsjövägen och avfart 172 från E4 mot Turebergsleden för OD E4.

3.3.4 Analys

Vid validering av trafikmodellen jämfördes de simulerade värdena från modellen med uppmätta motsvarande värden från en empirisk datakälla för samma tid och plats. I detta fall jämfördes resultat extraherade från Dynameq-modellen med den bearbetade probedatan.

3.3.4.1 Hastigheter

Utifrån den bearbetade probedatan samt resultaten från Dynameq-modellen kunde valideringsarbetet genomföras. Medelhastigheten beroende på tidpunkt sammanställdes i tabeller där både den aggregerade probedatan och resultaten från Dynameq-modellen utgjorde varsin kolumn. Med hjälp av dessa tabeller gick det att jämföra värdena från probedatan och Dynameq-modellen. Utefter dessa gick det att analysera huruvida modellen och probedatan gav tillräckligt liknande resultat för att kunna anses som fungerande. Utefter dessa tabeller gick det även att ta fram illustrerande grafer för hur medelhastigheten för samtliga analyserade platser varierade under förmiddagen beroende av tid för både probedatan och datan från Dynameq-modellen. För att modellen och probedatan skulle anses vara nog överensstämmande och därmed modellen som godkänd fick inte värdena skilja mer än 20%. Mindre än 10% skillnad ansågs vara bra. Dessa grafer användes för att illustrera och tydliggöra mönster och skillnader mellan de olika typerna av data. Dessa grafer och procentuella avvikelser redovisas i delkapitlet 3.2.1 Hastigheter.

3.3.4.2 *Restider*

Utifrån den bearbetade probedatan och resultaten från Dynameq-modellen jämfördes motsvarande värden på restider för de två dataseten. Jämförelsen mellan kortast restid, längst restid och medelrestid redovisades i form av stapeldiagram för respektive restidsgrupp. Dessa diagram gav möjligheten att tydligt se hur de olika metoderna förhöll sig till varandra enligt de tre restidsgrupperna för respektive sträcka. I tabellform redovisades de procentuella avvikelserna mellan metoderna. För att värdena skulle anses likvärdiga och därmed att modellen ansågs godkänd med hänsyn till den validerande källan fick skillnaden mellan resultaten från de två metoderna inte skilja sig mer än 20%. Mindre än 10% skillnad ansågs vara bra. Diagram och tabeller som grundade analyserna redovisas i delkapitel 4.2.2 Restider.

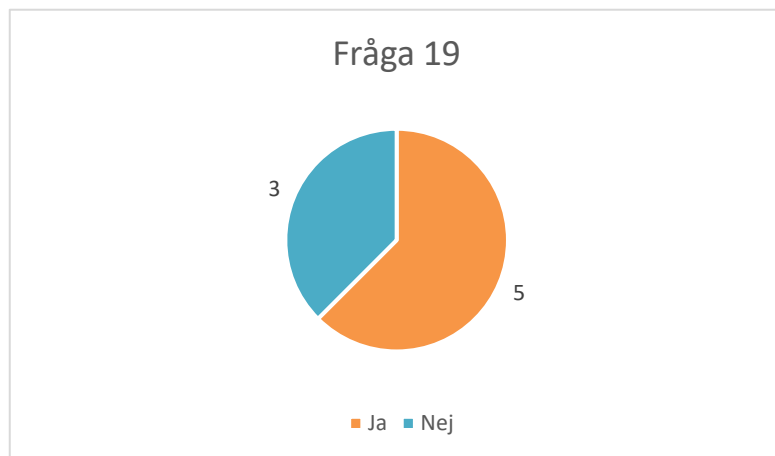
3.3.4.3 *Ruttval*

Resultaten från både Dynameq-modellen och probedatan redovisades i form av markeringar på karta. Utifrån dessa markeringar gick det att jämföra vilka ruttval som registrerats i probedatan med de simulerade trafikanternas ruttval. Eftersom modellen i Dynameq enbart tog hänsyn till rutter mellan två specifika länkar analyserades således fler fall för OD-paren för att få resultat som ansågs mer jämförbara med motsvarande från probedatan. Ruttvalen enligt de två metoderna redovisas i delkapitlet 3.2.3 Ruttval. För sex av de nio sträckor vars restider analyserades var de procentuella avvikelserna mellan medelrestider i Dynameq-modellen och medelrestiderna enligt probedatan tillräckligt små för att de kunde anses vara godkända. Av dessa sex var avvikelserna för fyra sträckor mindre än 10% vilket anses vara bra värden. Gällande resor med kortast restid var det tre av nio sträckor som ansågs ha bra procentuella avvikelser, och för resterande sträckor ansågs de ej vara godkända. För resorna med längsta registrerade restid ansågs tre av nio sträckor ha godkända eller bra avvikelser på mindre än 20% mellan de två metoderna. Ur resultaten går det således att tyda att avvikelserna för medelrestiderna var betydligt mindre jämfört med när restiderna var som kortast respektive längst.

4 Resultat

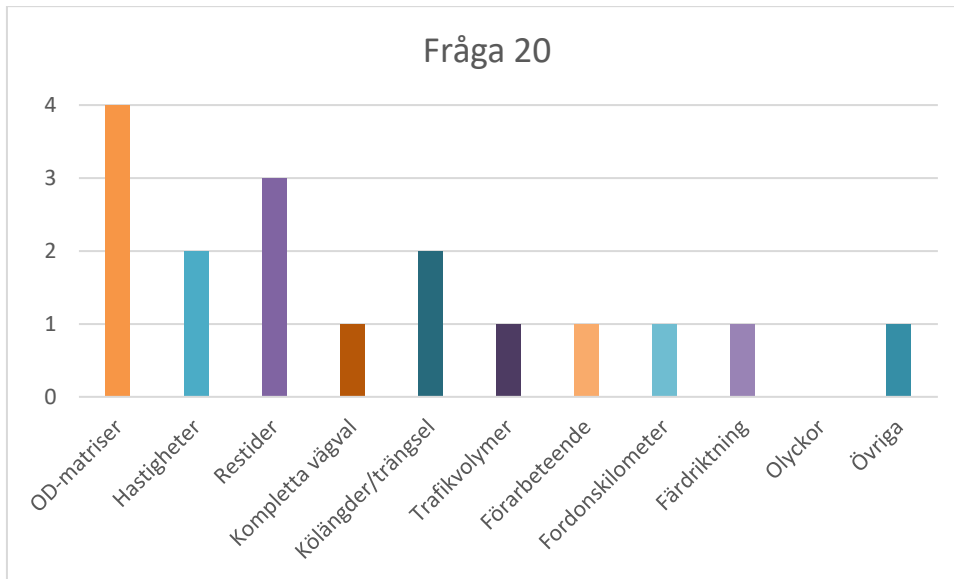
4.1 Enkätstudien

Totalt åtta personer deltog i enkätstudien. Samtliga åtta av dessa respondenter arbetade på Ramboll varav fem tillhörde enheten Smart Mobility, en på Road and Spatial Planning, en på Transport Systems och en på Samferdsel i Norge. I Figur 10 syns det att av åtta respondenter svarade fem att de har erfarenhet av att arbeta med GPS-data i syftet att validera trafikmodeller. De respondenter som svarade nej dirigerades direkt vidare till fråga 24 vilket var en följdfråga på varför de inte använt sig av GPS-baserade probedata för att validera trafikmodeller. Två av tre aktuella respondenter svarade på fråga 24, där en tillfrågad svarade *"Jag arbetar inte med validering av trafikmodeller"*. Den andra av de tillfrågade svarade *"Vid modellering behöver jag observationer från samtliga fordonstyper. GPS-data är inte lämpligt för det"*.



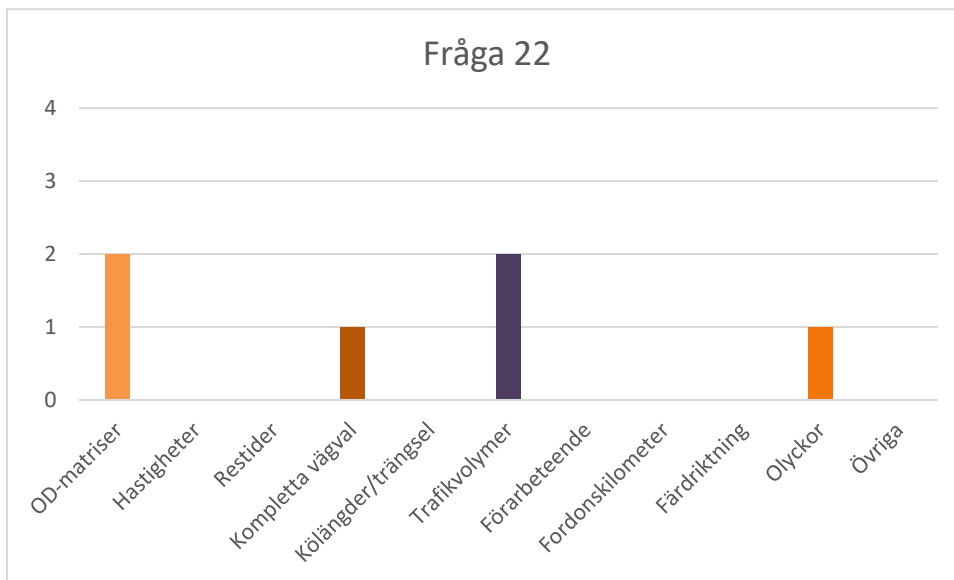
Figur 10. Svar på fråga 19: Have you been using GPS probe data for the purpose of validating traffic models?

I fråga 20, Figur 11, fick respondenterna möjlighet att beskriva vilka positiva erfarenheter de hade av ett urval av parametrar som kan användas vid validering av trafikmodeller. Fyra av de svarandena ansåg att information om efterfrågan mellan olika zoner (OD-matriser) var en framgångsrik parameter vid validering av trafikmodeller. Tre hade positiva erfarenheter av att jämföra restider. Hastigheter samt köer och trängsel hade två svar vardera. Resors kompletta vägval, trafikvolym, förarbeteende, totalt antal fordonskilometer och färdriktning har fått ett svar vardera. Som följdfråga fanns det genom fråga 21 möjligheten att utveckla sitt svar på fråga 20, ingen av de tillfrågade registrerade ett svar på den frågan.



Figur 11. Svar på fråga 20: For which parameters do think that GPS probe data has been working successfully when validation traffic models?

Fråga 22 gav respondenterna möjligheten att besvara vilka parametrar som de ansåg sig haft problem med eller att resultaten ej givit det utfall som de hoppades på. Svaren på denna fråga redovisas i Figur 12 och påvisar att flest respondenter upplevde att validering med OD-matriser och trafikvolym var svårt eller inte gav önskade resultat. Även analysering av resors kompletta rutter samt olyckor hade upplevts problematiskt. Följande fråga, fråga 23, gav respondenterna möjlighet att utveckla och förklara sina svar på fråga 22. En av de tillfrågade svarade som utveckling på sitt svar att OD-matriser fungerade mindre bra eftersom ”det är oklart hur stor del av trafiken som representeras”.



Figur 12. Svar på fråga 22: For which parameters has GPS probe data been difficult to use, or not given the desired results, when validating traffic models?

På fråga 25 fick deltagarna möjligheten att komma med förslag på nya eller relativt oprövade metoder i valideringsprocessen där GPS-baserade probedata kan utvecklas och har potential

att användas. Tre av respondenterna svarade på frågan varav en svarade "Ja". En annan respondent svarade "Finns en nytta i att validera samtliga trafikmodeller som tagits fram med GPS-data. Det jag tror är en nyckel i framtida arbete med GPS-data är betrakta det för precis vad det är, ett facit på hur människor kör och använder vägnätet". En av de tillfrågade svarade att GPS-baserade probedata har potential att användas för "Syntetiska trafikmätningar beroende på GPS-data och att kalibrera mot hastigheter".

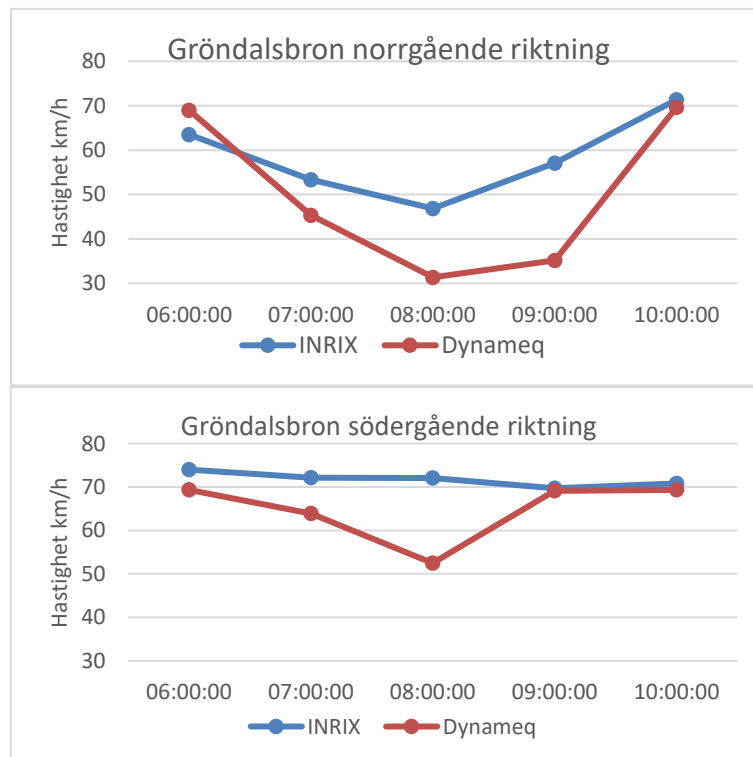
4.2 Fallstudie Region Stockholm

4.2.1 Hastigheter

Vid redovisning av procentuell avvikelse mellan de två dataseten innebär positiva värden att hastigheterna enligt Dynameq-modellen är högre jämfört med probedatan. Omvänt gäller för negativa värden.

4.2.1.1 Gröndalsbron

I Figur 13 redovisas hur medelhastigheten på Gröndalsbron förändras beroende på tidpunkt baserat på både probedata och data från Dynameq-modellen. I norrgående riktning följer de två dataseten liknande mönster med högst hastigheter i början och slutet av tidsintervallet och med de lägsta hastigheterna mellan klockan 08.00 och 09.00. Däremot är värdena från Dynameq-modellen lägre än motsvarande från probedatan, främst när hastigheterna är som lägst för de båda dataseten. Som mest skiljer det sig cirka 22 km/h, vilket enligt Tabell 3 motsvarar cirka 38% lägre hastighet.



Figur 13. Jämförelse medelhastigheter Gröndalsbron.

Tabell 3. Procentuella avvikelser vid jämförelse mellan hastigheter uppmätta i Dynameq och från probedata, Gröndalsbron. Negativa värden innebär lägre medelhastighet enligt Dynameq-modellen, positiva värden högre medelhastigheter enligt Dynameq-modellen.

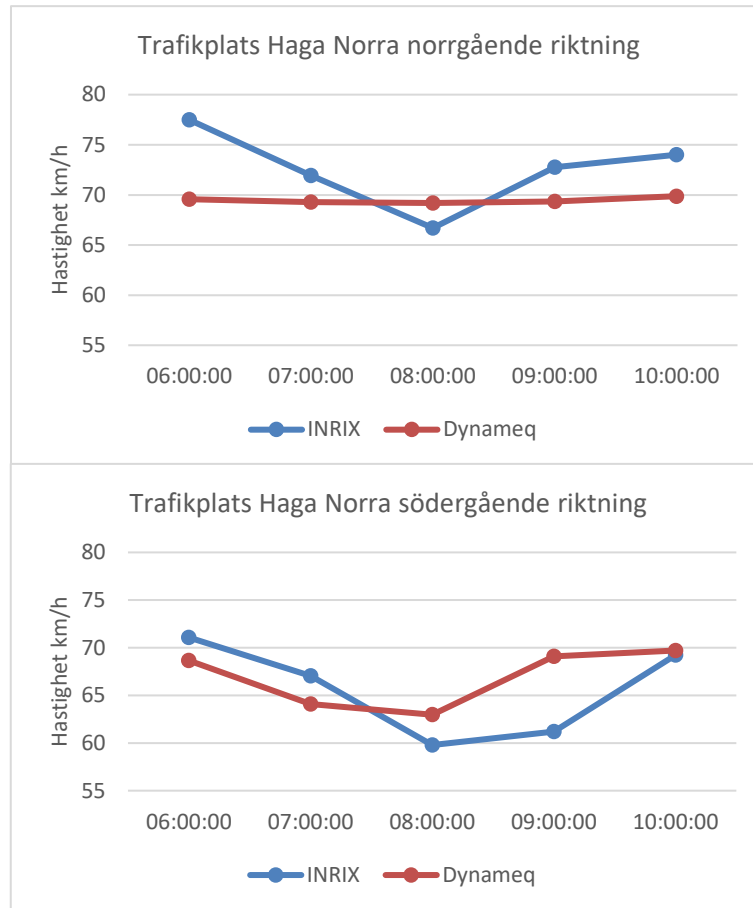
Tid	Norrgående	Södergående
06:00:00	9%	-6%
07:00:00	-15%	-11%
08:00:00	-33%	-27%
09:00:00	-38%	-1%
10:00:00	-2%	-2%

I södergående riktning går det att se tydliga skillnader på mönstret för de två olika dataseten. Baserat på probedatan är hastigheten nästintill konstant under hela intervallet. Resultatet från Dynameq-modellen har däremot en tydlig sänkning av medelhastigheten under de mest trafikerade timmarna, men med liknande värden som probedatan i början och slutet av intervallet. Störst skillnad är det vid klockan 08.00 där hastigheterna i Dynameq-modellen enligt Tabell 3 är cirka 20 km/h lägre, vilket motsvarar cirka 27%.

I båda fallen visar probedatan generellt på högre hastigheter jämfört med resultaten från Dynameq-modellen. Då skillnaderna var som störst i båda färdriktningarna överskred de 20% avvikelse och därmed uppnåddes ej kraven på modellen med hänsyn till vad som ansågs vara godkänt i dessa fall.

4.2.1.2 Trafikplats Haga Norra

I norrgående riktning vid Trafikplats Haga Norra går det ur Figur 14 att tyda viss skillnad i mönstret för hur medelhastigheten förändras beroende på tidpunkt vid jämförelse mellan probedata och resultaten från Dynameq-modellen. I Dynameq visar det på att trafiken flödar med en liknande hastighet över hela förmiddagen. Probedatan tyder däremot på att medelhastigheten sjunker runt klockan 08.00 men att hastigheten före och efter är högre jämfört med motsvarande i Dynameq. Skillnaderna kan däremot anses som små och är som störst vid 06.00 där hastigheten enligt Tabell 4 från probedatan är cirka 8 km/h högre, vilket motsvarar cirka 10%.



Figur 14. Jämförelse medelhastigheter Trafikplats Haga Norra.

Tabell 4. Procentuella avvikelser vid jämförelse mellan hastigheter uppmätta i Dynameq och från probedata, Trafikplats Haga Norra. Negativa värden innebär lägre medelhastighet enligt Dynameq-modellen, positiva värden högre medelhastigheter enligt Dynameq-modellen.

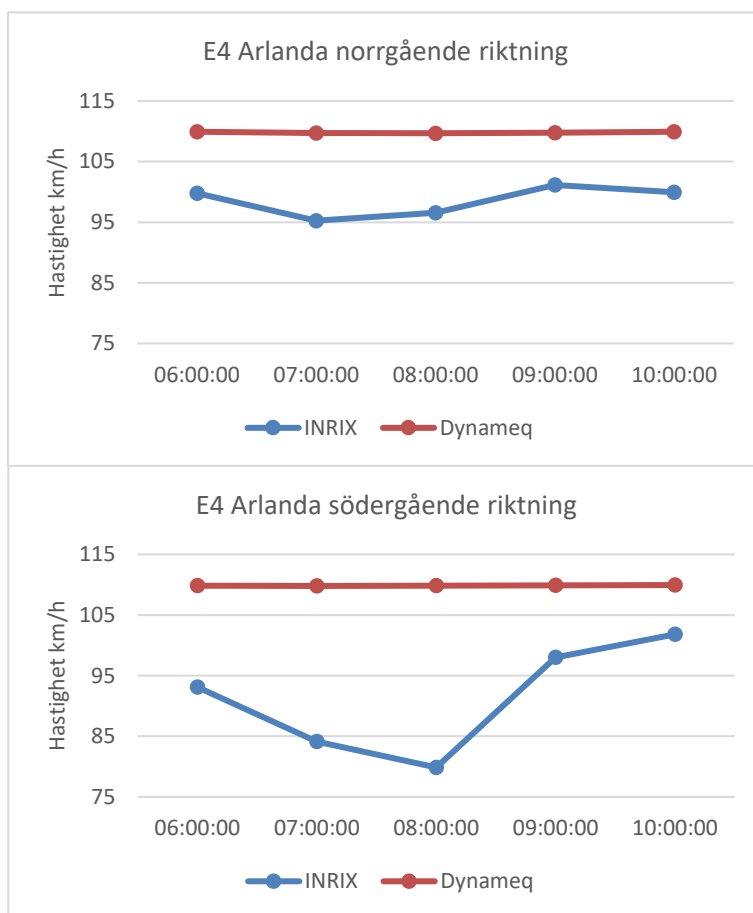
Tid	Norrgående	Södergående
06:00:00	-10%	-3%
07:00:00	-4%	-4%
08:00:00	4%	5%
09:00:00	-5%	13%
10:00:00	-6%	1%

I södergående riktning följde de två insamlingsmetoderna ett mer liknande mönster än i norrgående riktning. Hastigheterna är som högst i början och slutet av intervallet och har sin lägsta punkt vid klockan 08.00. Skillnaden mellan de två dataseten var att modellen i Dynameq snabbare går tillbaka till högre hastigheter medan mätningarna från probedata visar på lägre hastigheter en timme efter den tidpunkt där de var som lägst. Således var skillnaden som störst vid klockan 9.00, vilket enligt Tabell 4 var cirka 8 km/h, motsvarande cirka 13%.

För båda fallen visade Dynameq-modellen på mer konstanta hastigheter jämfört med probedata som båda hade högst maximala värden samt lägst minimala värden. I båda färdriktningarna var de största procentuella avvikelserna mindre än 20% och därmed ansågs värdena godkända.

4.2.1.3 E4 Arlanda

I Figur 15 redovisas resultaten från mätningar utmed väg E4 strax söder om Arlanda flygplats. I både norr- och södergående riktning går det att se liknande mönster. Hastigheterna registrerade i Dynameq är nästintill konstanta samtidigt som hastigheterna från probedata skiljer sig mer beroende på tidpunkt med sina lägsta punkter mer mot mitten av intervallet. I båda fallen är dessutom hastigheterna konstant högre i Dynameq jämfört med de som registrerats i probedata. Som mest skiljer det sig för den norrgående trafiken med cirka 14 km/h och 15% respektive cirka 30 km/h och 38% för södergående trafik enligt Tabell 5. Dessa värden anses godkända respektive icke godkända.



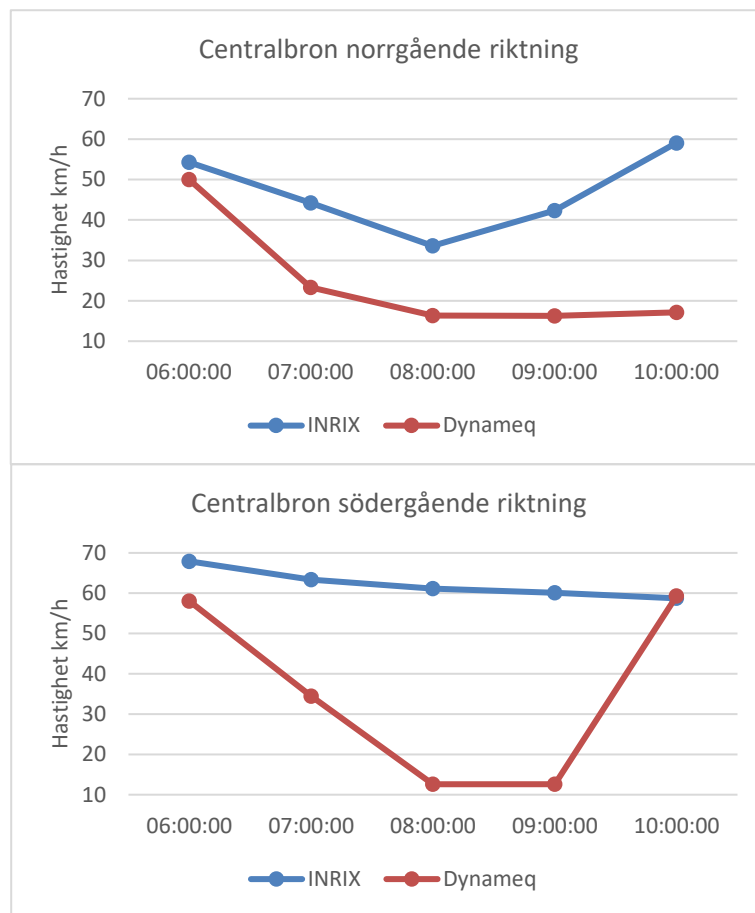
Figur 15. Jämförelse av medelhastigheter för E4 Arlanda.

Tabell 5. Procentuella avvikelser vid jämförelse mellan hastigheter uppmätta i Dynameq och från probedata, E4 Arlanda. Negativa värden innebär lägre medelhastighet enligt Dynameq-modellen, positiva värden högre medelhastigheter enligt Dynameq-modellen.

Tid	Norrgående	Södergående
06:00:00	10%	18%
07:00:00	15%	31%
08:00:00	14%	38%
09:00:00	9%	12%
10:00:00	10%	8%

4.2.1.4 Centralbron

I Figur 16 går det att tyda skillnader för medelhastigheterna som samlats in med probedatan jämfört med resultaten från Dynameq-modellen i båda färdriktningarna. I norrgående riktning är resultaten liknande vid klockan 06.00 och under de första två timmarna följer de liknande mönster med en sjunkande hastighet fram till 08.00, dock med lägre hastighet för den genererade trafiken i Dynameq. Sedan återhämtar sig hastigheterna i probedatan och blir högre för nästkommande timmar samtidigt som hastigheten från trafiken i Dynameq är nästintill konstant vid sin lägsta medelhastighet. Således är skillnaden vid klockan 10.00 mycket stor, cirka 42 km/h vilket enligt Tabell 6 motsvarar ungefär 71% lägre hastighet från Dynameq-modellen.



Figur 16. Jämförelse av medelhastigheter för Centralbron.

Tabell 6. Procentuella avvikelser vid jämförelse mellan hastigheter uppmätta i Dynameq och från probedata, Centralbron. Negativa värden innebär lägre medelhastighet enligt Dynameq-modellen, positiva värden högre medelhastigheter enligt Dynameq-modellen.

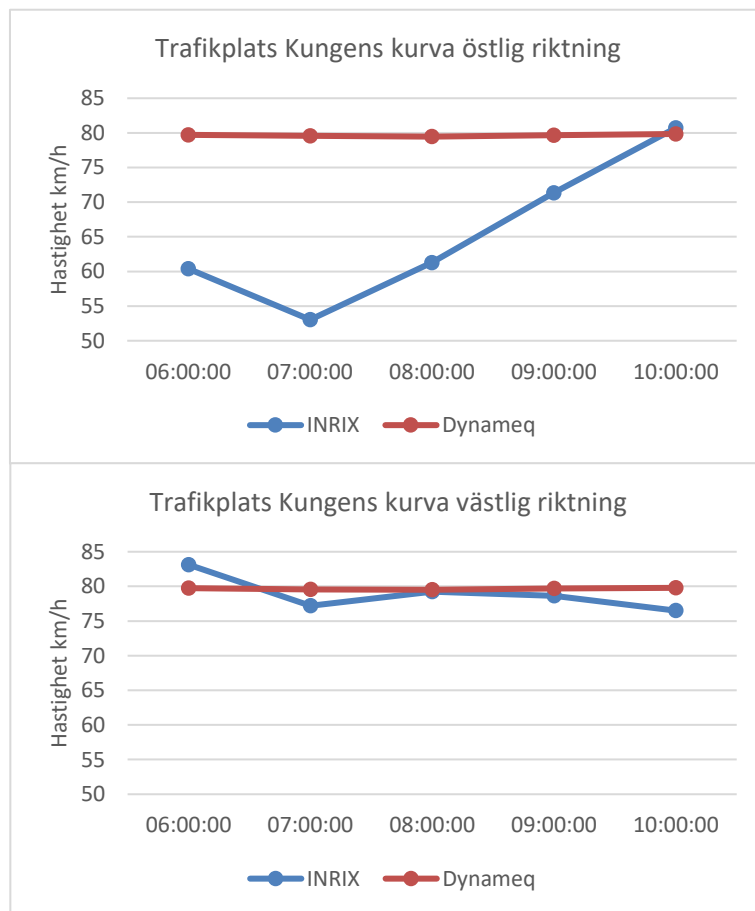
Tid	Norrgående	Södergående
06:00:00	-8%	-15%
07:00:00	-47%	-46%
08:00:00	-51%	-79%
09:00:00	-62%	-79%
10:00:00	-71%	1%

Även för trafik i södergående riktning finns det skillnader mellan de två olika dataseten. För trafiken som registrerades vid insamling av probedatan är hastighetsskillnaden över förmiddagen liten jämfört med fördelningen enligt Dynameq-modellen som har betydligt lägre hastigheter mellan klockan 07.00 och 09.00. Däremot är hastigheterna liknande vid början och slutet av tidsintervallet. Som störst är hastighetsskillnaden cirka 49 km/h, vilket enligt Tabell 6 motsvarade 79% vid klockan 08.00.

För båda färdriktningarna var hastigheterna från Dynameq-modellen lägre jämfört med hastigheterna som registrerades i probedatan. De procentuella avvikelserna var för de båda färdriktningarna markant större än 20% och därmed ej godkända.

4.2.1.5 Trafikplats Kungens kurva

Hur medelhastigheterna öster om trafikplats Kungens kurva vid väg E4/E20 varierar över förmiddagens timmar redovisas i Figur 17. I östlig riktning går det att urskilja tydlig skillnad mellan de resulterande hastigheterna från de två dataseten. Medan resultatet från Dynameq-modellen visar på att trafiken håller en jämn hastighet över förmiddagens timmar visar probedatan på en betydligt lägre hastighet mellan klockan 06.00 och 09.00 och efter det en ökning av medelhastigheten. Skillnaden är som störst cirka 27 km/h vid klockan 07.00, vilket motsvarar, enligt Tabell 7, cirka 50% lägre hastigheter registrerade med probedatan, enligt Tabell 7.



Figur 17. Jämförelse av medelhastigheter för Trafikplats Kungens kurva.

Tabell 7. Procentuella avvikelser vid jämförelse mellan hastigheter uppmätta i Dynameq och från probedata, Trafikplats Kungens kurva. Negativa värden innebär lägre medelhastighet enligt Dynameq-modellen, positiva värden högre medelhastigheter enligt Dynameq-modellen.

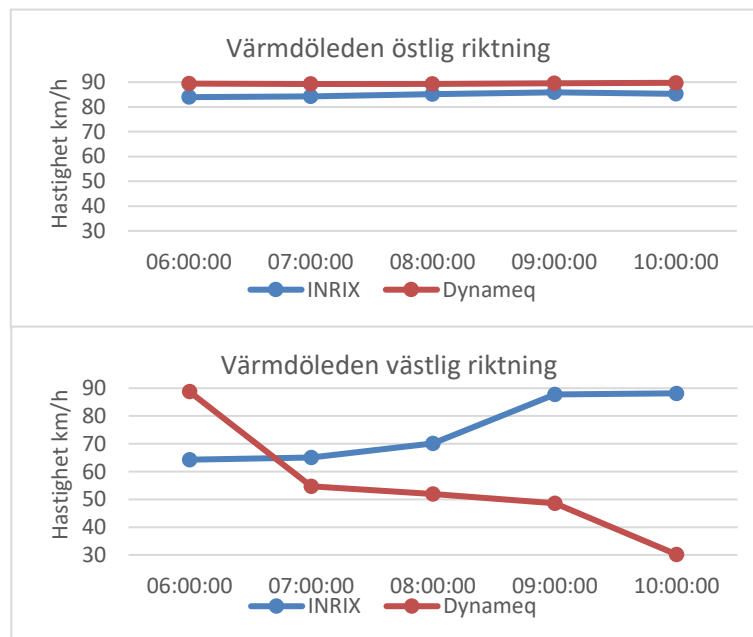
Tid	Östlig	Västlig
06:00:00	32%	-4%
07:00:00	50%	3%
08:00:00	30%	0%
09:00:00	12%	1%
10:00:00	-1%	4%

I västlig riktning däremot är hastigheterna mer lika än i motsatt färdriktning. Skillnaden är som störst vid klockan 06.00 och 10.00 där hastigheterna registrerade i probedatan är högre respektive lägre än de som Dynameq-modellen resulterar i. Mellan klockan 07.00 och 09.00 är skillnaderna mellan de två dataseten mindre och följer ett liknande mönster. Vid klockan 10.00 är skillnaden i resulterad hastighet cirka 3 km/h vilket enligt Tabell 7 motsvarar cirka 4%.

I östlig riktning var den procentuella avvikelser mellan de två dataseten som störst cirka 50% vilket ej ansågs vara godkänt. I västlig färdriktning påvisades däremot en procentuell skillnad på som mest cirka 4%, vilket ansågs som bra värden.

4.2.1.6 Värmdöleden

I Figur 18 redovisas hur hastigheternas förändring över tid skiljer sig mellan de två dataseten utmed Värmdöleden. I östlig färdriktning är hastigheterna för de båda dataseten nästintill konstanta där hastigheterna registrerade i probedatan är lägre än resulterade hastigheter från Dynameq-modellen. Skillnaden är som störst vid klockan 06.00 där hastigheterna från probedatan är cirka 6 km/h lägre, vilket enligt Tabell 8 motsvarande cirka 7%.



Figur 18. Jämförelse av medelhastigheter för Värmdöleden.

Tabell 8. Procentuella avvikelser vid jämförelse mellan hastigheter uppmätta i Dynameq och från probedata, Värmdöleden. Negativa värden innebär lägre medelhastighet enligt Dynameq-modellen, positiva värden högre medelhastigheter enligt Dynameq-modellen.

Tid	Östlig	Västlig
06:00:00	7%	38%
07:00:00	6%	-16%
08:00:00	5%	-26%
09:00:00	4%	-45%
10:00:00	5%	-66%

I motsatt färdriktning skiljer sig de två graferna mycket. I Figur 18 syns en tydlig sänkning av medelhastigheten för varje timme enligt Dynameq-modellen och enligt probedatan ökar hastigheten däremot under förmiddagen. Skillnaden är således som störst vid klockan 10.00 och är då cirka 58 km/h, vilket enligt Tabell 8 motsvarar ungefär 66%.

För de registrerade fordonen med östlig färdriktning var den procentuella avvikelser som störst cirka 7%, vilket ansågs visa på bra värden. I västlig riktning däremot var avvikelser som störst cirka 66% vilket var större än 20% och därmed inte godkänt.

4.2.1.7 Sammanfattning

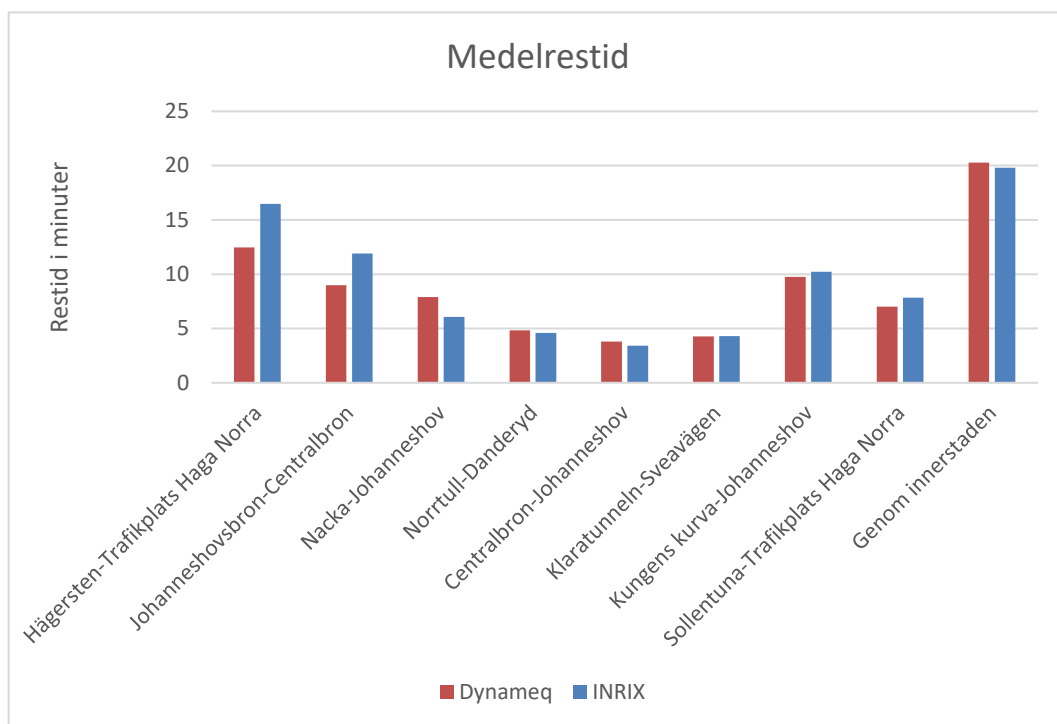
Av de sju fall där de procentuella avvikelserna var för stora var de resulterade hastigheterna enligt Dynameq högre än de uppmätta hastigheterna från probedatan hos två av fallen. För resterande fem fall med för stor avvikelse var hastigheterna i Dynameq lägre än de registrerade hastigheterna i probedatan. Vid jämförelse mellan färdriktningarna för platserna med för stora avvikelser hade tre av fem punkter med färdriktning mot Stockholm lägre hastigheter i Dynameq jämfört med probedatan. För två av fallen var hastigheterna i Dynameq högre än i probedatan. För resor med färdriktning från centrala staden var det för samtliga två platser lägre hastigheter i Dynameq jämfört med probedatan och inget fall där hastigheterna i Dynameq var högre än de uppmätta i probedatan. Dessa resultat påvisar att hastigheterna i Dynameq-modellen i fler fall underskattas jämfört med probedatan, främst i de fall där belastningen förmodas vara något lägre, i detta fall för resande bort från centrala Stockholm under morgontrafiken. För de länkar där belastningen förmodas högre var fördelningen mer jämn.

4.2.2 Restider

I Figur 19, Figur 20 och Figur 21 redovisas medelrestid, kortast restid samt längsta restid för respektive sträcka från de två dataseten. I Tabell 9 redovisas den procentuella avvikelser mellan de två dataseten för respektive parameter. Positiva värden indikerar att de resulterade restiderna i Dynameq-modellen var högre än de uppmätta restiderna i probedatan. Negativa värden visar således på lägre resulterande restider i Dynameq-modellen jämfört med de som mättes vid registrering av probedatan.

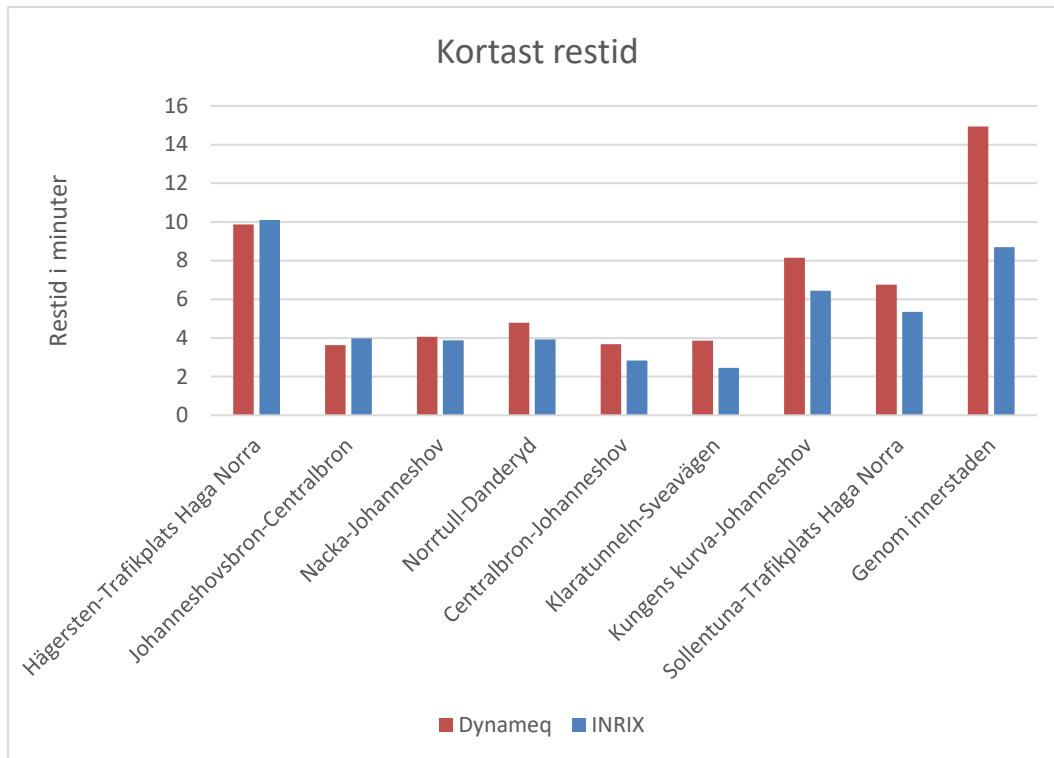
För medelrestider, Figur 19, var det fem sträckor där restiderna var kortare i Dynameq-modellen jämfört med de uppmätta restiderna hos probedatan. Dessa sträckor var Hägersten-Trafikplats Haga Norra, Johanneshovsbron-Centralbron, Klaratunneln-Sveavägen, Kungens kurva-Johanneshov och Sollentuna-Trafikplats Haga Norra. Utmed sträckorna Nacka-Johanneshov, Norrtull-Danderyd, Centralbron-Johanneshov och sträckan genom

innerstaden var värden i Dynameq däremot högre. För fyra av dessa sträckor var den procentuella avvikelsen inom 10% vilket ansågs vara bra värden, för två var avvikelsen inom 20% vilket ansågs godkänt och för tre sträckor var avvikelsen större än 20%, vilket ansågs ej vara godkända värden.



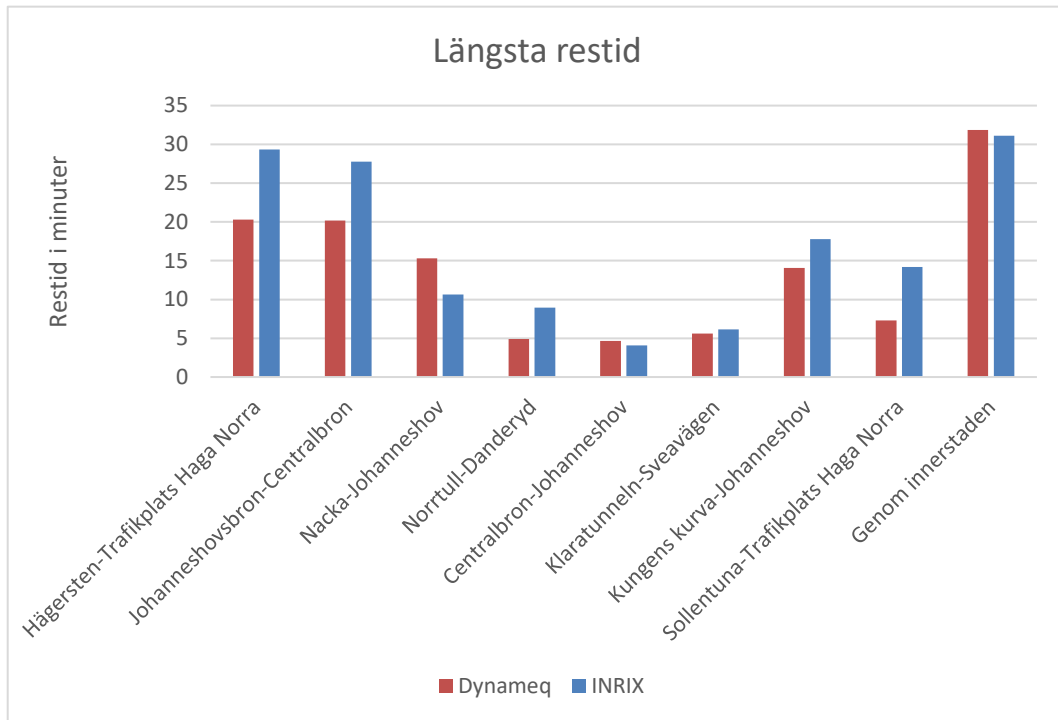
Figur 19. Jämförelse av resultat i Dynameq och från probedata för medelrestid utmed respektive sträcka.

Två av de analyserade rutterna hade kortast restider registrerade i Dynameq-modellen, Figur 20, vilket var sträckorna Hägersten-Trafikplats Haga Norra och Johanneshovsbron-Centralbron. Resterande sju sträckor hade kortast registrerade restider i probedatan. Avvikelserna för de kortaste registrerade restiderna var för tre av rutterna mindre än 10% och ansågs således vara bra. För resterande sex sträckor var avvikelsen större än 20% och ansågs således inte vara godkända.



Figur 20. Jämförelse av resultat i Dynameq och från probedata för kortast restid utmed respektive sträcka.

Analysen av resorna med längsta registrerade restid, Figur 21, påvisade vid tre av fallen att de resulterande restiderna i Dynameq-modellen var längre än restiderna registrerade i probedatan. Sträckorna som detta gällde var Nacka-Johanneshov, Centralbron-Johanneshov och sträckan genom innerstaden. För resterande sex sträckor var det från probedatan som de längsta restiderna registrerades. För sex av fallen var de procentuella avvikelserna mellan de två datasetens registrerade längsta restider större än 20% och således inte godkända. För en rutt var avvikelsen mindre än 20% och för två fall mindre än 10%. Dessa värden ansågs således vara godkända respektive bra.



Figur 21. Jämförelse av resultat i Dynameq och från probedata för längsta restid utmed respektive sträcka.

Resultaten visar därmed att medelrestiderna enligt de två dataseten är mer överensstämmande än både de kortaste och de längsta restiderna. Det går i resultaten även tyda att restiderna enligt Dynameq-modellen har en jämnare fördelning då det var i probedatan som det registrerades flest resor med både kortast och längst restid. I fem av fallen var det resultat som visade på konsekvent förhållande mellan de två dataseten, det vill säga att den ena metoden påvisade kortare, alternativt längre, restider än den andra metoden för samtliga av de tre restidsgrupperna. Rutterna Hägersten-Trafikplats Haga Norra och Johanneshovsbron-Centralbron påvisade konsekvent kortare restider enligt Dynameq-modellen och utmed sträckorna Nacka-Johanneshov, Centralbron-Johanneshov och genom innerstaden påvisades konsekvent längre restider hos Dynameq-modellen.

Tabell 9. Procentuell avvikelse mellan Dynameq-modell och probedata för respektive sträcka. Negativa värden innebär kortare restider enligt Dynameq-modellen, positiva värden längre restider enligt Dynameq-modellen.

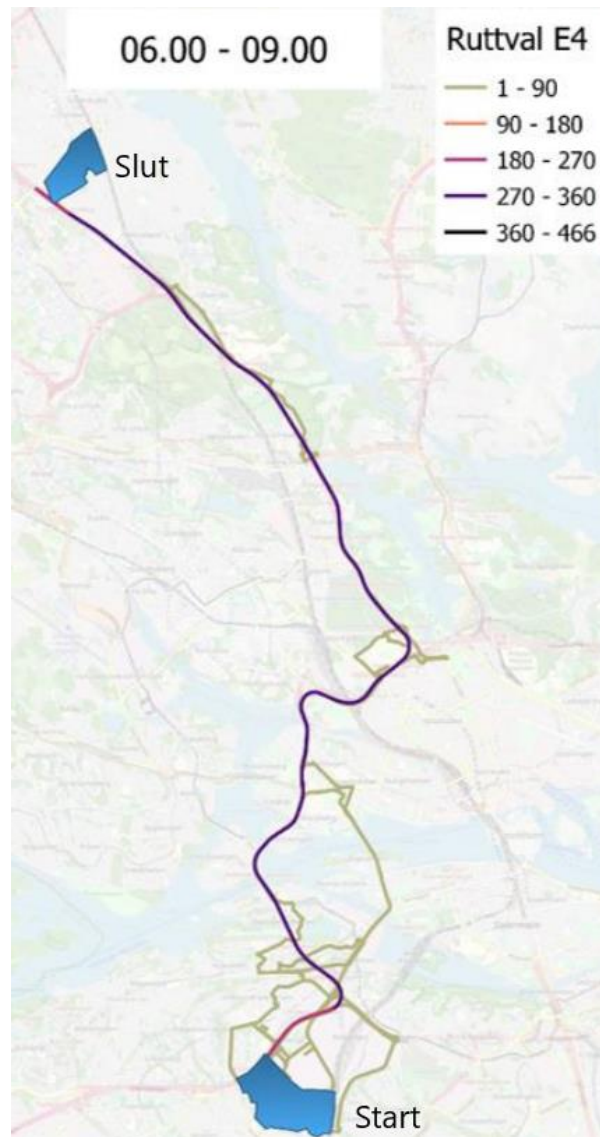
	Sträcka	Medelrestid	Kortast restid	Längst restid
1	Hägersten-Trafikplats Haga Norra	-24%	-2%	-31%
2	Johanneshovsbron-Centralbron	-24%	-9%	-27%
3	Nacka-Johanneshov	30%	5%	44%
4	Norrtull-Danderyd	5%	22%	-45%
5	Centralbron-Johanneshov	11%	30%	15%
6	Klaratunneln-Sveavägen	-1%	58%	-9%
7	Kungens kurva-Johanneshov	-5%	26%	-21%
8	Sollentuna- Trafikplats Haga Norra	-11%	26%	-48%
9	Genom innerstaden	2%	72%	2%

För sex av de nio sträckor vars restider analyserades var de procentuella avvikelserna mellan medelrestider i Dynameq-modellen och medelrestiderna enligt probedatan tillräckligt små för att de kunde anses vara godkända. Av dessa sex var avvikelserna för fyra sträckor mindre än 10% vilket anses vara bra värden. Gällande resor med kortast restid var det tre av nio sträckor som ansågs ha bra procentuella avvikelser, och för resterande sträckor ansågs de ej vara godkända. För resorna med längsta registrerade restid ansågs tre av nio sträckor ha godkända eller bra avvikelser på mindre än 20% mellan de två metoderna. Ur resultaten går det således att tyda att avvikelserna för medelrestiderna var betydligt mindre jämfört med när restiderna var som kortast respektive längst.

4.2.3 Ruttval

4.2.3.1 OD E4

I Figur 22 redovisas det resultat som Ahlberg et al. (2021) fick gällande ruttval för OD E4 från probedatan.



Figur 22. Ruttval för OD E4 baserat på probedata (Ahlberg et al., 2021).

Ur figuren går det att avläsa att majoriteten av resorna mellan de två zonerna genomfördes via väg E4 utmed hela sträckan. Det fanns däremot även alternativa rutter registrerade av ett mindre antal fordon, främst kring startzonen. Detta kan tyda på att trafikanterna gjorde olika vägval i början på sin resa för att ta sig till väg E4/E20 på kortast alternativt snabbast sätt.

I Figur 23 och Figur 24 redovisas resultaten av de simulerade trafikanternas ruttval för OD E4. I Figur 23 är både start- och slutlänk placerade utmed väg E4/E20. I Figur 24 är startlänken placerad utmed väg E4/E20 och slutlänken är placerad på avfarten mot Turebergsleden. Fler kombinationer av start- och slutlänkar testades i arbetsprocessen i

Dynameq. Däremot gav de i de flesta fall inga genererade resor alternativt mycket få och således exkluderas dessa från rapporten. Båda figurerna visar resultaten för ruttval mellan klockan 08.00 och 09.00, liknande resultat gavs för övriga timmar mellan klockan 06.00 och 09.00. Dessa redovisas under Bilaga 4. Notera att värdena för färgskalorna varierar för de två fallen.



Figur 23. Ruttval OD E4 enligt Dynameq-modell.



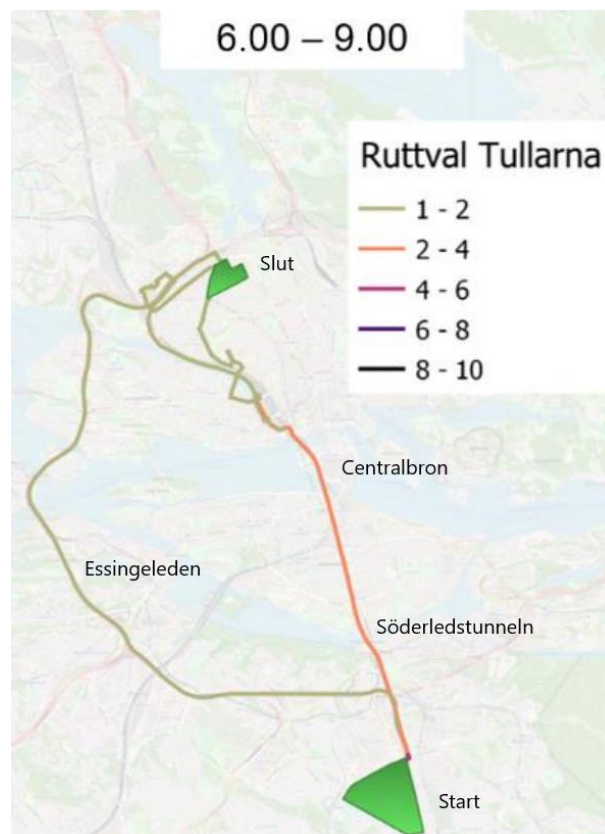
Figur 24. Ruttval OD E4 enligt Dynameq-modell.

Figur 23 och Figur 24 indikerar att det även enligt Dynameq-modellen främst är väg E4/E20 som används för att transportera sig mellan de två zonerna. De två scenarier i Dynameq som redovisas är nästintill helt lika när det kommer till vilka rutter som valdes, en avvikelse skedde där ett mindre antal fordon i Figur 23 valde att svänga av väg E4/E20 för att sedan vända tillbaka och köra på igen vid nästa trafikplats. Mängden fordon som registrerats i scenariot med slutlänk utmed E4 är större än de som registrerades i scenariot med slutlänk vid avfart på Trafikplats Tureberg. En anledning till detta förmodas vara att E4 förbi Trafikplats Tureberg är mer trafikerad än avfarten i anslutning till samma väg. Detta styrks av Trafikverkets ÅDT-mätningar från området vilket visar på mellan fyra och fem gånger så mycket passerande fordon per dygn jämfört med de som använder avfarten (Trafikverket, 2022b).

En jämförelse mellan resultaten från den insamlade probedatan och simuleringen av Dynameq-modellen visar att det finns det både likheter och skillnader. Den främsta likheten var att det för båda metoderna var en majoritet av fordonen som färdades utmed E4/E20 utmed hela sträckan. Däremot visar resultaten från probedatan på fler varianter av rutter jämfört med simuleringen av Dynameq-modellen, främst i anslutning till startzonen.

4.2.3.2 OD Tullarna

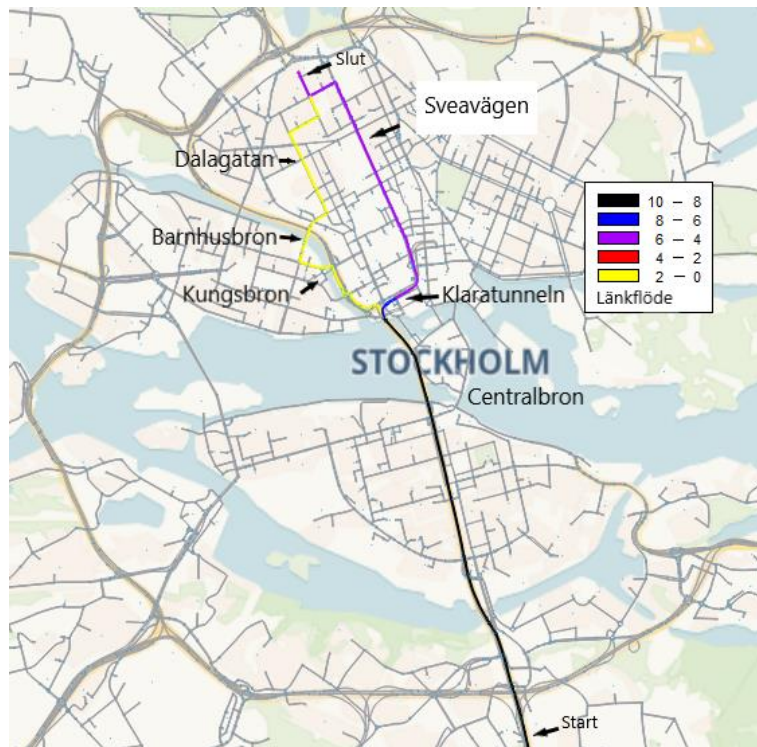
I Figur 25 redovisas det resultat som Ahlberg et al. (2021) fick gällande ruttval för OD Tullarna från probedatan.



Figur 25. Ruttval för OD Tullarna baserat på probedata (Ahlberg et al., 2021).

I Figur 25 går det att se att majoriteten av trafikanterna valde att resa mot norra delen av innerstaden genom Söderledstunneln och fortsättningsvis över Centralbron. Efter

Centralbron valde trafikanterna att antingen färdas utmed Klarastrandsleden alternativt utmed Torsgatan och Sankt Eriksgatan. Utöver dessa trafikanter valde ett fåtal att färdas utmed Essingeleden och åka runt innerstaden för att sedan nå målet från norr.



Figur 26. Ruttval OD Tullarna enligt Dynameq-modell.



Figur 27. Ruttval OD Tullarna enligt Dynameq-modell.

I Figur 26 och Figur 27 redovisas resultaten från de simulerade trafikanternas ruttval. I båda fallen är startlänken utmed väg 73 vid Enskele. Resultaten i Figur 26 är från det scenario där slutlänken valdes till Norrtullsgatan i anslutning till slutzonen och Figur 27 har sin slutlänk i cirkulationsplatsen på Sankt Eriksgatan som ansluter till slutzonen. Under arbetet

i Dynameq testades även andra kombinationer av start- och slutlänkar som ej redovisas i rapporten. Främst var det andra startlänkar som användes, men liknande arbetet med OD E4 var att några av kombinationerna inte hade simulerade resor mellan valda länkar. Båda figurerna visar på hur resandet såg ut mellan klockan 07.00 och 08.00, liknande resultat fick för övriga timmar mellan klockan 06.00 och 09.00. Dessa fall redovisas under Bilaga 5. Notera att det är skillnader i storleksordningen för färgskalorna i både Figur 26 och 28.

För scenariot där Norrtullsgatan ansattes som slutlänk, Figur 26, valde majoriteten efter att ha passerat centralbron att färdas via Klaratunneln ut på Sveavägen och sedan norrut mot slutlänken. Ett fåtal fordon följde efter Centralbron Klarastrandsleden för att sedan via Kungsbron och Barnhusbron färdas på Dalagatan mot slutlänken. När slutlänken sattes i cirkulationsplatsen på Sankt Eriksgatan valde samtliga simulerade trafikanter att färdas via Torsgatan och vidare på Sankt Eriksgatan fram till cirkulationsplatsen. Majoriteten valde före det att använda sig av Klarastrandsleden för att nå fram till Torsgatan via Kungsgatan – Vasagatan.

Vid jämförelse mellan de två metoderna syns det att det fanns en tydlig likhet i att majoriteten av resorna genomfördes via Söderledstunneln och Centralbron i båda fallen. Däremot fanns det skillnader i ruttval mellan metoderna efter att Centralbron passerats. De simulerade trafikanterna i Dynameq-modellen valde i många fall att färdas via Sveavägen medan det inte registrerades någon sådan resa hos probedatan. Hos probedatan registrerades det resor via Klarastrandsleden hela vägen till slutzonen, däremot simulerades inte någon sådan resa i Dynameq-modellen. Båda metoderna registrerade resande utmed Torsgatan och Sankt Eriksgatan. Ur probedatan gick det att se att ett fåtal resenärer valde att resa utmed Essingeleden och följaktligen runt innerstaden för att nå sitt mål, i Dynameq-modellen var det ingen av de simulerade trafikanterna som valde den rutten utan alla valde att resa genom innerstaden.

5 Diskussion

5.1 Resultatdiskussion

5.1.1 Enkätstudie

Svaren i enkäten påvisade både positiva och negativa erfarenheter kring användning av GPS-baserade probedata i valideringsprocessen vid utvecklandet av en trafikmodell. Resultaten tyder på att de tillfrågade generellt är mer positiva till att använda GPS-baserade probedata för att validera trafikmodeller, men att det samtidigt finns problem som påverkar vad som är möjligt.

Av antal svar att döma förefaller det att den parameter som flest arbetat med är OD-matriser, vilket är en parameter för vilken många även har svarat att det är en parameter som fungerar bra att använda sig av vid validering. Samtidigt är det en av de parametrar som flest ansett fungera sämre. En anledning till det är enligt en av respondenterna att det är svårt att avgöra hur stor del av trafiken som representeras av den insamlade datan. Detta kan också vara en förklaring till varför flera tillfrågade även ansåg att trafikvolym är en parameter som fungerar sämre. Trafikvolym beror likt OD-matriser starkt på hur stor andel av trafiken som representeras, även kallat penetrationsgrad, vilket är svårt att uppskatta exakt. Däremot anser en svarande att det finns potential i att kunna använda ”syntetiska” trafikräkningar från GPS-baserade probedata som en typ av valideringskälla.

Bedömningen från enkätens svar är att restider, hastigheter samt köer och trängsel är de parametrar som fungerat bäst för att använda vid att validering av trafikmodeller. Dessa tre parametrar har relativt många respondenter svarat att de fungerat bra, men inga har svarat att de har fungerat sämre. En av de tillfrågade ansåg dessutom att hastigheter från probedatan har större potential att användas än vad det gör idag.

5.1.2 Fallstudie Region Stockholm

5.1.2.1 *Hastigheter*

Totalt genomfördes det 12 analyser av hastigheter i Dynameq och probedata för analysering av hastigheter. Dessa mätningar genomfördes i båda färdriktningarna på sex olika platser i Region Stockholm. Vid sju av dessa mätpunkter ansågs de procentuella avvikelserna mellan de två metoderna vid vissa tidpunkter vara för stora för att modellen skulle anses vara godkänd. Vid övriga fem mätpunkter bedömdes värdena vara antingen bra eller godkända. Majoriteten av de uppmätta punkterna hade således procentuella avvikelser som anses vara för stora för att modellen kan anses godkänd.

I de sju fallen där avvikelserna ansågs vara för stora hade fem färdriktning in mot centrala Stockholm och resterande två hade färdriktning bort från centrala Stockholm. För de fem

fall där avvikelserna betraktades som godkända var det endast ett fall med färdriktning mot centrala Stockholm och övriga fyra fall hade färdriktning från Stockholms innerstad. Dessa resultat påvisar således att modellen, utifrån hastigheter, är bättre anpassad för förhållanden där trafikbelastningen antas vara lägre jämfört med den absoluta högtrafiken som råder vid resor in mot centrala staden på förmiddagen. En ytterligare aspekt som styrker att modellen fungerar sämre vid de tidpunkter och platser där belastningen på trafiknätet är som störst är att det ur Tabell 3 – 8, går att avläsa att det i många fall endast var under ett par timmar som avvikelserna var för stora. Avvikelseerna var i de flesta fall för stora vid de tillfällen som de uppmätta hastigheterna var som lägst, och även sannolikt då belastningen var som störst.

En förklaring till varför hastigheterna för fordon som färdas på vägar med fritt flöde har högre medelhastighet enligt probedatan än för fordonen simulerade i Dynameq kan vara att de simulerade fordonen inte väljer att köra snabbare än tillåten hastighet. Däremot finns det trafikanter i verkligheten som väljer att köra snabbare än enligt lagen, om trafiksituationen tillåter, vilket möjliggör en högre medelhastighet utmed de länkar som under det analyserade tidsintervallet är mindre belastade. Det kan även vara så att modellen ej är kodad på rätt sätt, om kapaciteten på länkarna inte är korrekta kan även resultaten bli felaktiga. Om inte efterfrågan stämmer överens med verkligheten kan det i kombination med felaktiga länkkapaciteter ge missvisande flöden och även missvisande hastigheter.

5.1.2.2 Restider

För resorna med kortast restid går det i Tabell 9 avläsa att det, för samtliga sträckor med för stora avvikelser, är längst restider enligt Dynameq-modellen, och således kortare enligt probedatan. För resorna med längsta restid är resultaten däremot omvända, för dessa resor har samtliga sträckor med för stora avvikelser, förutom en, kortare restider enligt Dynameq-modellen. Dessa resultat påvisar således att restiderna för resorna registrerade i probedatan har både de lägsta och de längsta restiderna. Detta kan på att trafikanterna enligt probedatan är mer känsliga för trängsel och köer men att de samtidigt kan färdas snabbare när möjligheten för det finns. En anledning till de kortare restiderna vid friare flöde på vägarna är att trafikanter i verkligheten kan välja att köra snabbare än vad som är lagligt, vilket de simulerade trafikanterna i Dynameq ej väljer att göra. En annan anledning till detta resultat kan vara att trafikanterna i verkligheten har haft möjligheten att göra andra ruttval, vilket kan ge större variation på restider.

För båda sträckorna genom innerstaden påvisas bra värden på avvikelserna för medelrestid och längsta restid. Samtidigt går det att se att samma sträckor med stor marginal hade de största avvikelserna för de kortaste restiderna. Samtliga sträckor med riktning in mot staden hade för stora avvikelser för de längsta restiderna, två av dem hade för stora avvikelser för de kortaste restiderna och de två andra för stora deviationer för medelrestiden. Sträckorna med riktning bort från centrala staden har i båda fallen godkända värden för avvikelserna för medelrestid mellan de två dataseten och för stora avvikelser för de kortaste restiderna. För en av sträckorna var avvikelserna för stor för de längsta restiderna och för en var avvikelserna godkänd. Ur resultaten går det även att utläsa att de tre sträckor där avvikelserna för medelrestiden ej är godkända är de enda tre sträckor där avvikelserna för kortaste restiderna är godkända.

5.1.2.3 Ruttval

För de två OD-paren fanns det både likheter och skillnader i hur trafikanterna valde att resa enligt de två olika dataseten. Likheterna för OD E4 var att samtliga trafikanter valde att resa hela eller delar av sin resa via E4/E20 enligt båda metoderna. Likheterna mellan de två metoderna för OD Tullarna var att majoriteten av trafikanterna hos probedatan valde att färdas via Söderledstunneln och Centralbron tillika samtliga trafikanter i Dynameq.

Skillnaden för de två metoderna var att det för OD E4 fanns fler olika rutter enligt probedatan, i många fall korta avstick, men även längre omledningar förekom. De flesta avvikelser inträffade utmed ruttens södra halva, ofta i anslutning till startzonen. En möjlig anledning till att de mest frekventa avvikelserna finns i anslutning till startzonen är att trafikanterna har olika rutter som är snabbast, kortast eller som upplevs snabbast från sina startpunkter ut till väg E4/E20. Eftersom det i Dynameq endast gick att ha enskilda länkar som startpunkt var detta ett fenomen som ej gick att ta hänsyn till för den metoden. En annan omständighet som ej tas hänsyn till hos Dynameq-modellen är att en trafikant kan ha flera mål med sina resor, inte bara resa mellan en start- och slutpunkt. Exempelvis kan en trafikant på sin väg till jobbet, hämta upp eller släppa av passagerare, stanna för att tanka eller handla etcetera. Händelser som dessa kan förklara de avvikelser som förekommer hos OD E4.

För OD Tullarna påvisades de flesta avvikelserna gällande ruttval när trafikanterna valde resväg efter att ha passerat Centralbron. Vilka val av rutter som de simulerade trafikanterna i Dynameq gjorde var starkt beroende på vilken länk som valdes som slutlänk, vilket kan ses vid jämförelse mellan Figur 26 och Figur 27. Som tidigare beskrivits testades flera olika scenarion där start- och slutlänkarna varierades för de olika sträckorna för att möjliggöra analys av samtliga resor som bedrevs mellan motsvarande zoner hos probedatan. När dessa slås ihop finns det vissa ruttval som förekommer både hos de registrerade resorna i probedatan samt i Dynameq-modellen. Däremot finns det flera rutter som avviker, såsom att flera resor i Dynameq-modellen genomförs via Sveavägen utan motsvarande resor hos probedatan. En orsak kan vara att modellen underskattar de förhållanden som råder i Klaratunneln och utmed Sveavägen, vilket enligt probedatan håller trafikanter borta från denna rutt. Det går även att se att trafikanterna i Dynameq-modellen för detta fall är mindre benägna att använda större vägar med högre kapacitet, såsom Essingeleden och Klarastrandsleden jämfört med trafikanterna i probedatan. Denna jämförelse påvisar att trafikanternas resonemang vid val av rutt beroende på avstånd, upplevd och faktisk restid skiljer sig något. Jämförelsen påvisar att Dynameq-modellen är något mer förenklad än probedatan, då färre alternativ till rutter registreras trots att endast en mindre andel av fordonen registreras i probedatan.

Generellt påvisar dock resultaten från de två metoderna stora likheter och många av de avstickare som erhöles ur probedatan har endast registrerats ett fåtal gånger och de har därmed liten betydelse i det stora hela.

5.1.2.4 Jämförelse mellan parametrarna

Samtliga av de tre analyserade parametrarna påvisade både likheter och skillnader mellan de två dataseten. De tre parametrarna hade alla vissa aspekter där avvikelserna mellan de två metoderna var små vilket indikerar en god överensstämmelse. Samtliga parametrar hade även avseenden där avvikelserna var för stora för att anses godkända.

För sträckor som analyserades med hänsyn till restider genomfördes det även analys av hastigheter på länkar tillhörande sträckan. I vissa fall visade resultaten på logiska samband. Men i vissa fall påvisades resultat som motsade varandra, exempelvis att en sträcka som hade kortare restid enligt Dynameq påvisade enligt samma dataset lägre hastigheter hos en uppmätt punkt utmed sträckan. Ett sådant fall var sträckan Hägersten – Trafikplats Haga Norra och mätpunkten för hastigheter på Gröndalsbron utmed Essingeleden. Medelrestiden för sträckan var kortare enligt Dynameq samtidigt som lägre hastigheter i norrgående riktning på Gröndalsbron kunde noteras. För sträckan Johanneshovsbron – Centralbron gavs ett liknande samband där medelrestiden var kortare i Dynameq för hela sträckan men att hastigheterna på Centralbron i norrgående riktning var lägre enligt Dynameq. Däremot gav motsvarande sträcka i motsatt färdriktning ett mer logiskt samband då restiderna enligt Dynameq var längre samtidigt som hastigheterna enligt Dynameq var lägre. Sträckan Nacka – Johanneshov utmed Värmdöleden visade även den på resultat som följde logiska samband med den uppmätta punkten vid Värmdöleden. Medelrestiden för den sträckan var längre i Dynameq jämfört med probedatan samtidigt som hastigheterna för fordon på Värmdöleden i västlig riktning var lägre. För resor utmed sträckan Trafikplats Kungens kurva – Johanneshov var medelrestiden enligt Dynameq kortare än hos probetadan. Hastigheterna i östlig riktning vid Kungens kurva var samtidigt högre enligt Dynameq, vilket ansågs ge ett logiskt samband. Ett ytterligare samband som ansågs logiskt var att restiderna vid mindre belastning generellt var längre enligt Dynameq-modellen, samtidigt som hastigheterna vid dessa förhållanden var lägre enligt Dynameq-modellen jämfört med probedatan. Däremot syntes inte ett lika tydligt samband vid högre belastning. Hastigheterna i dessa fall var mer likvärdiga men med en majoritet av mätningarna där hastigheterna enligt Dynameq var lägre, trots att restiderna vid högre belastning generellt var kortare enligt Dynameq-modellen. Viktigt vid dessa jämförelser är att de punkter där hastigheten analyseras endast utgör en mindre del den totala sträckan där de innefattas. För att göra mer utförliga analyser bör fler punkter analyseras för hastigheter.

Vid jämförelse mellan restider och ruttval för sträckor innefattande i OD Tullarna visas det att trots snarlika medelrestider via Klaratunneln och Sveavägen för de två metoderna skiljer sig ruttvalen för OD-Tullarna. Enligt Dynameq-modellen väljer flera av trafikanterna att resa via Sveavägen, men enligt probedatan gör ingen av trafikanterna detta val. Att ruttvalen för denna delsträcka skiljer sig markant trots de snarlika restiderna ansågs besynnerligt. Enligt probedatan finns det trafikanter hos OD Tullarna som väljer att färdas via Essingeleden, vilket Dynameq-modellen ej påvisar. Samtidigt var medelrestiderna för sträckan Hägersten – Trafikplats Haga Norra, vilket till stor del innefattar Essingeleden, avsevärt lägre enligt Dynameq-modellen jämfört med probedatan. Detta ger ett underligt samband då kortare restider generellt är attraktivt för trafikanter vid val av rutter, vilket enligt jämviktsteorin bör generera resor utmed Essingeleden för OD Tullarna enligt Dynameq-modellen också om trafikanter enligt probedatan väljer den ruten trots längre restider. OD Tullarna gav därmed exempel som visade på att trafikanterna enligt de två olika dataseten kan ha resonerat olika med hänsyn till jämviktsteorin. De verkliga trafikanterna var kanske generellt mer känsliga för risken med köer och trängsel och därmed längre restider. Således kan de ha valt att färdas med rutter som var längre men som de ansåg mer konsekventa och därmed anser sig ha större chans att komma fram i tid. De verkliga trafikanterna verkade bete sig mer i enlighet med det som beskrivs som stokastisk användarjämnvikt, och modellens trafikanter som det som beskrivs som deterministisk användarjämnvikt. Det kan även bero på hur kapaciteten hos de olika länkarna i modellen stämmer överens med verkligheten, då detta enligt jämviktsteorin även är en relevant parameter vid ruttval.

5.2 Metoddiskussion

5.2.1 Enkätstudien

Den första delen av studien omfattade en enkätstudie där personer som arbetar med trafikrelaterade frågor inom Rambolls organisation deltog. Resultaten gav goda och givande svar på de enklare frågorna där endast parametrar som fungerat bra och sämre skulle nämnas. Däremot var svaren på frågorna där möjligheten fanns att utveckla sina svar och resonemang mer bristfälliga, då flera av de tillfråga ej besvarade dessa. Att få en motivering till varför något fungerar bra eller dåligt ger möjligheten att utveckla vidare det som fungerar, och ändra på det som fungerar sämre. Enkäten var uppbyggd på ett sådant sätt att inga frågor var obligatoriska och om detta hade gjorts annorlunda hade det möjligtvis ökat antalet utvecklande svar och därmed minskat rummet för egen tolkning av de tillfrågades svar. Ett alternativ hade varit att istället för en enkätstudie genomföra en intervjustudie för den del som innefattade validering av trafikmodeller. Eftersom antalet svaranden var relativt få hade en intervjustudie varit lämpligt. En intervjustudie hade sannolikt givit mer utvecklade och användbara svar då det ger en större möjlighet för deltagaren att diskutera. Däremot är en intervjustudie mer tids- och resurskrävande än en enkätstudie, vilket måste tas till hänsyn för vid ett tidsbegränsat arbete som detta. För att öka svarsfrekvensen hade personer från andra organisationer än bara Ramboll kunnat tillfrågas, detta hade även gett möjligheten att analysera svar från personer tillhörande andra arbetsroller, exempelvis beställare, och därmed ge fler synvinklar på GPS-baserad probedata. Men eftersom enkäten delvis hade som syfte att vara en intern kunskapsbas hos Ramboll tillfrågades endast personer som arbetade på Ramboll.

5.2.2 Fallstudie Region Stockholm

I fallstudien användes GPS-baserade probedata som köpts in från företaget INRIX. Den inköpta datan bestod av en betydande mängd waypoints, cirka 24,5 miljoner. Dessa punkter hade samtliga information rörande tidpunkt, position samt enhetens egenskaper. Att hantera denna data krävde således omfattande arbete i programmeringsverktyg för att kunna extrahera användbara data. För att arbetet med GPS-baserade probedata skulle kunna genomföras snabbt och effektivt krävs det således kompetens i användning av programmeringsverktyg. Då sådant arbete kan upplevas som komplext kan metoden anses brista något i användarvänlighet, vilket skapar en risk att metoden inte kan användas fullt ut och därmed hindra utvecklingen. För detta arbete upplevdes en del problem med just detta då författaren inte besatt någon erfarenhet av arbete med RStudio. På grund av att arbetet var tvunget att kombineras med upplärning av grunderna i programmering i RStudio blev arbetet sannolikt mindre effektivt genomfört än vad som är möjligt. Trots detta framställdes det genom datahanteringen goda resultat som var möjliga att analysera. Ett effektivare arbete hade möjliggjort att testa fler parametrar och fler fall för respektive parameter, vilket sannolikt hade skapat en bredare och mer verklighetsförankrad validering. Exempelvis vid analyseringen av ruttval hade fler OD-par kunnat undersökas om kunskapen kring hur den typen av data hanterades fanns.

Eftersom det endast är en liten andel av trafiken, cirka 1–2% i denna studie, som representeras i probedatan finns det en risk att resultaten ej avspeglar verkligheten fullständigt. Eftersom GPS-positionering är känsligt för störningar om det råder

missgynnsamma förhållanden såsom förekomst av fysiska hinder och missgynnsamma väderförhållanden under tiden probedatan registreras finns det därmed även risk för att viss del av datan är missvisande. Det kan exempelvis verka som ett fordon färdas på fel väg eller åker i fel hastighet, vilket skapar orimliga resultat. Eftersom enskilda resor innefattar många waypoints är sannolikheten för att någon eller flera waypoint påvisar felaktig information därmed påtaglig. Med hänsyn till att det är en liten andel av trafikanterna som representeras i probedatan tillsammans med den påtagliga sannolikheten för fel finns risken att penetrationsgraden blir lägre än vad som anses vara tillräckligt för att säkerställa representativa data.

Vid jämförelse mellan ruttval i Dynameq-modellen och probedatan påvisas de brister som finns hos en trafikmodell på grund av att det är en förenkling av verkligheten. För ruttvalen till exempel påvisar Dynameq-modellen betydligt färre alternativ jämfört med probedatan. Att Dynameq-modellen är en förenkling av verkligheten kan även vara en förklaring till varför Dynameq-modellen är mindre känslig än probedatan för restidsdifferentiering beroende på belastning och trängsel på vägarna. Det föreföll även finnas brister hos Dynameq som modelleringsverktyg vid analys av ruttval i jämförelse med informationen erhållen ur probedatan. En brist som upplevdes var det faktum att det ej gick att analysera ruttval mellan två zoner, vilket gick med probedatan. I Dynameq var det endast mellan två specifika länkar som ruttvalen kunde analyseras, i många fall kan det vara av intresse att analysera samtliga resor mellan två zoner och inte enbart mellan två specifika länkar. I denna studie var det dessutom ett problem vid jämförelsen mellan de två olika metoderna av ruttvalsanalys. Ska en helt korrekt jämförelse göras bör det vara med snarlika metoder. En ytterligare brist som upplevdes i fallstudien var att analysen av restider endast var möjlig för specifika rutter, och inte för samtliga resor mellan två punkter. I denna studie hade det varit av intresse att undersöka just detta. Noterbart är däremot att både ruttval mellan två specifika länkar och restider för specifika sträckor går att analysera från probedatan, vilket inte gjordes i denna studie på grund av författarens bristande kunskap inom datahantering med programmeringsverktyget. Det hade förmodligen gett mer ännu mer jämförbara värden mellan de två data-seten, men det hade krävt databearbetning som författaren ej behärskade.

5.3 Sammanfattande diskussion

Svaren på enkäten tydde på att det finns erfarenhet med att använda sig av GPS-baserade probedata i arbetet med att validera trafikmodeller. Flera av de svaranden uttryckte att det finns positiva aspekter med att använda probedata och ansåg att det finns både potential hos redan beprövade områden samt områden som är utforskade för den typen av data. Fallstudien påvisade att det för de tre parametrarna hastigheter, restider och ruttval går att ta fram data som är jämförbara med de resultat som går att extrahera och analysera från dynamiska trafikmodeller, i detta fall i Dynameq. Förmågan hos ett data-set att tillhandahålla värden jämförbara med utdata från modellen är enligt teorin för modellvalidering grundläggande. Arbetsprocessen i fallstudien påvisade flexibiliteten hos den GPS-baserade probedatan, det vill säga att det möjliggör analys för nästintill samtliga delar av vägnätet, vilket även den teoretiska bakgrunden förespråkade. Det finns som beskrivet i teorin parametrar som är mycket svåra eller tidskrävande att analysera med vissa metoder, men som med probedata i kombination med rätt kompetens går att arbeta med enklare och snabbare.

Fallstudien påvisade däremot även den komplexitet som finns i att använda GPS-baserade probedata som valideringskälla. Eftersom datahanteringen kräver ett omfattande arbete i programvaror, som för en oerfaren person utan den kompetensen kan innebära ett visst motstånd, kan processen med att få fram analyserbara och jämförbara data anses som komplex. Utan standardiserade och beprövade metoder finns risken att resultaten ej blir konsekventa samt att arbetet med och utveckling av metoden förhindras då färre personer har möjligheten att arbeta med det. En nackdel med GPS-baserade probedata som både uttrycks i enkätsvaren samt styrks i teorin är problemen som finns relaterade till penetrationsgrad och förekomst av mätningsfel vid GPS-positioneringen. Det finns en påtaglig risk vid insamling av GPS-baserade probedata att en liten andel av trafikanterna representeras samt att en för stor andel av resorna upplever något typ av mätningsfel. Eftersom penetrationsgraden dessutom förändras konstant försvåras de typer av dataanalyser som är i behov av penetrationsgraden. I fallstudien då hastigheter och restider analyserades var felaktiga mätningar ett problem som inte upplevdes som påtagligt, men när Alhberg et al. (2021) analyserade ruttval var det en större andel resor som upplevde felaktiga mätningar.

Vid jämförelse mellan utdata från Dynameq-modellen och den bearbetade probedatan går det att uttyda att det finns skillnader, även om värdena i stort visat sig vara jämförbara. Utifrån hastigheter och restider går det se att trafiken i verkligheten påverkas mer av systemets belastningsgrad jämfört med den simulerade trafiken i Dynameq-modellen. Ruttvalen vid OD Tullarna indikerar dessutom på att trafikanterna i verkligheten är räddare för trängsel och köer och därmed väljer att göra ruttval som anses ge säkrare och mer konsekventa restider. Dynameq-modellen verkar därmed jämfört med den GPS-baserade probedatan brista något i skillnaden i beteenden mellan olika trafikanter. Men att en trafikmodell tar mindre hänsyn till individuellt beteende kan anses vara rimligt då samtliga trafikanter i verkligheten tänker och beter sig olika i trafiken.

6 Slutsatser

Slutsatsen från det genomförda arbetet blir att hantering och analysering av GPS-baserade probedata har god potential att utvecklas till en vedertagen metod vid validering av trafikmodeller. Den GPS-baserade probedatan ger möjligheten att ta fram värden som är jämförbara med utdata från en trafikmodell, exempelvis i modelleringsverktyget Dynameq. Fördelarna är att GPS-data kan samlas in utmed hela vägnätet vilket skapar en bättre representation vid valideringen och eftersom den samlas in från fordon som färdas på vägarna representerar verkliga förhållanden. Insamling av GPS-baserade probedata kräver dessutom ingen extra installerad utrustning eller extra resurser utmed vägnätet. Däremot finns det områden där det krävs utveckling, såsom standardiserade och effektivare metoder för datahantering. Att utöka andelen fordon som representeras är också av vikt för att dels få säkrare data, dels för att ta höjd för resor med felaktiga mätningar. Viktigt är även att notera att hur trafikanters beteende skiljer sig är något som i hög grad påverkar den GPS-baserade probedatan. Det kan vara svårt att åstadkomma en trafikmodell där trafikanternas beteende speglar verkligheten fullt ut eftersom människor är individer och därmed svåra att efterlikna med en modell. Därmed kan värdena från den GPS-baserade probedatan, trots att de är jämförbara, påvisa tydliga skillnader jämfört med utdatan från modellen i vissa förhållanden. Fallstudien påvisar även att det finns problem i Dynameq-modellen då valideringen i flera fall visade på för stora avvikelser mellan de två data-seten. Dels kan trafikanternas beteende i modellen förändras för att mer efterlikna verkligheten, dels kan "fysiska" aspekter såsom kapacitet och efterfrågan förbättras.

7 Rekommendationer

För att öka användarpotentialen för personer vars kompetens inom programmering och datahantering inte är omfattande bör det arbetas med att framställa standardiserade metoder för den datahantering som krävs för att kunna använda probedatan fullt ut. Om det framställs standardiserade metoder ökar sannolikheten för konsekventa resultat samt effektivare arbetsprocesser. Det kan även bidra till att fler personer kan arbeta med det, vilket bör driva vidare utvecklingen av metoden ytterligare.

Eftersom den GPS-baserade probedatan enligt teorin och enkätstudien påvisade en stor variation av parametrar möjliga för analys bör det således göras liknande studier som denna där andra parametrar testas som den validerande datan.

I enkätstudien uttrycks det från flera svaranden att det finns problem i att skatta efterfrågan och flöden med hjälp av GPS-baserade probedata på grund av att det är svårt att ha ett exakt värde på penetrationsgraden, vilket är vitalt för att få korrekta värden vid en sådan skattning. Däremot finns det stort intresse för att finna metoder för att använda GPS-baserade probedata till att skatta efterfrågan och flöden med säkra resultat eftersom det med andra metoder är svårt att göra detta med empiriska data. Således bör det göras fortsatta studier i hur det med hjälp av säkrare penetrationsgrader går att utveckla skattning av efterfrågan och flöden med hjälp av GPS-baserade probedata.

Valideringen av Dynameq-modellen påvisade att trots att modellen redan kalibrerats och validerats finns det behov av ytterligare förbättringar. Ett område där modellen bör förbättras är hur trafikanterna beter sig och tänker. Resultaten från analysen av ruttval indikerade att trafikanterna i verkligheten är mer försiktiga och rädda för trängsel och efterföljande restidsökning. De verkliga trafikanterna är dessutom mer flexibla när det kommer till ruttval då fler olika rutter registrerades. Fallstudien påvisade även att det finns förbättringspotential hos modelleringsverktyget INRO Dynameq. Den GPS-baserade probedatan upplevdes mer flexibel då, om förmågan finns, fler parametrar och variationer på dessa kunde analyseras jämfört med i Dynameq. Exempelvis gick det i Dynameq endast att analysera restider för specifika rutter, men inte för samtliga resor mellan två punkter, jämfört med att båda varianterna kan analyseras med den GPS-baserade probedatan. Att kunna analysera ruttval för samtliga resor mellan två zoner och inte endast mellan två specifika länkar i Dynameq hade varit positivt.

Referenser

- Ahlberg, J., Danielsson, A., Drageryd, L., Gundlegård, D., Ramsey, J., Sjöholm, A. & Sjöstrand, S. (2021). *Probedata, förstudie kring användning av GPS-baserade probedata för skattning av hastigheter, länklöden och ruttval*. Linköping universitet och Ramboll: u.o
- Ahsani, V., Amin-Naseri, M., Knickerbocker, S & Sharma, A (2018). *Quantitative analysis of probe data characteristics: Coverage, speed bias and congestion detection precision*. Ames, Iowa, USA: Iowa State University
- Caas, J., Perarnau, J. & Torday, A. (2011). *The need to combine different traffic modelling levels for effectively tackling large-scale projects adding a hybrid meso/micro approach*. Barcelona, Spanien: University of Vic
- Chiu, Y. C., Bottom, J., Mahut, M., Paz, A., Balakrishna, R., Waller, T & Hicks, J. (2011). *Dynamic Traffic Assignment*. Washington DC, USA: Transportation Research Board.
- Eriksson, J., Lindborg, E., Adell, E., Holmström, A., Silvano, A.P., Nilsson, A., Henriksson, P., Wiklund, M. & Dahlberg, L. (2018). *Nya sätt att samla individuell resvaneinformation*. Linköping: VTI, Statens väg- och transportforskningsinstitut. Diarienummer: 2017/0508-8.3
- Geotab (2020). *What is GPS?* <https://www.geotab.com/blog/what-is-gps/> [2022-03-01]
- INRO (2022). *Dynameq*. <https://www.inrosoftware.com/en/products/dynameq/> [2022-03-16]
- INRIX (2022). *About*. <https://inrix.com/about/> [2022-05-23]
- Jamali, S (2021). *Global positioning system (GPS)* [Internt material]. Lund: Lunds Tekniska Högskola
- Johansson, A (2021). *Semistrukturerade intervjuer: Saker du behöver veta om intervjuer*. <https://andreasjohanssonux.se/semistrukturerad-intervju/>. [2022-05-23]
- Lantmäteriet (2022). *Felkällor vid GNSS-mätning*. <https://www.lantmateriet.se/sv/Kartor-och-geografisk-information/gps-geodesi-och-swepos/GPS-och-satellitpositionering/Metoder-for-GNSS-matning/Felkallor-vid-GNSS-matning/> [2022-02-22]
- Lindborg E., Johansson, M. & Sundbergh, P. (2016). *Resvaneundersökningar som indata till persontransportmodeller - problem, möjligheter och framtida behov i Sverige och Norge*. Stockholm: Trafikanalys.
- Markovic´, N., Sekula, P., Laan, Z., Andrienko, G & Andrienko, N (2018). *Applications of Trajectory Data From the Perspective of a Road Transportation Agency: Literature Review and Maryland Case Study*. IEEE transactions on intelligent transportation systems: u.o

Mitrovich, S., Montesano, G., Pugliese, F., Mancini, M., Valenti, G, (2010). *Using GPS-data from a sample of private cars for modelling the urban traffic*. Lissabon, Portugal: 12th WCTR, World Conference on Transport Research Society

Nationalencyklopedin (2022). *Kvantitativ metod*.

<https://www.ne.se/uppslagsverk/encyklopedi/l%C3%A5ng/kvantitativ-metod> [2022-04-22]

Open Street Map (2022). *Om*. <https://www.openstreetmap.org/about>. [2022-05-23]

Olstam, J., Köhler, J., Tapani, A., Perez R & Wahlstedt, J (2014). Generell metodbeskrivning. I Bång, K-L., Olstam, J & Köhler, J. (2014). *Handbok för kapacitetsanalys med hjälp av simulering*. Trafikverket: u.o, ss. 11-50

Szeto & Wong (2011) *Dynamic traffic assignment: model classifications and recent advances in travel choice principles*. Hong Kong: Department of Civil Engineering, The University of Hong Kong.

Söder, I (2020). *Trafikanalys Södra Storvreta*. Uppsala: WSP

Toledo, T & Koutsopoulos, H. N. (2004) *Statistical validation of traffic simulation models*. Cambridge, England: Center for transport and logistics.

Traffic forecasting resource (2022) *Model calibration and validation*.

https://tfresource.org/topics/Model_calibration_and_validation.html [2022-05-30]

Trafikverket (2021a). *Trafikverkets Forsknings och innovationsplan för åren 2021-2026*. Borlänge: Trafikverket.

Trafikverket (2021b) *PM – Kodningsprinciper och validering av regional Dynameq-modell V 0.8c* [internt material] Trafikverket: u.o

Trafikverket (2022a). *Sampers*. <https://www.trafikverket.se/tjanster/system-och-verktyg/Prognos--och-analysverktyg/Sampers/> [2022-03-17]

Trafikverket (2022b). *Trafikflödeskartan*. <https://vtf.trafikverket.se/SeTrafikinformation> [2022-05-03]

Wärnhjelm, M (2005). *Kostnadseffektiva resvaneundersökningar*. Stockholm: Vägverket Samhälle och Trafik.

Bilagor

Bilaga 1. Rubriker i tabeller för Trajectories

```
----- #Trajectories -----
-----
names(trajs)[names(trajs)=="X1"] <- "TripID"
names(trajs)[names(trajs)=="X2"] <- "DeviceID"
names(trajs)[names(trajs)=="X3"] <- "ProviderID"
names(trajs)[names(trajs)=="X4"] <- "Timezone"
names(trajs)[names(trajs)=="X5"] <- "TrajIdx"
names(trajs)[names(trajs)=="X6"] <- "TrajRawDistanceM"
names(trajs)[names(trajs)=="X7"] <- "TrajRawDurationMillis"
names(trajs)[names(trajs)=="X8"] <- "SegmentId"
names(trajs)[names(trajs)=="X9"] <- "SegmentIdx"
names(trajs)[names(trajs)=="X10"] <- "LengthM"
names(trajs)[names(trajs)=="X11"] <- "CrossingStartOffsetM"
names(trajs)[names(trajs)=="X12"] <- "CrossingEndOffsetM"
names(trajs)[names(trajs)=="X13"] <- "CrossingStartDateUtc"
names(trajs)[names(trajs)=="X14"] <- "CrossingEndDateUtc"
names(trajs)[names(trajs)=="X15"] <- "CrossingSpeedKph"
names(trajs)[names(trajs)=="X16"] <- "OnRoadSnapCount"
names(trajs)[names(trajs)=="X17"] <- "ErrorCodes"

trajshead<-trajs

write.csv(trajshead, "Examensarbete\\data\\data\\trajs\\trajshead.csv")
```

Bilaga 2. Kod i RStudio för att beräkna Hastigheter

```
-----#Gröndalsbron-----
#Norrgående riktning-----
##### Baserat på Trajectories #####
Gröndalsbron_Norr_Filter_Hastighet<-
  filter(trajshead,
         CrossingSpeedKph>0,
         CrossingSpeedKph<200,
         SegmentId %in%
c("45049310_0","45049310_1","45049310_2","45049310_3","45049310_4","45049
310_5","45049310_6","45049310_7","45049310_8","45049310_9"))%>%
  separate(CrossingStartDateUtc,
           sep=" ",
           into=c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
           sep=":",
```

```

        into=c("CrossingstartHour", "CrossingStartMinute",
"CrossingStartSecond"))
##### Baserat på Trajectories1 #####
Gröndalsbron1_Norr_Filter_Hastighet<-
  filter(trajshead1,
    CrossingSpeedKph>0,
    CrossingSpeedKph<200,
    SegmentId %in%
c("45049310_0", "45049310_1", "45049310_2", "45049310_3", "45049310_4", "45049
310_5", "45049310_6", "45049310_7", "45049310_8", "45049310_9"))%>%
  separate(CrossingStartDateUtc,
    sep=" ",
    into=c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
    sep=":",
    into=c("CrossingstartHour", "CrossingStartMinute",
"CrossingStartSecond"))
##### Baserat på Trajectories2 #####
Gröndalsbron2_Norr_Filter_Hastighet<-
  filter(trajshead2,
    CrossingSpeedKph>0,
    CrossingSpeedKph<200,
    SegmentId %in%
c("45049310_0", "45049310_1", "45049310_2", "45049310_3", "45049310_4", "45049
310_5", "45049310_6", "45049310_7", "45049310_8", "45049310_9"))%>%
  separate(CrossingStartDateUtc,
    sep=" ",
    into=c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
    sep=":",
    into=c("CrossingstartHour", "CrossingStartMinute",
"CrossingStartSecond"))
##### Baserat på Trajectories3 #####
Gröndalsbron3_Norr_Filter_Hastighet<-
  filter(trajshead3,
    CrossingSpeedKph>0,
    CrossingSpeedKph<200,
    SegmentId %in%
c("45049310_0", "45049310_1", "45049310_2", "45049310_3", "45049310_4", "45049
310_5", "45049310_6", "45049310_7", "45049310_8", "45049310_9"))%>%
  separate(CrossingStartDateUtc,
    sep=" ",
    into=c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
    sep=":",
    into=c("CrossingstartHour", "CrossingStartMinute",
"CrossingStartSecond"))

##### Gröndalssbron ihopslagen----
Gröndalsbron_Norr_filter_merge<-
  rbind(Gröndalsbron_Norr_Filter_Hastighet,
    Gröndalsbron1_Norr_Filter_Hastighet,
    Gröndalsbron2_Norr_Filter_Hastighet,
    Gröndalsbron3_Norr_Filter_Hastighet)
Gröndalsbron_Norr_filter_merge$CrossingEndDateUtc <-
ymd_hms(Gröndalsbron_Norr_filter_merge$CrossingEndDateUtc)+hours(2)

#Södergående riktning----
##### Baserat på Trajectories #####

```

```

Gröndalsbron_Söder_Filter_Hastighet<-
  filter(trajshead,
         CrossingSpeedKph>0,
         CrossingSpeedKph<200,
         SegmentId %in%
c("205889185_0","205889185_1","205889185_2","205889185_3","205889185_4","
205889185_5","205889185_6","205889185_7","205889185_8","205889185_9"))%>%
  separate(CrossingStartDateUtc,
           sep=" ",
           into=c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
           sep=":",
           into=c("CrossingstartHour", "CrossingStartMinute",
"CrossingStartSecond"))
##### Baserat på Trajectories1 #####
Gröndalsbron1_Söder_Filter_Hastighet<-
  filter(trajshead1,
         CrossingSpeedKph>0,
         CrossingSpeedKph<200,
         SegmentId %in%
c("205889185_0","205889185_1","205889185_2","205889185_3","205889185_4","
205889185_5","205889185_6","205889185_7","205889185_8","205889185_9"))%>%
  separate(CrossingStartDateUtc,
           sep=" ",
           into=c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
           sep=":",
           into=c("CrossingstartHour", "CrossingStartMinute",
"CrossingStartSecond"))
##### Baserat på Trajectories2 #####
Gröndalsbron2_Söder_Filter_Hastighet<-
  filter(trajshead2,
         CrossingSpeedKph>0,
         CrossingSpeedKph<200,
         SegmentId %in%
c("205889185_0","205889185_1","205889185_2","205889185_3","205889185_4","
205889185_5","205889185_6","205889185_7","205889185_8","205889185_9"))%>%
  separate(CrossingStartDateUtc,
           sep=" ",
           into=c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
           sep=":",
           into=c("CrossingstartHour", "CrossingStartMinute",
"CrossingStartSecond"))
##### Baserat på Trajectories3 #####
Gröndalsbron3_Söder_Filter_Hastighet<-
  filter(trajshead3,
         CrossingSpeedKph>0,
         CrossingSpeedKph<200,
         SegmentId %in%
c("205889185_0","205889185_1","205889185_2","205889185_3","205889185_4","
205889185_5","205889185_6","205889185_7","205889185_8","205889185_9"))%>%
  separate(CrossingStartDateUtc,
           sep=" ",
           into=c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
           sep=":",
           into=c("CrossingstartHour", "CrossingStartMinute",
"CrossingStartSecond"))

```

```
##### Gröndalssbron ihopslagen----
Gröndalsbron_Söder_filter_merge<-
  rbind(Gröndalsbron_Söder_Filter_Hastighet,
        Gröndalsbron1_Söder_Filter_Hastighet,
        Gröndalsbron2_Söder_Filter_Hastighet,
        Gröndalsbron3_Söder_Filter_Hastighet)
Gröndalsbron_Söder_filter_merge$CrossingEndDateUtc <-
ymd_hms(Gröndalsbron_Söder_filter_merge$CrossingEndDateUtc)+hours(2)

#Hastigheter och plottar----
##### Plottning Medelhastigheter----
#Norrgående riktning----

Gröndalsbron_Norr_vardag<-filter(Gröndalsbron_Norr_filter_merge,
!CrossingStartDate %in%c("2019-09-21", "2019-09-22", "2019-09-28", "2019-09-
29", "2019-10-05", "2019-10-06", "2019-10-12", "2019-10-13", "2019-10-
19", "2019-10-20"))

Gröndalsbron_Norr_TidHast<-
aggregate(list(Medelhastighet=Gröndalsbron_Norr_vardag$CrossingSpeedKph),

list(Dag=cut(Gröndalsbron_Norr_vardag$CrossingEndDateUtc, "1 hour")),

      mean)%>%

  separate(Dag,
           sep = " ",
           into = c("Datum", "Tid"))

Gröndalsbron_Norr_TidHast$Tid <- strptime(Gröndalsbron_Norr_TidHast$Tid,
format = "%H:%M:%S")
Gröndalsbron_Norr_TidHast$Tid <-
as.POSIXct(Gröndalsbron_Norr_TidHast$Tid, format="%H:%M:%S")

Gröndalsbron_Norr_TidHast<-
Gröndalsbron_Norr_TidHast[order(Gröndalsbron_Norr_TidHast$Tid),]
Gröndalsbron_Norr_TidHast_medel<-Gröndalsbron_Norr_TidHast %>%
  group_by(Tid) %>%
  summarise(ave = mean(Medelhastighet))
write.csv(Gröndalsbron_Norr_TidHast_medel,
"Examensarbete\\data\\data\\trajs\\Gröndalsbron_Norr.csv")

Gröndalsbron_Norr_plot <- ggplot(Gröndalsbron_Norr_TidHast_medel,
aes(x=Tid, y=ave, color=ave)) +
  geom_point()+
  theme(axis.text.x = element_text(angle = 45))+
  xlab("Tid")+
  ylab("Medelhastighet")+
  labs(color="Medelhastighet
        (km/h",
        title = "Gröndalsbron Norrgående riktning")+
  scale_x_datetime(breaks = date_breaks("1 hour"), labels =
date_format("%H:%M", tz = "Europe/Stockholm"), expand=c(0,0))
Gröndalsbron_Norr_plot

#Södergående riktning----
Gröndalsbron_Söder_vardag<-filter(Gröndalsbron_Söder_filter_merge,
!CrossingStartDate %in%c("2019-09-21", "2019-09-22", "2019-09-28", "2019-
09-29", "2019-10-05", "2019-10-06", "2019-10-12", "2019-10-13", "2019-10-
19", "2019-10-20"))
```

```

Gröndalsbron_Söder_TidHast<-
aggregate(list(Medelhastighet=Gröndalsbron_Söder_vardag$CrossingSpeedKph)
,
list(Dag=cut(Gröndalsbron_Söder_vardag$CrossingEndDateUtc, "1 hour")),
        mean)%>%
  separate(Dag,
            sep = " ",
            into = c("Datum", "Tid"))

Gröndalsbron_Söder_TidHast$Tid <-
strptime(Gröndalsbron_Söder_TidHast$Tid, format = "%H:%M:%S")
Gröndalsbron_Söder_TidHast$Tid <-
as.POSIXct(Gröndalsbron_Söder_TidHast$Tid, format="%H:%M:%S")

Gröndalsbron_Söder_TidHast<-
Gröndalsbron_Söder_TidHast[order(Gröndalsbron_Söder_TidHast$Tid),]
Gröndalsbron_Söder_TidHast_medel<-Gröndalsbron_Söder_TidHast %>%
  group_by(Tid) %>%
  summarise(ave = mean(Medelhastighet))
write.csv(Gröndalsbron_Söder_TidHast_medel,
"Examensarbete\\data\\data\\trajs\\Gröndalsbron_Söder.csv")

Gröndalsbron_Söder_plot <- ggplot(Gröndalsbron_Söder_TidHast_medel,
aes(x=Tid, y=ave, color=ave)) +
  geom_point()+
  theme(axis.text.x = element_text(angle = 45))+
  xlab("Tid")+
  ylab("Medelhastighet")+
  labs(color="Medelhastighet
        (km/h",
        title = "Gröndalsbron södergående riktning")+
  scale_x_datetime(breaks = date_breaks("1 hour"), labels =
date_format("%H:%M", tz = "Europe/Stockholm"), expand=c(0,0))
Gröndalsbron_Söder_plot

```

Bilaga 3. Kod i RStudio för att beräkna restider

```

#Hägersten-Tp1 Haga Norra----
#SegmentID=691607316_0 46101898_0

HägerstenHaga <-
  filter(trajshead,
         SegmentId %in% c("691607316_0", "46101898_0"),
         is.na(ErrorCodes))
HägerstenHaga1 <-
  filter(trajshead1,
         SegmentId %in% c("691607316_0", "46101898_0"),
         is.na(ErrorCodes))
HägerstenHaga2 <-
  filter(trajshead2,
         SegmentId %in% c("691607316_0", "46101898_0"),
         is.na(ErrorCodes))
HägerstenHaga3 <-
  filter(trajshead3,
         SegmentId %in% c("691607316_0", "46101898_0"),
         is.na(ErrorCodes))

```

```

HägerstenHaga_merge<-
  rbind(HägerstenHaga,
        HägerstenHaga1,
        HägerstenHaga2,
        HägerstenHaga3)
HägerstenHaga_merge <- HägerstenHaga_merge %>%
  mutate(a = case_when(
    SegmentId=="691607316_0" ~ -1,
    SegmentId!="691607316_0" ~ 0
  ))
HägerstenHaga_merge <- HägerstenHaga_merge %>%
  mutate(b = case_when(
    SegmentId=="46101898_0" ~ 2,
    SegmentId!="46101898_0" ~ 0
  ))

setDT(HägerstenHaga_merge)[ , c := sum(a, b), by = TripID]

HägerstenHaga_merge <- HägerstenHaga_merge %>%
  filter(c==1)
HägerstenHaga_a <- HägerstenHaga_merge %>%
  filter(a==-1)
HägerstenHaga_b <- HägerstenHaga_merge %>%
  filter(b==2)

names(HägerstenHaga_b)[names(HägerstenHaga_b)=="...1"] <- "...1b"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="TripID"] <- "TripIDb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="DeviceID"] <- "DeviceIDb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="ProviderID"] <-
"ProviderIDb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="Timezone"] <- "Timezoneb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="TrajIdx"] <- "TrajIdxb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="TrajRawDistanceM"] <-
"TrjRawDistanceMb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="TrajRawDurationMillis"]
<- "TrjRawDurationMillisb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="SegmentId"] <-
"SegmentIdb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="SegmentIdx"] <-
"SegmentIdxb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="LengthM"] <- "LengthMb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="CrossingStartOffsetM"] <-
"CrossingStartOffsetMb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="CrossingEndOffsetM"] <-
"CrossingEndOffsetMb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="CrossingStartDateUtc"] <-
"CrossingStartDateb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="CrossingEndDateUtc"] <-
"CrossingEndDateb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="CrossingSpeedKph"] <-
"CrossingSpeedKphb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="OnRoadSnapCount"] <-
"OnRoadSnapCountb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="ErrorCodes"] <-
"ErrorCodesb"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="a"] <- "a(b)"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="b"] <- "b(b)"
names(HägerstenHaga_b)[names(HägerstenHaga_b)=="c"] <- "c(b)"

HägerstenHaga_ab <- cbind(HägerstenHaga_a,

```

```
HägerstenHaga_b)
HägerstenHaga_ab$CrossingStartDateUtc <-
ymd_hms(HägerstenHaga_ab$CrossingStartDateUtc)+hours(2)
HägerstenHaga_ab$CrossingEndDateb <-
ymd_hms(HägerstenHaga_ab$CrossingEndDateb)+hours(2)

HägerstenHaga_ab$Timediff <- as.double(HägerstenHaga_ab$CrossingEndDateb
- HägerstenHaga_ab$CrossingStartDateUtc, units = "secs")

HägerstenHaga_ab <- HägerstenHaga_ab %>%
  separate(CrossingStartDateUtc,
           sep = " ",
           into = c("CrossingStartDate", "CrossingStartTime"))%>%
  separate(CrossingStartTime,
           sep = ":",
           into = c("Hour", "Minute", "Second"))
HägerstenHaga_ab <- HägerstenHaga_ab %>%
  filter(Timediff > quantile(HägerstenHaga_ab$Timediff,0.05),
         Timediff< quantile(HägerstenHaga_ab$Timediff, 0.95),
         Hour %in% c("06", "07", "08", "09"))
Medelrestid_HägerstenHaga <- mean(HägerstenHaga_ab$Timediff)/60
Medelrestid_HägerstenHaga
Minrestid_HägerstenHaga <- min(HägerstenHaga_ab$Timediff)/60
Maxrestid_HägerstenHaga <- max(HägerstenHaga_ab$Timediff)/60
```


Bilaga 4. Ruttval enligt Dynameq OD E4





Bilaga 5. Ruttval enligt Dynameq OD Tullarna

