

Utvärdering av ankomstprognoser för tåg

Examensarbete:
Linus Rydstedt

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

© Copyright Linus Rydstedt

LTH School of Engineering
Lund University
Box 118
SE-221 00 Lund
Sweden

LTH Lunds Tekniska Högskola
Lunds universitet
Box 118
221 00 Lund

Printed in Sweden
Media-Tryck
Biblioteksdirektionen
Lunds universitet

LTH, Institutionen för Teknik och samhälle
CODEN: LUTVDG/(TVTT-5353)/1-53/2023
ISSN 1653-1922 Tryckt i Sverige av Media-Tryck, Lunds universitet
Lund 2023

Examensarbete

Thesis / Lunds Tekniska Högskola,
Institutionen för Teknik och samhälle,
Trafik och väg

CODEN: LUTVDG/(TVTT-5353)/1-53/2023

ISSN 1653-1922

Author: Linus Rydstedt

Title: Utvärdering av ankomstprognoser för tåg

English title: Evaluation of train arrival forecasts

Language: Swedish

Year: 2023

Keywords: Train delay; traffic management; evaluation; arrival forecast;
prediction.

Citation: Linus Rydstedt, Utvärdering av ankomstprognoser för tåg. Lund, Lunds universitet, LTH, Institutionen för Teknik och samhälle. Trafik och väg 2023. Thesis 386.

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

1. Abstract

In this master thesis, an evaluation of arrival forecasts for trains is made to see how various prediction sources compare to each other and how well these reflect reality. Furthermore, this thesis is a sprint of a research project that is carried out and conducted at Lund University of Technology on behalf of the Swedish Transport Administration. Unlike the research project, this thesis is based solely on a first preliminary data dump and not on data from a specific time period regarding a certain test track.

The purpose of the thesis is to examine and evaluate arrival forecasts for trains. This is done by analyzing the ability of the underlying prediction sources to create reliable arrival forecasts. The main focus is to increase knowledge of both the strengths and weaknesses of the prediction sources, as well as to see if there are situations where one or the other seems particularly suitable.

Currently, there are different ways to create arrival forecasts for trains. The most trivial is through manual calculations, but it can also be done based on automatic calculations. The automatic calculations use digital graphs and in recent years it has become increasingly common for these to be based on AI solutions, such as AIRT.

The data that has been used for analysis in this master thesis is provided by the Swedish Transport Administration and is the same that they use when creating arrival forecasts. The data consists of nationwide data points regarding arrivals and departures during the time period 2022-10-01 – 2022-11-15. The analyzed data set consists of 5,530,048 lines of code and, with the help of SQL-Server Management Studio, has been used to structure and sort.

In the master thesis, a test was made to see which of the evaluation methods Mean value, MAE and RMSE should be used when evaluating the arrival forecasts. The result shows that MAE is the best suited and has thus continued to be used in the analysis of subsequent problems.

The Swedish Transport Administration currently uses five different prediction sources when calculating arrival forecasts for trains: Amber, Tågprognos, Manual, STEG and AIRT. When analyzing which prediction source provides the most reliable arrival predictions, it emerged that AIRT provides significantly better predictions than any of the other four. The recommendation is therefore to drop the prediction sources Amber, Tågprognos, Manual and STEG and use AIRT exclusively. This is because, on average, this prediction source creates about 2.3 times better predictions than the second-best prediction source, STEG.

To answer the last problem and to follow up and see if the ranking of how well the different prediction sources performed was constant or random, a test was performed. The test was to see how the various prediction sources performed when the traffic situation on the track in the previous hour had been good, normal and bad. The results of the test showed that the ranking remained constant, with AIRT emerging as the best alternative.

Keywords

Train delay, traffic management, evaluation, arrival forecast, prediction

2. Sammanfattning

I detta examensarbete görs en utvärdering av ankomstprognoser för tåg för att se hur de olika prediktionskällorna står sig mot varandra och hur pass väl dessa efterspeglar verkligheten. Vidare är detta examensarbete en sprint av ett forskningsprojekt som genomförs och bedrivs på Lunds Tekniska Högskola på uppdrag av Trafikverket. Till skillnad från forskningsprojektet bygger detta examensarbete enbart på en första preliminär datadump och inte på data från en specifik tidsperiod avseende en viss testbana.

Syftet med examensarbetet är att undersöka och utvärdera ankomstprognoser för tåg. Detta görs genom att analysera de bakomliggande prediktionskällornas förmåga att skapa tillförlitliga ankomstprognoser. Det huvudsakliga fokuset är att öka kunskapen om såväl styrkor som svagheter hos prediktionskällorna, samt att se om det finns situationer där den ena eller andra ter sig särskilt lämplig.

I dagsläget finns det olika sätt att skapa ankomstprognoser för tåg. Det mest triviala är genom manuella beräkningar, men det kan även göras baserat på automatiska beräkningar. De automatiska beräkningarna använder sig av digitala grafer och under de senaste åren har det blivit alltmer vanligt förekommande att dessa baseras på AI-lösningar, som exempelvis AIRT.

Datan som i examensarbetet har använts för analys är tillhandahållen av Trafikverket och är densamma som de använder sig av vid skapande av ankomstprognoser. Datan består av rikstäckande datapunkter gällande ankomster och avgångar under tidsperioden 2022-10-01 – 2022-11-15. Den analyserade datamängden består av 5 530 048 rader kod och har med hjälp av SQL-Server Management Studio använts för att strukturera och sortera.

I examensarbetet görs först en granskning över vilken av utvärderingsmetoderna Medelvärde, MAE och RMSE som bör användas vid utvärdering av ankomstprognoserna. Resultatet visar på att MAE är det bäst lämpade och har således fortsatt använts vid analys av efterföljande problemställningar.

Trafikverket använder sig i dagsläget av fem olika prediktionskällor vid beräkning av ankomstprognoser för tåg: Amber, Tågprognos, Manuell, STEG och AIRT. Vid analys av vilken prediktionskälla som ger mest tillförlitliga ankomstprediktioner framkom att AIRT utgör markant bättre prediktioner än någon av de andra fyra. Rekommendationen är därför att framöver slopa prediktionskällorna Amber, Tågprognos, Manuell och STEG och uteslutande använda sig av AIRT. Detta eftersom denna prediktionskälla i medeltal skapar ca 2,3 gånger bättre prediktioner än den näst bästa prediktionskällan, STEG.

För att besvara den sista problemställningen samt för att följa upp och se ifall rangordningen på hur väl de olika prediktionskällorna presterade var konstant eller slumpmässig utfördes ett test. Testet gick ut på att se hur de olika prediktionskällorna presterade då trafikläget på banan förgående timme varit bra, normalt och dåligt. Resultatet av testet visade att rangordningen stod sig konstant, där AIRT framstår som det bästa alternativet.

Nyckelord

Tågförsening, trafikledning, utvärdering, ankomstprognos, förutsägelse

3. Förord

Detta examensarbete har utförts under vårterminen 2023, som slutmoment i utbildningen till civilingenjör inom väg- och vattenbyggnad vid Lunds Tekniska Högskola. Ämnet föreslogs av Carl-William Palmqvist, forskare på institutionen, och utgör en sprint inom ett projekt som bedrivs på uppdrag av Trafikverket.

Först och främst vill jag passa på att tacka alla som har varit involverade i detta examensarbete. Stort tack till Karl Åkerlund på Trafikverket som hjälpt till att få loss data och ge tillgång till denna. Utan datan hade arbetet inte kunnat genomföras. Vidare riktas även ett stort tack till övriga på Trafikverket som godkänt att datan använts för externt bruk. Sist men inte minst vill jag rikta ett speciellt stort tack till Carl-William Palmqvist, min handledare på Lunds Tekniska Högskola för all tid han ägnat till att vägleda, förklara och stödja mig med framför allt strukturen av den stora mängden data. Förhoppningsvis kommer denna avhandling att vara användbar för läsare och framtida forskning inom liknande ämnen eller relaterade fältarbeten.

Lund, Juni 2023

Innehåll

1. Abstract	1
2. Sammanfattning	2
3. Förord	3
4. Introduktion	1
4.1. Mål	4
4.2. Problemställning	4
4.3. Avgränsningar	4
5. Bakgrund	5
5.1. Förseningar	5
5.2. Punktlighet	6
5.3. Prognoser för tågankomster	6
6. Teknisk bakgrund	9
6.1. Litteraturstudie om tidigare forskning	9
6.1.1. Artificiell intelligens	9
6.1.2. Maskininlärning	9
6.1.3. Former av försening	10
7. Tågplanen	12
8. Grafiska tidtabeller	14
9. Digital graf	16
10. Prediktionskällor	18
10.1. STEG	18
10.2. Amber	18
10.3. Tågprognos	18
10.4. Manuell	18
10.5. AIRT	18
11. Utvärderingsmått	19
11.1. Medelvärdet	19
11.2. Mean absolute error (MAE)	20
11.3. Root mean square error (RMSE)	20
12. Metod och resultat	22
12.1. Databeskrivning	22
12.2. Datahantering	24
12.3. Vilket utvärderingsmått bör användas	25
12.3.1. Sammanfattning & rekommendation gällande lämpligast utvärderingsmått ..	27
12.4. Vilken prediktionskälla ger mest tillförlitlig ankomsttid	27
12.4.1. Sammanfattning & rekommendation gällande mest tillförlitlig ankomsttid ..	29
12.5. Vilken framförhållning är väl avvägd	29

12.5.1.	Sammanfattning & rekommendation gällande framförhållning	32
13.	Diskussion och slutsats.....	33
13.1.	Vilket utvärderingsmått bör användas	33
13.2.	Vilken prediktionskälla ger mest tillförlitlig ankomsttid.....	34
13.3.	Vilken framförhållning är väl avvägd.....	35
13.4.	Framtida utvecklingsmöjligheter	36
13.5.	Vem riktar sig prognoserna till	37
14.	Källhänvisning	39
14.1	Källförteckning	39
14.2.	Figurförteckning	43
15.	Bilaga	44

4. Introduktion

Transporter står för 23% av världens energirelaterade utsläpp av växthusgaser och nästan 75% kommer från vägförknippade fordon. Trots att man idag, globalt, i politiken aktivt arbetar för att fasa ut de fossila bränslena står dessa för 95% av världens transportenergi (United Nations, 2021). Att i stället välja resor med tåg bidrar till ett minskat utsläpp och kan göra stor skillnad för framtida livskvalitet på vår planet. I enlighet med Parisavtalet har Sverige, för att minska klimatpåverkan, skapat klimatpolitiska ramverk. Det långsiktiga klimatmålet innebär att Sverige senast år 2045 ska ha nettonollutsläpp av växthusgaser till atmosfären (Naturvårdsverket, 2022). Ett steg i rätt riktning för att uppnå detta mål är att succesivt styra bort människor från bilresor till ett mer kollektivt resande. Med en vardag som präglas av skenande bränslepriser och ökad inflation kommer allt fler att vara öppna för en övergång till kollektivt resande. Det är av stor vikt att riksdag och regering tar vara på denna chans och möjliggör för branschen att leverera punktliga resor där tidtabeller kan hållas.

Resor med tåg har historiskt sett varit en viktig del i det svenska transportnätet och antalet tågresor har under de senare åren fortsatt att öka (Gummesson, 2019). Arbetet med underhåll av järnvägen har inte följt samma takt vilket skapat stora störningar i järnvägssystemet (Honauer & Ödeen, 2019). På senare tid har tågtrafiken präglats av ett ökat antal förseningar och för att skapa en bättre punktlighet måste arbete kring kvalitét systematiseras (Trafikverket, 2022a). Med en ökande befolkning och en politik som präglas av allt större fokus inom energi och miljö får inte den redan eftersatta järnvägen tas förgiven. Under juli månad år 2022 ankom 91,1% av persontågen sina slutdestinationer i rätt tid, i enlighet med tidtabellen, och under de senaste åren har punktligheten legat omkring 90% (Trafikverket, 2022c). Detta anses av branschen som en för låg siffra, då målet är satt till 95%. Tillsammans med Trafikverket arbetar branschen gemensamt för att uppfylla detta punktlighetsmål. Trafikverket äger och förvaltar ca 90% av den svenska järnvägen och ansvarar således bland annat för den långsiktiga infrastrukturplaneringen, byggnation av ny järnväg, drift och underhåll samt sköter informationsflödet till resenärer angående rådande trafikförhållanden (Trafikverket, 2019).

Stora delar av det svenska järnvägssystemet har sina rötter från mitten av 1800-talet. Det eftersatta underhållsarbetet gör att systemet i dagsläget verkar på gränsen till sin maximala kapacitet, vilket gör att järnvägsnätet är extra känsligt för störningar (Bårström & Granbom, 2012). Det är ofrånkomligt att helt undvika störningar i tågtrafiken, men den bristfälliga kommunikationen kring störningar och dess påverkan på ankomsttider är i nuläget oacceptabel. För att ändra på människors vanor och främja tågresor framför exempelvis bil är en viktig faktor att motsvara resenärernas förväntningar på punktlighet. Att arbeta med, och utveckla predikterade realistiska tidtabeller som motsvarar verkligheten är därför av största vikt. Ett stort fokus bör därför läggas på ny innovativ teknik, ökat underhåll, realistiska tidtabeller och ett mer kommunikativt sätt att förmedla den aktuella reseinformationen. En stor mängd människor ratar idag resor med tåg på grund av dess ständiga försening och i många fall är det ovissheten och den bristfälliga informationen som ligger till grund. Genom att därför skapa och förmedla en korrekt mer realistisk förutsägelse av hur stor förseningen faktiskt är skulle resenärers och därmed allmänhetens syn på tågtrafiken kunna ändras.

Järnvägen med dess infrastruktur har under de senaste decennierna väckt flera heta samhällsdebatter och i dagsläget diskuteras järnvägens framtid av riksdag och regering (Ödebrink, 2022). För att ställa om det svenska transportsystemet och bidra till en stärkt utveckling med tillväxt, ökad konkurrenskraft och klimatomställning är det av stor vikt att maximalt kapacitetsutnyttjande finns på våra järnvägar. Genom att skapa en bättre uppfattning över hur väl ankomst- och avgångsprognosernas utfall sammanfaller med verkligheten skulle tidtabellen kunna optimeras ytterligare. Således skulle möjligen kapaciteten på järnvägen även kunna utnyttjas på ett mer optimalt sätt. Infrastrukturdebatterna har medfört att forskning inom järnvägssystem och tågtrafiken har fått ett uppsving. Järnvägsnätet är ett komplext system

där en stor mängd komponenter och tåg tvingas interagera och samsas om ett begränsat antal spår, vilket gör det extra sårbart för eftersläpande- och försenade tåg.

I en studie som avser det holländska järnvägsnätet skriver Yuan et al, (2022) att den enskilt största orsaken till störningar är fortplantning av förseningar. Med detta menar Yuan et al att ett framförvarande tåg, som är försenat, även påverkar punktligheten för ett bakomliggande tåg. Detta utgör ett stort problem för tillförlitligheten för såväl driftledningscentralen som för passagerarna. *Bild 1* nedan visar ett exempel på hur en försening av tåg 1 kan fortplanta sig till andra bakomliggande tåg.

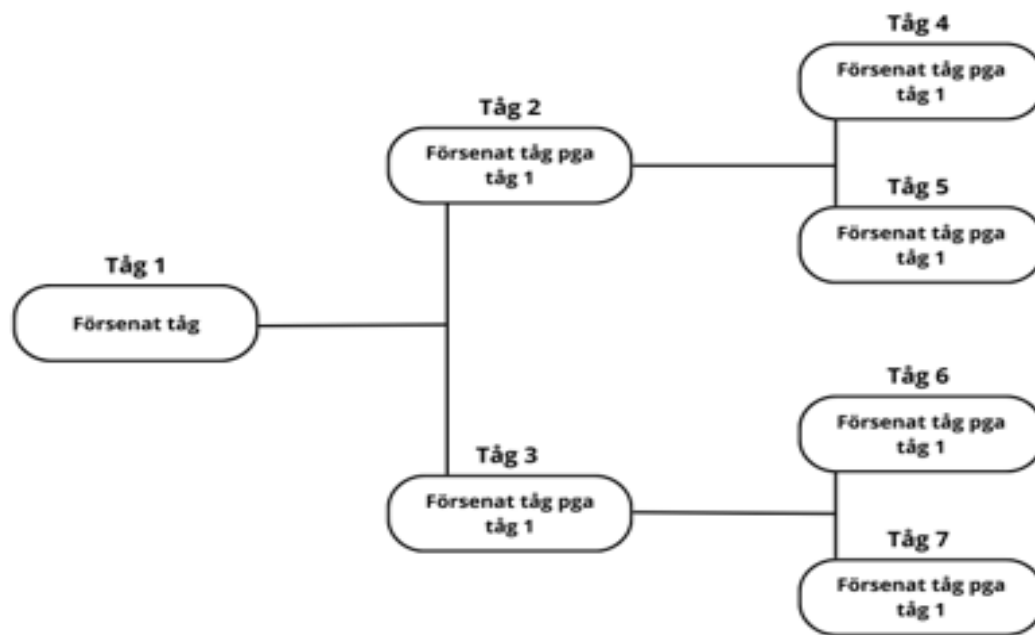


Bild 1 Schematisk bild över hur en försening av tåg 1 kan generera förseningar för bakomvarande tåg

En försening kan bero på oändligt många saker. Enligt Goverde (2005) kan man dela in förseningar i två kategorier, primära- och sekundära förseningar. Goverde menar att primära förseningar definieras av avvikelser från tidtabellen orsakade av variationer inom ett händelseförlopp. Ett exempel på en sådan process kan vara ett ovanligt högt antal av- och påstigande resenärer vid en station. Några andra exempel skulle kunna vara lokförarens körbeteende, om ett signalfel tvingar ett tåg till stopp, om tågen tvingas köra i reducerad hastighet på grund av banarbete eller om tåget inte kan avgå från en station i tid på grund av att en dörr inte går att stänga. Således är alltså primära förseningar sådana som direkt kan hänföras till det enskilt berörda tåget.

Vidare menar Goverde att sekundära förseningar är förseningar som genereras på grund av ett annat tåg, det vill säga händelser som inte direkt kan hänföras till det enskilt berörda tåget. Ett exempel på en sekundär försening är fallet då ett tåg hinner i kapp ett långsammare tåg. Ytterligare exempel på sekundära förseningar är om ett tåg tvingas sänka sin hastighet på grund av att ett tåg framför kör med reducerad hastighet, eller om tåget tvingas vänta på att få tillstånd att köra in till stationsplattformen på grund av att där redan står ett annat tåg. Vid rusnings- och

tät trafik kan en mindre primär försening generera stora effekter på andra tåg som befinner sig långt därifrån. Sekundära förseningar beror på oplanerade interaktioner mellan tåg. För att försöka minimera problematiken med sekundära förseningar och spridning av förseningar placeras bufferttider in i tidtabellerna. Denna bufferttid sätts in mellan två på varandra följande tåg. I *Bild 1* ovan skulle således buffertiderna placeras in mellan samtliga tåg 1-7.

4.1. Mål

Målet med examensarbetet är att undersöka och utvärdera ankomstprognoser för tåg. Detta görs genom att analysera de bakomliggande prediktionskällornas förmåga att skapa tillförlitliga ankomstprognoser. Det huvudsakliga fokuset är att öka kunskapen om såväl styrkor som svagheter hos prediktionskällorna, samt att se om det finns situationer där den ena eller andra ter sig särskilt lämplig.

4.2. Problemställning

- Vilket utvärderingsmått bör användas?
- Vilken prediktionskälla ger mest tillförlitlig ankomsttid?
- Vilken framförhållning är väl avvägd?

4.3. Avgränsningar

Det pågår ett forskningsprojekt, inom samma område, parallellt med detta examensarbete. Forskningsprojektet har en avsevärt längre tidsperiod och för att detta examensarbete skall hålla sig inom sin tidsram har avvikelser från forskningsprojektet behövt vidtas. Den data som analyseras i examensarbetet utgörs endast av en första preliminär datadump och således inte av data generad från en hel tremånadersperiod på en testbana, som forskningsprojektet gör.

Den data som använts är tillhandahållen av Trafikverket och avser datapunkter som är rikstäckande. Datan sträcker sig från norr till söder och innefattar datapunkter från samtliga 8 trafikledningscentraler: Malmö, Göteborg, Norrköping, Hallsberg, Stockholm, Gävle, Ånge och Boden. Datamängden avser 5 530 048 datapunkter som sträcker sig mellan datumen 2022-10-01 – 2022-11-15.

Detta examensarbete avser att utvärdera ankomstprognoser och kommer därför inte att belysa tågens avgångar. Vidare tas ingen hänsyn till hur stor eller liten förseningen är. Att en försening finns är illa nog.

5. Bakgrund

Ankomstprognoser är något alla resenärer är angelägna att ta del av, inte minst för att planera och få en guidning om när tågen beräknas ankomma. Eftersom ankomstprognoser lägger grunden till resenärernas förväntan av när tåget skall ankomma är det viktigt att denna prognos är tillförlitlig. Detta kapitel ämnar ge läsaren en kontext av ankomstprognoser genom att beskriva dessa ur olika perspektiv, samt en bakgrund kring problematiken i dagsläget.

5.1. Förseningar

Förseningar kan bedömas och tydas på olika vis. Ett sätt är, enligt Palmqvist (2019), att förseningar är skillnaden mellan tågens planerade ankomsttid och dess aktuella ankomst. Med andra ord är således förseningens storlek differensen mellan tiden som tåget enligt tidtabell ska ankomma och tiden som tåget faktiskt ankom. Enligt denna definition kan storleken av förseningen för ankomster vara såväl positiv som negativ. Detta gör att ett ankommande tåg som exempelvis fått värdet 7 indikerar att tåget ankommit 7 minuter för sent, och vice versa indikerar värdet -7 att tåget ankommit 7 minuter för tidigt.

Trafikverkets definition av en försening mäts som andelen tåg som ankommer senast 5 minuter och 59 sekunder efter angiven ankomsttid. Är förseningen större än detta skall en orsakskod anges och om förseningen är mindre utelämnas orsakskod helt. Detta mått gäller för såväl kort-, medel- och långdistanståg som för godståg. Inställda avgångar anses som separat händelse och är således inte en del av detta punktlighetsmått (Trafikverket, 2020b). I praktiken betyder denna definition att ett tåg som exempelvis ankommer 5 minuter sent noteras för en försening men med utebliven orsakskod. För passagerare kan även små förseningar leda till missade anslutande avgångar och till betydligt större totala förseningar (Parbo, Nielsen & Prator, 2016). Således krävs en försening på minst 6 minuter för att den skall rapporteras och anges med en orsakskod. Det finns fem olika orsakskategorier: infrastruktur, olyckor/tillbud, järnvägsföretag, följdorsaker och driftledning.

Palmqvist (2019) konstaterar i sin studie att de flesta förseningar är orsakade på grund av mindre störningar, upp till någon minut. Dessa mindre störningar sker ofta på eller i anslutning till driftplatser där tågen skall stanna, och sedan inte klarar av att köra vidare i tid. Exakt vad som är roten till störningarna är svårt att säga men ofta är uppehållstiden, den tid som tågen ska befinna sig på stationen, för kort. Små förseningar kan under en tågresor ackumuleras vilket således i slutändan kan utgöra en betydande roll för problematiken med opålitliga tidtabeller. Detta kan, som tidigare nämnts, till och med ske utan att tåget noteras med någon orsakskod, eftersom varje enskild försening understiger 5 minuter (Palmqvist, 2019). Palmqvists studie baseras på data på mer än 200 miljoner observationer över ca 7,5 miljoner resor på det svenska järnvägsnätet under år 2011-2017. Studien påvisade även att 78% av fördröjningstiden utgjordes av fördröjningshändelser som var 3 minuter eller mindre. Vidare konstaterades att 85% av fördröjningstimmarna utgjordes av fördröjningshändelser som var 5 minuter eller mindre. Således var det endast, 15%, som stod för händelser som noterades med en orsakskod till förseningen (Palmqvist, 2019). Punktligheten vid slutet av en resa är ett resultat av en lång serie med förseningar och återhämtningar, där de flesta är för små för att ges orsakskoder, även om tågen till slut mycket väl kan vara opunktliga.

5.2. Punktlighet

Punktlig het kan bestämmas på många olika sätt. För en resenär är det av stor vikt att tåget är punktligt i den formen att tåget ankommer i rätt tid, men för att tåget inte skall generera sekundära förseningar till andra tåg är det av stor vikt att tåget också avgår i rätt tid. På grund av detta kontrolleras såväl när tåget avgår från stationen som när det ankommer. Distinktionen av vad som anses vara acceptabelt för att benämnas som punktlighet varierar och beror till stor del bland annat på hur lång tågresa n är och vilken tågtyp som körs. Att på ett tydligt vis kunna beskriva punktligheten till ett flertal stationer i ett och samma punktlighetsmått är svårt. En resenär är sannolikt intresserad av hur punktligt tåget är till den stationen hen skall kliva av på. Tidigare har punktlighetsmättet avsett punktligheten mellan tågets start- och slutstation. Ett tåg gör oftast flertalet stopp och uppehåll från start till slutstation. För passagerare som stiger av på någon av mellanstationerna är intresset av om tåget var punktligt på slutstation obefintligt. Som tidigare nämnt är det svårt att på ett tydligt sätt kunna beskriva punktligheten till flera stationer i ett och samma mått.

Järnvägsnätet är ett integrerat och sammanlänkat system där många tåg samsas om en begränsad infrastruktur med ett specifikt antal spår. I dagsläget baseras tågresa n på tidtabeller och scheman som är sårbara för förseningar. Avvikelser från tidtabellerna orsakas av en mängd olika faktorer, allt från förutsägbara till oförutsägbara händelser. Exempelvis spelar väderrelaterade faktorer en stor roll. Antalet av- och påstigande resenärer är en annan faktor och listan kan göras lång (Lee et al. 2016). För att tillhandahålla en god servicenivå är det av största vikt att förutse dessa förseningar och skapa en punktlighet och träffsäkerhet gällande resenärernas avgångs- och ankomsttider. Enligt Palmqvist (2019), Yuan et al. (2002) & Huang, Z. Li et al. (2021) är fördröjningsutbredning en huvudorsak till avvikelser i järnvägsnätet. Detta är till största del på grund av att tågen är så starkt beroende av framförvarande tåg. En avvikelse på ett framförvarande tåg kan generera stora följavvikelser för många eftervarande tåg, vilket utgör ett komplext problem.

5.3. Prognoser för tågankomster

I dagsläget finns det olika sätt att skapa ankomstprognoser för tåg. Det mest triviala är genom manuella beräkningar, men dessa kan även göras baserat på automatiska beräkningar. De automatiska beräkningarna använder sig av digitala grafer och under de senaste åren har det blivit alltmer vanligt förekommande att dessa baseras på AI-lösningar, som exempelvis AIRT (Trafikverket, 2020a) och TrainBrain (Mattsson, 2021). Under de senaste årens vidareutveckling och förfining av tekniker inom storskaliga datamängder har forskning fokuserat på maskininlärningsalgoritmer och artificiell intelligens (AI) för att lära sig tyda allmänna regler och mönster i komplexa system, exempelvis järnvägssystem. Detta för att kunna leverera tillförlitlig kundservice i form av pålitliga tidtabeller som överensstämmer med när tågen faktiskt ankommer och avgår. Toppmoderna modeller för att prediktera förseningar kan delas in i två kategorier: deterministiska- och stokastiska prediktionsmodeller. De deterministiska prediktionsmodellerna som exempelvis kan bygga på neurala nätverk och ensembleinlärning, använder sig vanligen av tidigare observationer och data för att producera ett fast bästa uppskattningsvärde av förseningen eller ankomsttiden. De stokastiska förutsägelsemodellerna som exempelvis bayesianska nätverk, bygger i stället på en sannolikhetsfördelning som utgör grunden för prediktionen (Gorman, 2009). Baserat på vilken prediktionskälla som använts kan tågets beräknade ankomsttid opraktiskt nog variera. Hur tillförlitliga och exakta de olika prediktionskällorna är, och i vilka situationer den ena eller andra lämpar sig bäst finns det i dagsläget inga riktlinjer för. Dock är själva syftet med

ankomstberäkningen densamma oavsett prediktionskälla, nämligen att beskriva vid vilken tid som tåget kommer att anlända till, eller avgå från, en specifik station.

Prognoserna skapas för att underlätta för resenärer att planera sina behov utefter sin resa. Ankomstprognoser för tåg utgör en betydande och viktig information för såväl trafikledare som för resenärer att kunna ta del av. För att skapa predikterade och tillförlitliga prognoser som avser framtida förseningar krävs att ett stort antal parameter vägs in. Detta kan vara parametrar som exempelvis tid och plats, sträcka mellan stationerna, stationernas utformning, tågets tidigare försening samt andra angränsade tågs förseningar.

En prognos för ett tåg innebär i praktiken en prognos över flera punkter samtidigt, detta på grund av att ett tåg passerar flertalet mellanstationer på sin resa från start- till slutdestination. I en utvärdering av en prognos kan det uppstå komplikationer eftersom vissa driftplatser anses viktigare, och därmed i modellerna och algoritmerna skattas högre, än andra. Anledningen till detta är att tågen vid dessa, högre skattade driftplatser, är mer utsatta för störningar, exempelvis på grund av större och mer omfattande resenärutbyten vilket ökar riskerna för konflikter mellan tågen. Prognosens utfall baseras därav till stor del på hur denna skattning gjorts och det är därför väsentligt att det finns riktlinjer för hur skattningarna skall utföras. En annan faktor som påverkar och försvårar skattningen är att avståndet till och från de prognostiserade punkterna inte är konstanta, utan varierar. Exempelvis är det en sak att prognostisera när ett tåg från Malmö C anländer till Lund C, och en annan sak att prognostisera när samma tåg, från Malmö C, anländer till Stockholm C. Vissa prognospunkter ligger enstaka minuter framåt i tiden, medan andra ligger flera timmar och mil bort. Sannolikheten att ge en tillförlitlig prognos ökar med ett minskat avstånd till den prognostiserade driftplatsen. Det blir följaktligen statistiskt sett mer sannolikt att prognoserna för pendeltåg blir mer tillförlitliga än vad de blir för snabbtåg, eftersom avstånden är kortare. Dock kan det vara av större vikt att ha en mer tillförlitlig prognos för snabbtågen, med tanke på dess längre körsträcka. Prognoserna tillför mer om punktligheten är lägre, och resorna längre.

En och samma prognos står sig inte under en hel tågfärd. Efterhand som mer information blir känd uppdateras och finjusteras prognoserna för att vara aktuella och tillförlitliga. Exempelvis är det, som tidigare nämnts, avsevärt enklare att på ett mer pålitligt sätt prognostisera när ett tåg anländer till Stockholm C när det närmar sig Södertälje Syd än när det lämnade Malmö C. Med andra ord görs det flertalet upprepande prognoser över när tåget skall ankomma, med ett minskande antal driftplatser. Att enbart utvärdera och basera information på den första eller sista prognosen vore inte rättvist. Samtidigt finns det en poäng med att inte skapa förvirring för resenärer med överväldigande och osäker information. Den information som resenärer får sig tilldelad påverkar deras val och beteenden. Det finns därför en baktanke med att inte gå ut med information om att ett tåg kommer att vara exempelvis 1 timme försenat, för att sedan korrigera prognosen med att förseningen enbart blir 10 minuter. Detta eftersom det vore miserabelt om resenärer väljer att lämna stationen baserat på den första prognosen, och sedan missar tåget när prognosen ändras. Å andra sidan kan många resenärer känna en stor frustration när en prognos ständigt ändras och förlängs med ett visst antal minuter. Kanske vore det då bättre att redan från början säga att tåget blir 1 timme sent. Det är därför av stor vikt att beakta detta vid utvärdering av prognoser. En annan viktig aspekt att beakta är till vem prognosen och informationen är ämnad. En person sittandes ombord på ett tåg är förmodligen mest intresserad av när tåget som senast kommer att ankomma till stationen, för att kunna planera in vilka eventuella anslutande färdmedel som kommer att passa. Däremot kan en annan person som är ståendes på perrongen och inväntar samma tåg vara mer intresserad av när tåget som tidigast kommer att ankomma till stationen, för att inte missa avgången. Det råder alltså en asymmetri i prognoserna. Vidare behöver inte resenärer ta del av all information som genereras, utan enbart nödvändig information, eftersom det annars lätt kan bli rörigt. Därför är en annan viktig aspekt hur gränssnittet skall utformas för att särskilja på prognoser som skapas för internt bruk och de som skapas för att kommuniceras externt. Dock kan man inte skilja på vilken information som skall

visas för de olika resenärerna eftersom samtliga, oavsett om man befinner sig på tåget eller på perrongen, måste tilldelas samma information då det annars lätt kan uppstå konflikter.

Tidtabeller är inte enbart skapade för resenärer utan används även aktivt av flertalet andra yrkesgrupper, så som lokförare, planerare på järnvägsföretagen och trafikledare. Tidtabellerna för resenärer och lokförare ser snarlika ut men skiljer sig åt från övriga yrkesgrupper. Resenärernas och lokförarnas gränssnitt är med angivelser för när tågen kommer att avgå- och ankomma till de olika driftsplatserna. Den största skillnaden mellan resenärernas och lokförarnas tidtabeller är hur många driftplatser som tas med och visas. För en resenär är huvudsaken att kunna se när tåget avgår från en station och ankommer till en annan, samt avgångstider för eventuella anslutande byten. Lokföraren behöver däremot i sin tidtabell kunna se hålltider för avsevärt fler driftplatser. För trafikledarna är det helhetsbilden över alla tågen som vistas på banan som är av vikt. För att trafikledarna inte skall överväldigas av siffor använder de sig i stället av grafiska tidtabeller, mer och detta senare. Tidtabellerna för planerarna på järnvägsföretagen är skapade för att generera information kring omlopp för tågen, vilka dagar tågen skall gå, vart och när dem skall gå och ankomsttider. Detta är av vikt för att tågpersonalen ska kunna ta sig tillbaka till sin utgångsplats vid avslutat arbetspass, samt för att veta vart tågen hamnar. Tidtabell gäller för ett specifikt och unikt tågnummer under hela årets tågplan, vilken kommer att presenteras vidare längre fram i rapporten.

6. Teknisk bakgrund

För att ett järnvägssystem skall fungera med tillförlitliga tågprognoser krävs att flertalet delsystem samarbetar. Detta kapitel startar med en litteraturstudie för att belysa tidigare forskning inom liknande projekt och därmed introducera specifika modeller som tidigare använts, bland annat artificiell intelligens och maskininlärning. Vidare innehåller detta kapitel en teoretisk bakgrund om tågplanen, vilken sedan följs upp och avslutas med en förklaring av både grafisk- och digital tidtabell.

6.1. Litteraturstudie om tidigare forskning

Det har tidigare forskats och gjorts studier inom prediktioner för tågförseningar och tidtabeller. Studiernas slutsatser påvisar goda resultat där man lyckats precisera och på ett bättre, mer tillförlitligt sätt, ange tidsrymden av den predikterade förseningen.

6.1.1. Artificiell intelligens

Så sent som i mars år 2023 slutfördes ett projekt kallat AIRT, AI-baserad Realtidsprognostisering av Trafikinformation. I detta projekt användes Artificiell Intelligens för att analysera tågdata. Själva syftet med projektet var att stärka förmågan att kunna förutse tågstörningarnas storlek för att på så sätt förbättra trafikinformationen. Ytterligare gjordes inom projektet ny forskning genom att väga in väderinformation i prognoserna. Tanken här var att skapa förbättrade mer korrekta prognoser genom att inkludera parametrar som vädrets påverkan har på infrastrukturen. AIRT utgör en av de 5 prediktionskällorna som ligger till grund för ankomstprognoserna som utreds i detta examensarbete. Mycket information har därför kunnat införskaffas tack vare detta projekt (Trafikverket, 2023).

Yaghini et al. (2013) genomförde med hjälp av ett artificiellt neuralt nätverk (ANN-nätverk) en undersökning av hur väl ANN-nätverket förutspådde tågförseningar hos passagerartåg. Studien påvisade stor noggrannhet och låg inlärningstid. Yaghini et al, använde sig av multinominala logistiska regressionsmodeller för att utvärdera kvaliteten på resultaten med hjälp av ANN. Marković et al. (2015) byggde sedan vidare på Yaghini et al. (2013) om förutsägelser av förseningar för passagerartåg och presenterade utfall som påvisade att tillämpningen av stödvektorregression (SVR) uppnådde bättre resultat än ANN. För att kunna uppskatta vilka grundorsaker och effekter som förseningar kan leda till, samt vilka orsaker en handling har på att flera andra händelser inträffar efter varandra, konstruerades ett beslutsträd (Lee et al, 2016).

6.1.2. Maskininlärning

Larsen (2022) studerade och jämförde olika maskininlärningsmodeller med syftet att förutsäga förseningar av tågens avgångar, vilket gjordes från Oslo till Trondheim med tillhörande mellanstationer. Larsen använde för detta tre olika datamodeller, RNN, LSTM och GRU. Genom historiska tågdata försökte förseningarnas storlek att förutsägas. Skjøren (2022) undersökte med hjälp av historiskdata prestandan av den modell för förutsägelse av tågförseningar som används på de norska järnvägarna. Skjøren fokuserade studien på tre bansträckor i Oslo-området. Här analyserades och jämfördes skillnaden mellan den uppskattade och den faktiska förseningen. För att utvärdera precisionen av modellen använde sig Skjøren av en Pearson-korrelationskoefficient, vilken granskar förhållandet mellan den faktiska förseningen och den absoluta uppskattningsavvikelsen, det vill säga skillnaden mellan beräknad

och planerad avgång. Skjøren använde även utvärderingsmått Mean Absolute Error (MAE) och Root Mean Squared Error (RMSE) för att mäta prediktionernas prestanda och för att kunna jämföra med andra studiers resultat.

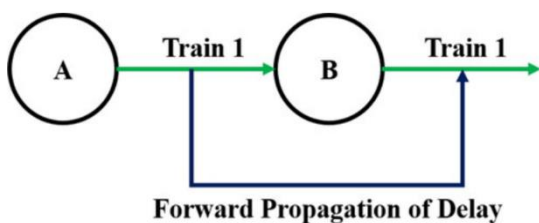
Oneto et al. (2017) skapade ett Dynamiskt tågfördröjningssystem (DTDPS-system) baserat på Shallow Extreme Learning Machine (SELM) och Deep Extreme Learning Machine (DELM) som använde en upprepad inlärningsprocess med historiska tågrörelsedata och väderdata för långtidsförutsägelse av tågförseningar i ett storskaligt järnvägsnätverk. Oneto et al. (2020) skapade en hybridmetod för att slå samman dessa kunskaper tillsammans med operatörernas erfarenhet och datadrivna modeller baserade på driftsdata och information. Detta gjordes för att förutsäga körtid, uppehållstid, tågförseningar, straffavgifter och tågomkörningar.

I en studie gjorde Spanninger et al. (2022) en överblick av olika metoder för att förutsäga tågförseningar. I deras studie dras slutsatsen att datadrivna maskininlärningsmetoder som exempelvis Neurala nätverk och Random Forest uppnår hög nivå och goda resultat vad gäller noggrannhet i förutsägelse. Dessa metoder har blivit mycket populära eftersom det i järnvägssystemen finns en stor mängd data att tillgå. Tillgången till storskaliga data medför även att modellkalibreringen och testprestandan förbättras. I studien påpekas risken för överanpassning, det vill säga, risken att maskininlärningsmetoden ger korrekta förutsägelser för träningsdata men inte för den ”nya” datan.

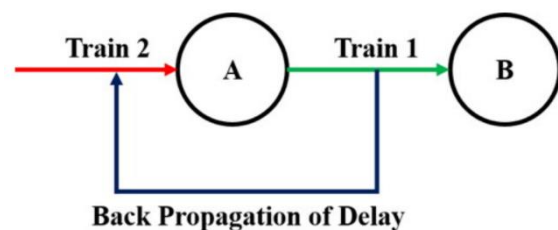
Med hjälp av järnvägsdata och de tre maskininlärningsmetoderna: Random Forest, ANN och linjär regression studerade Kosolsombat och Limprasert (2017) ankomsttider för tåg. I deras resultat presenteras vikten av funktionernas betydelse. De kom nämligen fram till att exempelvis olika dagar, veckor och stationer har olika inverkan på den förväntade förseningen. Milinković et al. (2013) föreslog en Fuzzy Petri Net (FPN) modell för att simulera primära förseningar i ett järnvägssystem. I denna studie användes FPN-modellen med bland annat egenskaperna hierarki och tid för att simulera trafikprocesser och tågrörelser i ett järnvägssystem.

6.1.3. Former av försening

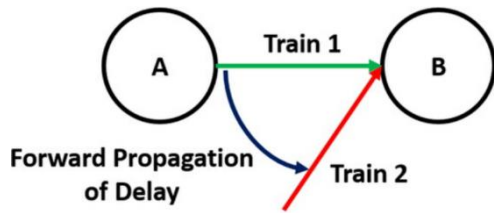
Förseningar kan mätas i olika former. En form kallas reaktionär försening vilken av Taleongpong et al. (2022) definitionenligt är en försening som genererats på grund av att ett annat tidigare tåg varit försenat. Med andra ord är en reaktionär försening någonting som skapats på grund av att en incident primärt gjort att framförvarande tåg blivit försenat och således sekundärt påverkat att de bakomvarande tågen blivit försenade. Ett intressant inslag i studien är de fem olika mekanismerna som kan leda till tågförseningar. *Figur 1-5* nedan visar de olika mekanismerna.



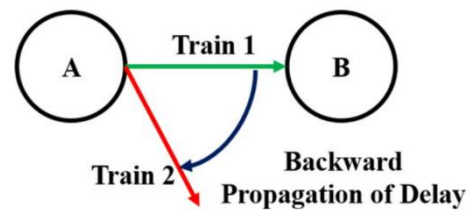
Figur1 Förseningsmekanism 1 - självförökning



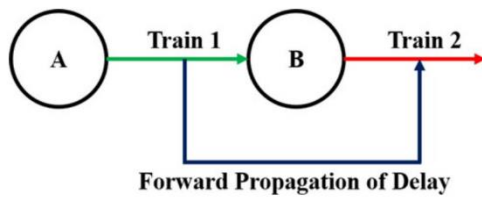
Figur 2 Förseningsmekanism 2 - ankomst bakåtförseningsutbredning



Figur 1 Förseningsmekanism 3 - ankomst framåtförseningsutbredning



Figur 2 Förseningsmekanism 4 - avgång bakåtförseningsutbredning



Figur 3 Förseningsmekanism 5 - avgång framåtförseningsutbredning

Förseningsutbredningen fokuserar på hur ett exempeltåg, tåg 1, som färdas från station A till station B fortplantar sig till ett annat tåg, tåg 2. Med undantag för förseningsutbredningsmekanism 1 i figur 1 vilket enbart avser tåg 1.

Förseningsmekanism 1 – självförökning. Detta sker när tåg 1 lämnar station A med en försening vilket leder till att tåg 1 även kommer att vara försenat när det lämnar station B, det blir således en fortplantning av förseningen. Se *figur 1*.

Förseningsmekanism 2 – ankomst bakåtfördröjningsutbredning. Detta sker när tåg 1 lämnar station A med en försening vilket leder till att tåg 2 ankommer till station A med en försening. Se *figur 2*.

Förseningsmekanism 3 – ankomst framåtförseningsutbredning handlar om att tåg 1 ankommer till station B med en försening, vilket medför att tåg 2 också blir försenat till station B. Se *figur 3*.

Förseningsmekanism 4 – avgång bakåtförseningsutbredning. Tåg 1 avgår från station A med en försening vilket även resulterar i en avgångsförsening för tåg 2 från station A. Se *figur 4*.

Förseningsmekanism 5 – avgång framåtförseningsutbredning. Tåg 1 ankommer med en försening till station B vilket resulterar i att tåg 2 avgår med en försening från station B. Se *figur 5*.

7. Tågplanen

Allt fler svenskar väljer att resa med tåg och under de senaste 25 åren har persontrafiken ökat med 70%. Det är förändrade resvanor, ökande arbetspendling och miljömedvetenhet som gjort att intresset för resor med tåg nu är större än någonsin (Trafikverket, 2016). Denna positiva utveckling ställer samtidigt krav på järnvägen, där det redan blivit trångt på spåren. Den svenska järnvägen har hamnat i en situation där kapaciteten inte längre räcker till för att bemöta den stora efterfrågan, som ständigt expanderar. Samtidigt som järnvägsföretagen vill öka och köra fler tåg måste den redan eftersatta och slitna järnvägen rustas upp och underhållas under många år framöver. Ytterligare en stor utmaning med den ökande efterfrågan av tågtrafik, både vad gäller gods- och persontrafik, är att tågen behöver samsas om en järnväg som till två tredjedelar består av enkelspår och de flaskhalsar som enkelspåren medför. Järnvägsnätet är stort och det är många olika aktörer som vill använda det.

Det är Trafikverket som genom en så kallad tågplan ansvarar för att bemöta den ökande efterfrågan av utnyttjande av tider i spåren. Tågnätet är ett begränsat system och då det råder tät trafik påverkas specifika tågs ankomsttider potentiellt sett av flertalet andra tåg. Ett stort problem är, som tidigare nämnts, att tågen enbart är begränsade till de enkel- och dubbelspår som finns, hur stationerna är utformade och hur förbindelserna till dessa ser ut. För att skapa tillförlitliga ankomstprediktioner hos ett specifikt tåg krävas att flertalet andra tågs förväntade ankomsttider beaktas och vägs in. Detta eftersom trafikstörande händelser hos andra tåg, på angränsande linjer, kan påverka ankomsttiden hos det specifika tåget.

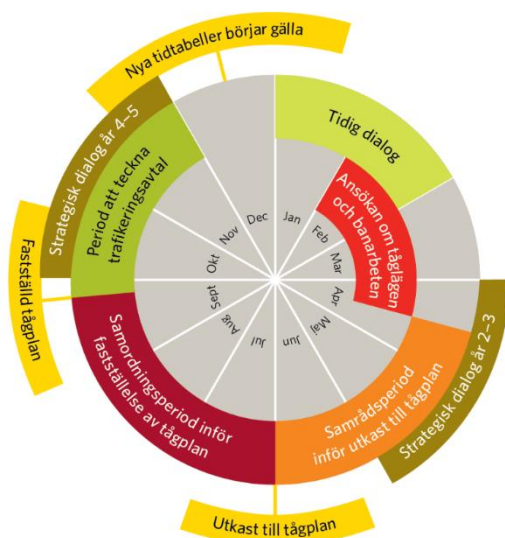
Tågplanen beskriver kapacitetsutnyttjandet av statens spåranslagning och gäller till och med ett år från startdatum. Startdatum är detsamma varje år, andra lördagen i december. Syftet med tågplanen är att se till att trafiken med de cirka 4000 gods- och persontåg som varje dag färdas på järnvägen flyter på utan konflikter. Det är även i tågplanen som de inplanerade tiderna för bland annat banarbeten planeras in. När tågplanen konstrueras och pusslas ihop måste därför konstruktörerna även se till att alla de cirka 2200 banarbeten som varje år utförs får tider i tidtabellen. För att detta ska fungera på ett vedertaget sätt måste tidtabellen vara strukturerat uppbyggd. Att skapa en tidtabell för ett enskilt tåg är ingen svår process men att få ihop över 4000 tåg som samtidigt varje dag trafikerar spåren utgör ett komplext system (Trafikverket, 2016).

Att skapa en fungerande tågplan är ett digert arbete och tar cirka ett år. Processen startar med att Trafikverket tar fram en järnvägsnätsbeskrivning där förutsättningar för att bedriva tågtrafik, och ansökan om att få använda kapacitet i järnvägsnätet redogörs. I början av varje år bjuder Trafikverket in alla tågbolag till ett möte för en inledande dialog. Därefter, från februari till april, kan alla bolag ansöka om tider för tåglägen och banarbeten till den kommande tågplanen. Efter ansökningsperiodens slut bjuder Trafikverket in de bolag som gjort en ansökan, till samrådsmöten. På dessa möten görs kompromisser för att försöka tillmötesgå företagens olika önskemål. I järnvägsbranschen finns det en vedertagen medvetenhet om att man möjligen kommer att få andra tågtider och tåglägen än de man ansökt. Detta eftersom det är många som vill använda och trafikera spåren. Med tanke på att man måste ansöka om tågtider cirka 1 år i förväg finns det även tider som ansöks av strategiska skäl, som säkerhet, i form av ”back-up” tider. När ansökningsperioden är stängd läggs alla ansökningar in i ett system och det stora konstruktionsarbetet med tidtabellen kan då påbörjas. Konstruktionsarbetet behandlar bland annat konfliktreglering för tågtider som ansökts och ligger omlott, på samma tid. Två tåg får inte vara för tätt inpå varandra och kan inte befinna sig på samma plats samtidigt. Om så är fallet reglerar konstruktörerna tågtiderna genom att flytta tågens tid i tidtabellsgrafnen, vilket medför att starttiden flyttas fram eller bak för att passa in. Målet är att skapa en hållbar tågplan med så många nöjda aktörer som möjligt.

En del sträckor och banor är enklare att skapa tidtabeller för än andra. Vissa sträckor kan nämligen trafikeras av fler tåg och tågtyper av olika karaktär. Exempelvis är en del sträckor

trafikerade av såväl lokaltåg, regionaltåg, snabbtåg, fjärrtåg som godståg, vilka både har olika hastighetsprofiler och egenskaper. Samtidigt finns det andra sträckor som enbart trafikeras av lokaltåg. Detta medför ett mycket komplext pussel för konstruktörerna att lägga för att undvika konflikter mellan de olika tågen och dess olika hastighetsprofiler samtidigt som de större underhållsarbeten också planeras in. Olika tågtyper som trafikerar en linje stannar på olika många stationer på sin färd mot slutdestinationen. För snabbtågen utgör majoriteten av stationerna längs med sträckan endast passager utan något stopp, medan regionaltågen har betydligt fler uppehåll.

Efter ytterligare samrådsmöten med aktörerna kan de sista pusselbitarna läggas. Under sommaren skickas det första utkastet om den nya tågplanen ut till alla ansökande. Det är därefter möjligt för bolagen att lämna in synpunkter på eventuella missnöjen vilka behandlas och går genom innan sacken knyts ihop, och tågplanen fastställs i slutet på september. Om flera tågbolag ansöker om att framföra tåg samtidigt uppstår en intressekonflikt (Trafikverket, 2016). Enligt 6 kap. 9 § i järnvägslagen (*Tilldelning av infrastrukturkapacitet och tillhandahållande av tjänster – Kapacitetsplanering*) står följande ”Infrastrukturförvaltaren skall i så stor utsträckning som möjligt tillmötesgå alla ansökningar om infrastrukturkapacitet samt beakta den ekonomiska effekten på sökandenas verksamhet och andra förhållanden av betydelse för sökandena.” (Järnvägslagen, 2004). Med andra ord skall ansökningarna alltså i första hand tillgodoses. I fallet då denna konflikt inte går att lösa under samordningsperioden tillämpas särskilda förutbestämda prioriteringskriterier och beräkningsverktyg för att avgöra vilka tåg som får företräde i tågplanen. Denna beräkningsmodell utgår från och använder prioriteringskriterier för att finna den lösning som ger störst samhällsekonomisk nytta. Tanken här är inte att ställa enskilda tåg mot varandra utan i stället ge en helhetsbedömning. För att säkerställa robustheten i tidplanen på de mest krävande och kapacitetstighta banorna simuleras dessa separat. Om ett problem skulle uppstå måste konstruktörerna åtgärda detta och på nytt simulera den nya tidplanen. Detta sker i en iterativ process som slutningen förhoppningsvis mynnar ut i en mer robust tidtabell. Den nya tidtabellen för tågen och tidplanen för banarbeten träder i laga kraft andra lördagen i december, och gäller sedan i ett år framöver (Trafikverket, 2016). En schematisk bild över planeringsprocessen kring tågplanen syns i *Figur 6* nedan.

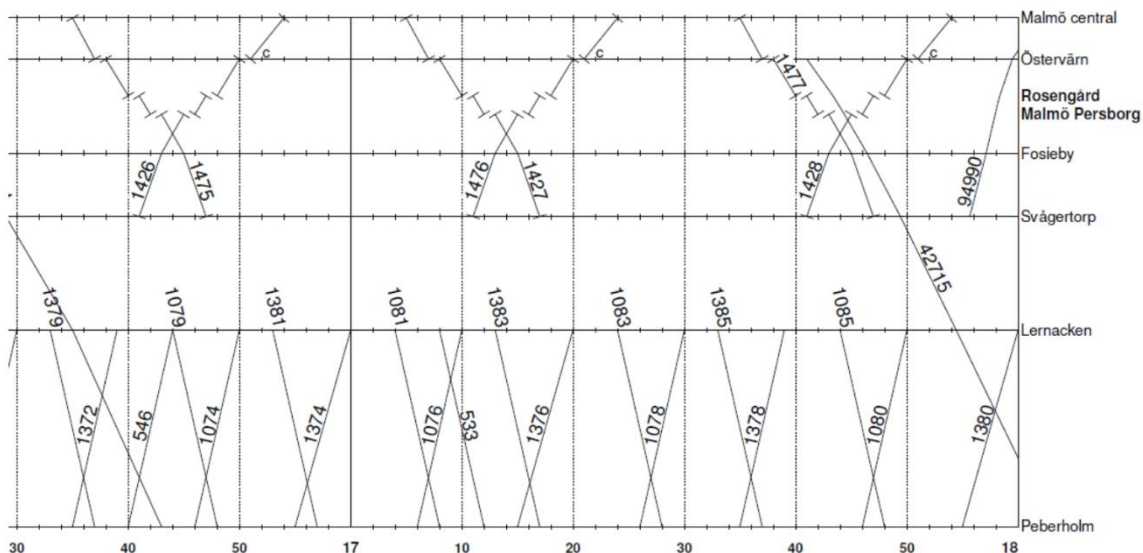


Figur 4 Den svenska planeringsprocessen av tågplanen (Trafikverket, 2016)

8. Grafiska tidtabeller

I tågtidtabellernas begynnelse fram tills för något decennium sedan sköttes all tågplanering med hjälp av penna och papper. Då var det en tågkларerare som manuellt ritade in samtliga tåg som skulle trafikera sträckan, för att sedan pussla ihop en fungerande tidtabell, på en pappersgraf. I dagsläget görs fortfarande en del prognoser manuellt. Dessa prognoser är skapade av trafikledare som sitter på trafikcentralerna runt om i landet och bygger på deras erfarenheter kring olika situationer som sker på järnvägen. Här spelar således den mänskliga faktorn uteslutande roll i hur väl de predikterade framtida prognosernas utfall stämmer överens med verkligheten.

Det finns många olika sätt att skapa och beskriva en tidtabell på. En form är att framföra den i tabellformat, men för att på ett enklare och mer överskådligt sätt få en helhetsbild över flera tåg på en längre sträcka är det mer praktiskt att illustrera dem grafiskt. De grafiska tidtabellerna är uppbyggda i ett xy-diagram där x-axeln visar tiden och y-axeln visar driftplatserna (Lindner, 2000; Andersson, 2014). I *Figur 7* nedan visas ett typexempel på hur en grafisk tidtabell kan se ut.



Figur 5 Exempel på hur grafisk tidtabell kan se ut. Trafikverkets dagliga grafer (Trafikverket 2020).

De fyra- och femsiffriga numren i grafen ovan definieras av vilka tåg, tågets id, som trafikerar spåren. De diagonala linjerna indikerar var tågen är planerade att befinna sig vid olika tidpunkter. I grafen visas alla de tåg som är tänkta att trafikera en specifik bana under tidtabellens giltighetsperiod, i detta fall visas 16:30-18:00. Vidare kan man i grafen utläsa att varje tåg motsvarar ett streck, där lutningen på strecket avser hur fort tåget skall köra. En brantare linje medför således att tåget håller högre hastighet och vice versa (Andersson, 2014). En brist i de grafiska tidtabellerna är att det inte framgår om banan har dubbel- eller enkelspår, utan det är upp till konstruktören själv att hålla reda på det. I fallet med figuren ovan kan man exempelvis mellan Peberholm och Lernacken anta att det är dubbelspår eftersom tågen annars varken hade kunnat mötas eller köra förbi varandra ute på linjen, sträckan, mellan driftplatserna. Då linjerna i grafen inte längre är diagonala utan övergår till att vara horisontella indikerar detta på att tågen har uppehåll och står stilla.

Gångtiden för ett tåg definieras som den tid det tar för ett tåg att köra från en punkt till en annan, exempelvis mellan två driftplatser. När en svensk tidtabell konstrueras beräknas den minimala

gångtiden mellan två driftplatser baserat på bansträckningens karaktäristiska egenskaper och utformning, schablonmässiga värden samt vilken typ av tåg som är ämnade att trafikera banan (Caimi, et al., 2016). I exempelvis Storbritannien gör man på annat vis vid beräkning av gångtiden. Där använder man i stället historiska data över hur lång tid det tagit för tågen att ta sig fram mellan olika driftplatser (Palmqvist, 2019).

Uppehållstid är den tid ett tåg står inne på en driftplats. Dessa uppehållstider följer ingen vedertagen praxis utan är baserade på tumregler, vilka sällan är jämförda och kontrollerade med data över hur det tidsmässigt ser ut i verkligheten (Palmqvist, 2019). När ansökningarna om tid i spår görs, lägger järnvägsföretagen själva in hur långa uppehållstider de har för avsikt att följa. Då det inte finns någon vedertagen praxis att följa kan det av konstruktören opraktiskt nog planeras in onödigt stora eller små tidstillägg för uppehållstiden. Järnvägsföretagen betalar för den tid de trafikerar spåren och det kan därför ekonomiskt vara fördelaktigt för dem att lägga in korta uppehållstider. Detta medför således att tidtabellkonstruktören skapar en tidtabell baserat på uppehållstider vilka järnvägsföretagen själva har fyllt i vid ansökan (Hellström, 2014).

För att i tidtabellen, bland annat, ta hänsyn till att alla tåg inte har samma prestanda, lokförare har olika körbeteende och att det råder olika väderförhållanden, lägger man in en viss procent med extra tidstillägg (Hoogheimstra & Teunisse, 1998). Palmqvist (2019) skriver i sin rapport att själva storleken på tidstillägget varierar i olika länder, men att det i Europa är vanligt att ha ett tidstillägg på 3-7% av den minimala gångtiden. I Nordamerika är tidstilläggen lite större. Där använder man praxis om 6-8% tidstillägg (Pachl, 2018). Jensen (2015) skriver att det på en del sträckor till och med kan förekomma tidstillägg som är så mycket som upp emot 13% av den minimala gångtiden. Utöver detta förespråkar internationella järnvägsförbundet, UIC, att ytterligare ett tidstillägg på 1,5 minuter skall råda för var 10:e mil som tåget färdas (Palmqvist, 2019). Den stora anledningen till att införa tidstillägg i tidtabellerna är att minimera risken för förseningar. Å andra sidan får man på grund av detta, ökade körtider och högre operativa kostnader. För att slippa onödigt långa körtider och dyra operativa kostnader krävs därför att tidstilläggen är korrekt anpassat utefter bandelarnas förhållanden (Jensen, 2015).

Ett annat viktigt begrepp vid skapande av tidtabeller är Headway-tid. Med detta avses tidsdifferensen mellan två på varandra följande tåg (Hansen, 2010). Headway-tidens minimivärde blir således, den tid som krävs för att det efterföljande tåget skall kunna köra obehindrat utan att hinna ifatt framförvarande tåg (Klabes, 2010). Detta styrs till stor del av avståndet mellan två på varandra följande signaler, blocksträcka, och tågets hastighet (Jensen, 2015). Inom järnvägsteknik är en blocksträcka en sträcka som med hjälp av spårledning och signaler garanterar att enbart ett tåg i taget kan belägga blocksträckan. Ett tåg som befinner sig på en blocksträcka blockerar denna del av banan för alla andra tåg. Det är först när detta tåg passerar ut ur blocksträckan som ett annat tåg kan köra in på denna del av banan. Med andra ord är en blocksträcka en del av linjen som avgränsas av två signaler och som av säkerhetsskäl endast ett tåg i taget får befinna sig på. Ju längre avståndet är mellan blocksträckorna och ju högre hastigheten är desto längre headway-tid krävs. Där med sagt kan man inte enbart minska blocksträckorna för att minska headway-tiden eftersom man då tummar på kapaciteten. Blocksträckorna brukar normalt vara 1-3 km långa och beror på vilken största tillåtna hastighet, STH, som råder på banan (Khoshniyat & Peterson, 2017).

9. Digital graf

Digital graf är ett datorprogram som Trafikverket använder för att bedriva operativ planering digitalt. Genom en vidareutveckling av de manuella grafiska tidtabellerna underlättar Digital Graf för trafikledningen att planera in och skapa nya ankomst- och avgångstider för tågen (Angel, 2020). Programmet bygger på digitala grafer och beräkningar som kan meddela vilka spår som är mest lämpliga att använda för att tågen ska ta sig fram för att undvika trubbel. Digital graf effektiviserar arbetet genom att göra beräkningar på alla tåg samtidigt, medan en människa enbart kan fokusera på en beräkning åt gången (Angel, 2020).

För att fortsätta utvecklingen och optimera tid och plats på den redan kapacitetsfyllda järnvägen har ett projekt kallat C-DAS, Connected - Driver Advisory System, tagits fram. Projektet har initierats och drivits fram genom ett branschsamarbete inom TTT, Tillsammans för Tåg i Tid. C-DAS är ett system som i realtid sammankopplar trafikledningen, fjärrtågklarare, med lokföraren i en konstant uppdaterad tidplan. Ombord på varje lok finns en förarplatta där C-DAS är installerad. Detta gör att alla tåg via applikationen kan kommunicera ut sin position och hastighet.

På trafikledningscentralen arbetar man med den operativa planen för dagens tåg i ett program kallat STEG (Styrning av Tågtrafik via Elektronisk Graf). När ett tåg är redo för avfärd startas C-DAS applikationen på förarplattan i loket. C-DAS applikationen beräknar den rekommenderade hastigheten som lokföraren skall hålla för att komma i tid till de planerade uppehållen. Fjärrtågklararen får informationen om tågets egenskaper och prestanda skickad till sin digitala graf. Utefter denna information kan planering för tågföringen göras efter de faktiska omständigheterna som råder ute på banan. C-DAS har sin utgångspunkt i den operativa tidplan som tågklararen skapat i den digitala grafen i STEG. Därefter beräknar C-DAS körtiden med hjälp av informationen om tågets prestanda och banans STH (största tillåtna hastighet). C-DAS gör att informationen mellan fjärrtågklararen, på ledningscentralen, och lokföraren blir omedelbar och tillgänglig för de som behöver ta del av den. Lokföraren kan även se om det tillkommer, bortfaller eller planeras om tågmöten. Under tågets färd får trafikledningen ständigt uppdaterad information om tågets position och hastighet.

Om allt går som planerat tar tåget sig fram med god marginal enligt tidtabellen. Den eventuella extra tid som finns i tidtabellen kan användas för att sänka farten, så att energi sparas, samt till att ge resenärerna bättre komfort. I fallet då det inträffar en incident, exempelvis om ett tåg framför drabbas av ett tekniskt fel, kan tidtabellen och den operativa planen som lokföraren följer behöva justeras utefter de nya förutsättningarna. Direkt när planen är anpassad skickas de nya tiderna från trafikledningen till C-DAS som då kan beräkna en ny rekommendation åt lokföraren. Baserat på uppdateringen kan även trafikinformationen uppdateras så att resenärerna som står och väntar på stationen får information om den eventuella förseningen (Trafikverket, 2022b). I *Bild 2* och *Bild 3* nedan visas schematiskt hur C-DAS fungerar.

Nuläge

Utmaningar



Bild 2 Principskiss över nuläget för trafikledare, förare, resenär och godskund. Första testet av C-DAS tillsammans med SJ genomfört (Trafikverket 2020).

C-DAS

Connected Driver Advisory System

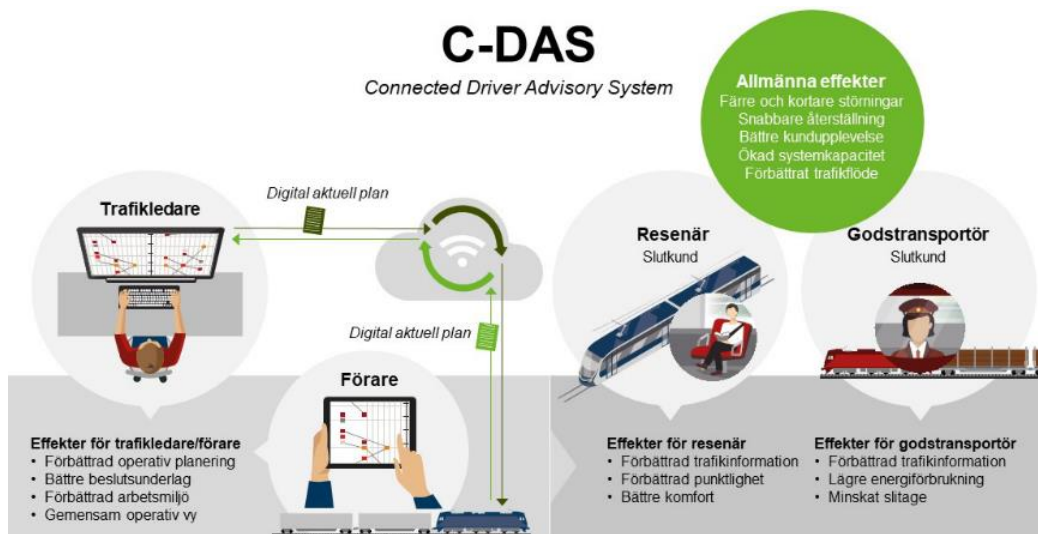


Bild 3 Principskiss som visar fördelar med C-DAS. Första testet av C-DAS tillsammans med SJ genomfört (Trafikverket 2020).

10. Prediktionskällor

För att skapa predikterade ankomstprognoser finns det i dagsläget 5 olika tillvägssångsätt. Tillvägagångssätten beror på vilken av de 5 olika bakomliggande prediktionskällorna som används och opraktiskt nog kommer prognoser att ha varierande resultat baserat på vilket av alternativen som väljs. De olika prediktionskällorna presenteras noggrannare nedan.

10.1. STEG

STEG är som tidigare nämnts det digitala graf-verktyg som Trafikverket använder sig av vid operativ planering av tidtabeller.

10.2. Amber

Amber-Amber och Amber-Stå är två automatiska prognoser för pendeltåg skapade i systemet Amber, vilka baseras på gångtider.

Amber-Amber är automatiska prognoser för pendeltåg baserade på försening eller beräkningsrapporter. Exempelvis om ett tåg är x-minuter sent till Sundbyberg räknar systemet ut när det tidigast kan vara på efterföljande stationer.

Amber-Stå fungerar som Amber-Amber men baseras på att ett tåg inte har ankommit eller avgått. Amber-Stå är alltså automatiska prognoser för pendeltåg baserat på utebliven rapport. Exempelvis om ett tåg inte har avgått i tid från Sundbyberg, så beräknar Amber-Stå när tåget som tidigast kan vara framme vid efterföljande station.

10.3. Tågprognos

Tågprognos är en automatisk prognos som Trafikverket har på prov i deras tågprognos-tjänst och används inte i nuläget i fält utan är enbart som ett test.

10.4. Manuell

Manuell innebär att det är en manuellt skapad prognos, alltså ingen automatiskt skapad prognos.

10.5. AIRT

AIRT är en algoritm som genom AI beräknar och skapar prediktiva prognoser för tågens ankomsttider. De fyra andra, ovan nämnda, prediktionskällorna använder sig inte av maskininlärning för att generera prognoser, men detta är precis vad AIRT gör. Vid användning av historiska data och maskininlärning går det att skapa ankomstprognoser som är automatiserade och enbart baserar sig på att känna igen mönster kring hur situationer historiskt sett ut i likartade lägen.

11. Utvärderingsmått

För att kunna utvärdera prognoserna är en central del att datan är väl sorterad. Den insamlade datan måste bearbetas och struktureras för att kunna bli analyserad. I vissa fall kan felaktigheter i datan behöva anpassas för att göra den användbar till prognosmodeller. Det skall dock tilläggas att datan aldrig ändras eller på något sätt förändras. Vid analys av historiska data kan trender och mönster identifieras. Analysen kan innefatta att syna tidigare förseningar, jämföra tidtabellsdata med faktiska ankomster- och avgångar för att upptäcka eventuella förseningar, genomsnittliga restider, vanliga resvägar eller andra faktorer som påverkar trafiken på järnvägen, exempelvis underhållsarbete.

När en prognos har skapats kan den justeras och uppdateras med hjälp av realtidsdata. Det kan inkludera information om eventuella förseningar, avvikelser eller trafikstörningar som uppstår under resans gång. Dessa uppdateringar kan sedan användas för att ge aktuell information till resenärerna. Det är sedan baserat på den insamlade och analyserade datan som prognosmodeller används för att förutspå ankomsterna. Modellerna kan vara statistiskt baserade men kan även innefatta användning av maskininlärningstekniker skapade för att finna mönster och samband i datan. Modellerna och prognosalgoritmerna bör utformas och kalibreras med äldre data, tidigare kända faktorer som erkänt påverkat trafiken. Genom att använda detta kan man få insikt i vilka faktorer som faktiskt påverkar och skapar förseningar, för att sedan kunna förbättra och finjustera prognosernas tillförlitlighet och precision. När prognoserna skapats görs utvärderingar och jämförelser med faktiska data för att kunna bedöma prognosernas tillförlitlighet. Då det råder avvikelser eller rent av felaktigheter utförs ändringar i modellerna och insamlingen av datan för att prognoserna framöver ska kunna förbättras.

För att kunna utvärdera tillförlitligheten i de olika prediktionskällorna: STEG, Amber, Tågprognos och Manuell, används i detta arbete tre vanligt förekommande statistiska utvärderingsmått. Modellernas matematiska formler visas och förklaras nedan.

11.1. Medelvärde

Det första och enklaste utvärderingsmått är medelvärdet, vilket är detsamma som medelfelet. Detta är en statistisk mätning som används för att beräkna genomsnittet av en serie numeriska värden, och representerar det mest centrala värdet eller det mest typiska värdet i en datamängd. Vid ankomstprediktioner för tåg utgör medelvärdet hur stort det genomsnittliga predikterade felet är. För att beräkna medelvärdet adderar man alla värden i datamängden och dividerar sedan summan med antalet värden. Med andra ord definieras formeln som summan av alla fel dividerat med antalet fel. x_1-x_n är felen och n är antalet fel. Se formeln nedan:

$$\text{Medelvärdet} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Medelvärdet ger en indikation på den centrala tendensen eller genomsnittliga nivån av de observerade värdena. Detta kan vara användbart för att få en översiktlig uppfattning om datamängden och för att jämföra olika variabler. En nackdel med medelvärdet är att det dessvärre kan påverkas av extremvärden i datamängden, och kan därför inte alltid ge en fullständig bild av fördelningen av värdena. Viktigt att notera är även att medelvärdet är känsligt för förändringar i datamängden och kan påverkas av avvikande värden.

11.2. Mean absolute error (MAE)

MAE, Mean Absolute Error, är ett vanligt mått på prognosfel eller felbedömning i statistik och maskininlärning. Det används för att mäta genomsnittlig absolut avvikelse mellan de förutsagda värdena och de faktiska värdena i en datamängd. För att beräkna MAE jämförs varje förutspått värde med det faktiska värdet och avvikelsen, som är absolutvärdet av skillnaden mellan dessa två värden. Därefter beräknas genomsnittet av dessa absoluta avvikelser för att få fram MAE. Detta ger ett mått på den genomsnittliga storleken av felet mellan de förutsagda och de faktiska värdena. Eftersom MAE tar hänsyn till absolutvärdena av avvikelserna, elimineras eventuella negativa och positiva avvikelser, vilka således skulle kunnat ta ut varandra. Ju lägre MAE-värdet är, desto bättre överensstämmer de förutsagda värdena med de faktiska värdena. Med dessa egenskaper är detta ett vanligt och användbart mått för att utvärdera prognosmodeller och jämföra deras prestanda.

I MAE mäts den genomsnittliga storleken på felet för den specifika modellens prediktion. Enheten anges i minuter. Detta utvärderingsmått tar inte någon hänsyn till prediktionernas riktning. Det eftersträvas ett så lågt MAE-värdet som möjligt. Den matematiska formeln för medelabsolutfelet ser ut som följer:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{faktisk} - y_{predikterad}|$$

$y_{faktisk}$ är tågets faktiska försening och $y_{predikterad}$ är tågets predikterade försening. Enligt beskrivning av Trafikverkets data, vilket presenteras i nästa kapitel, kan man utläsa att $y_{aktuell} - y_{predikterad}$ är lika med definitionen av värdena i kolumnen precision i datan. Den matematiska formeln kan alltså skrivas om som följande:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{precision}|$$

11.3. Root mean square error (RMSE)

RMSE är likt MAE ett vanligt mått på prognosfel eller felbedömning som används inom statistik och maskininlärning. RMSE används för att mäta den genomsnittliga kvadratrot av kvadratiska avvikelser mellan de förutsagda värdena och de faktiska värdena i en datamängd. För att beräkna RMSE jämförs varje förutsagt värde med det faktiska värdet och kvadraten av avvikelsen beräknas. Därefter beräknas genomsnittet av dessa kvadratiska avvikelser, och för att sedan få fram RMSE tas roten av resultatet. RMSE utgör ett mått på den genomsnittliga storleken av ett fel mellan de förutsagda och faktiska värdena. Genom att kvadrera avvikelserna och sedan ta roten ur undviks negativa avvikelser som kan ta ut varandra. RMSE lägger större vikt vid stora avvikelser jämfört med MAE och kan vara mer känslig för extrema värden. RMSE används ofta för att utvärdera prestanda av prognosmodeller, där lägre RMSE-värden indikerar bättre överensstämmelse mellan förutsagda och faktiska värden.

RMSE mäter alltså skillnaden mellan det faktiska värdet, det vill säga det sanna värdet, och modellens predikterade värde. Ju lägre RMSE-värdet är, desto närmare befinner sig det

predikterade värdet det faktiska, sanna, värdet. Den matematiska formeln för RMSE ser ut som följande:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{faktisk} - y_{predikterad})^2}$$

Precis som för MAE kan RMSE skrivas om till följande:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{precision})^2}$$

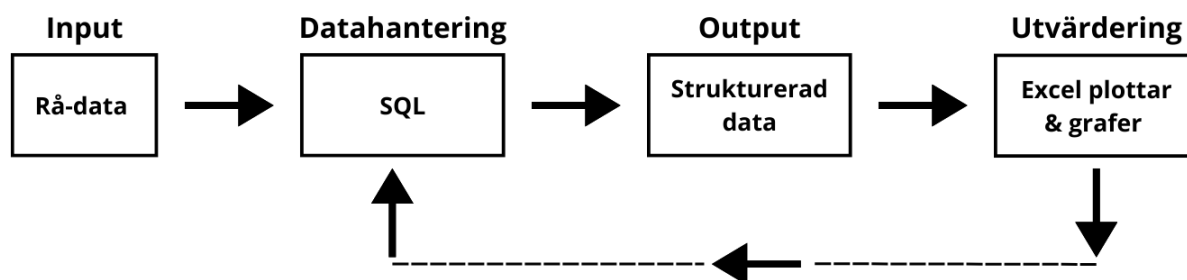
Viktigt att notera är att RMSE är känsligare för extremvärden än vad MAE är. Detta beror på att felen kvadreras innan medelvärdet beräknas, således lägger RMSE en relativt hög vikt på stora, extrema, fel. Detta gör att RMSE med fördel används när stora fel är särskilt oönskade.

Alla tre utvärderingsmetoderna, Medelvärdet, MAE och RMSE kan variera från 0 till oändligheten, ∞ . Vidare är de så kallat negativt orienterade vilket betyder att ett lägre värde anger att prediktionen är närmare det sanna, faktiska, värdet. En annan värdefull notis är att RMSE alltid kommer att vara större än eller lika med MAE. Ju större skillnaden är, mellan RMSE och MAE, desto större är variansen, spridningen, i det enskilda felet. I fallet då RMSE är lika med MAE indikeras detta av att felen i de båda prediktionskällorna är lika stora.

12. Metod och resultat

Detta kapitel utgör själva kärnan av rapporten där försök till att besvara de tre frågeställningarna som nämns i problemställningen görs. Frågeställningarna är starkt sammankopplade till varandra, men metoderna och kriterierna som ligger till grund för analysen skiljer sig självfallet åt mellan frågorna. Själva arbetsgången för respektive del presenteras och beskrivs i inledningen till samtliga frågeställningar. Varje del avslutas med en kortare sammanfattning och rekommendation över hur resultatet bör användas i framtida arbete. Dessa ligger sedan till grund för vidare diskussion i kapitlet om Diskussion och slutsats.

Bilden nedan illustrerar övergripligt den metod som använts i detta arbete, Input – Datahantering – Output – Utvärdering. Själva inputen till arbetet består av samma data som den Trafikverket använder sig av vid skapande av sina ankomstprognoser. I datahanteringsdelen används SQL för att koda och strukturera upp output-datan och plocka fram de parametrar som vidare skall undersökas och utvärderas. Detta leder oss vidare till utvärderingsdelen vilken bygger på output-datan. Denna data matas in i Excel och grafer skapas för att visuellt kunna se hur de olika källornas genererade prediktioner står sig mot varandra och verkligheten. Varje delmoment inom processtrådet i *Figur 6* kommer nedan att förklaras utförligt.



Figur 6 Schematisk illustration av arbetets metod

12.1. Databeskrivning

För att kunna utvärdera ankomstprognoserna har en stor mängd data analyserats. Denna data omfattar en tidsserie gällande ankomster och avgångar för tåg vilka återfinns för samtliga driftsplatser runt om i det svenska tågnätet under tidsperioden 2022-10-01 – 2022-11-15. Den analyserade datamängden består av 5 530 048 rader kod och har med hjälp av SQL-Server Management Studio använts för att sorteras. Rådatan bestod av följande kolumner med information: Operativttåg, utgångsdatum operativttåg, platssignatur, driftledningsområde, aktivitetstyp, källa, prediktionsnummer, prediktionstidpunkt, predikterad tid, planerad tid, rapporterad tid, framsynthet, precision, predikterad försening, trafikläge och punktlighet. Nedan följer en förklaring av de olika kolumnerna.

Operativttåg avser en sifferkombination, vanligen om 3-4 siffror, som indikerar vilket tåg som framförs. Sifferkombinationen är tågspecifikt och unikt för varje tåg och kan således ses som ett id eller personnummer för tåget.

Utgångsdatum operativttåg syftar till vilket datum det specifika operativa tåget har framförts.

Platssignatur är initialerna för den driftplats som prediktionen gäller och som tåget befinner sig vid. Driftplatserna består oftast av 1-3 bokstäver och är unika för varje driftplats.

Exempelvis har driftplats Malmö C platssignaturen M och Stockholm Central har platssignaturen Cst.

Driftledningsområde innefattar vilket geografiskt område som tåget befinner sig i. Sverige är indelat i 8 driftledningsområden: Boden, Ånge, Gävle, Stockholm, Norrköping, Hallsberg, Göteborg och Malmö. Beroende på var tåget befinner sig är det från dessa platser som man ombesörjer, kommunicerar och ansvarar för att leverera tåglägen, likt ett kontrolltorn som tillhandahåller luftrummet för flygplan inom ett specifikt område. Inom varje trafikledningsområde finns en ledningscentral varifrån trafiken leds inom och genom de olika områdena.

I kolumnen **Aktivitetstyp** framgår ifall tåget har ankommit eller avgått från driftplatsen.

I kolumnen **Källa** framgår vilken prediktionskälla som har triggat datapunkten och skapat prognosen. I Trafikverkets data finns 5 olika prediktionskällor: STEG, Amber, Tågprognos, Manuell och AIRT.

Prediktionsnummer avser en siffra som i storleksordning indikerar vilken prediktion datadumpen har. Exempelvis får ett tåg en predikerad ankomsttid till driftplats x då tåget påbörjar sin resa, denna prediktion får prediktionsnummer 1. Ankomsttiden kan komma att justeras allteftersom mer information blir känd under tågets färd. När systemet skapar en ny prediktion blir detta prediktionsnummer 2, nästa prediktion får värde 3 och så vidare. Samma procedur sker i en iterativ process fram till dess att tåget ankommit till driftplats x. Värdet på antalet prediktioner och därmed prediktionsnumret kan komma att variera beroende på olika omständigheter på banan.

Prediktionstidpunkt är den tidpunkt som prediktionen skapades/gjordes.

Predikerad tid är tidpunkten som prognostiserats att tåget skall avgå eller ankomma.

Planerad tid avses som den tänkta avgångs- eller ankomsttid som tåget enligt tidtabell och tågplan är planerad för.

Rapporterad tid är den faktiska tid som tåget avgick eller ankom till driftplatsen.

Framsynthet utgörs av differensen mellan rapporterad tid och prediktionstidpunkt. Det vill säga differensen mellan den faktiska tid som tåget avgick eller ankom och den tidpunkt som prediktionen skapades. Denna differens utgör ett mått på hur lång tid i förväg som prediktionen skapades innan tåget faktiskt avgick eller ankom. Ju större värdet på framsyntheten är desto tidigare, innan tåget faktiskt avgick eller ankom, skapades prediktionen. Exempelvis avser värdet 30 på framsynthet att prediktionen skapades 30 minuter innan tåget faktiskt avgick eller ankom.

Precision är differensen mellan den rapporterade tiden och den predikerade tiden. Det vill säga differensen mellan den faktiska tid som tåget avgick eller ankom och den tidpunkt som prognostiserats fram att tåget skall avgå eller ankomma. Denna differens utgör ett mått på hur god precision prediktion har. Ett negativt värde exempelvis -6 indikerar att tåget ankom eller avgick 6 minuter före den predikerade tiden. Ett positivt värde exempelvis 6 indikerar att tåget ankom/avgick 6 minuter efter den predikerade tiden. Optimalt ligger värdet på 0, vilket indikerar att prediktionen överensstämde med tiden då tåget faktiskt avgick/ankom och således var precisionen av prediktionen träffsäker.

Predikerad försening avses som differensen mellan den predikerade tiden och den planerade tiden. Det vill säga differensen mellan tidpunkten som prognostiserats fram att tåget skall avgå eller ankomma och den tänkta avgångs- eller ankomsttid som tåget enligt tidtabell och tågplan är planerad för. Denna differens utgör ett mått på hur stor den predikerade förseningens tidsrymd är. Ett negativt värde exempelvis -8 indikerar att värdet på prediktionen är 8 minuter tidigare än tågets planerade avgångs eller ankomsttid. Ett positivt värde exempelvis 8 indikerar att den predikerade avgångs eller ankomsttiden är 8 minuter senare än tidtabellens avgångs

eller ankomsttid. Optimalt är den predikterade förseningens värde 0, vilket indikerar att prediktionen överensstämde med tågets tidtabell.

Punktlighet indikerar, i procent, hur många tåg som ankommit till stationen plus eller minus 6 minuter från tidtabellsangiven ankomsttid.

Trafikläge är indelat i tre olika kategorier: Bra, Normalt och Dåligt och med detta menas hur väl trafiken har flutit på den senaste timmen. SQL-koden är skriven så att om medelvärdet av punktligheten plus ett tillägg av en standardavvikelse av punktligheten är större än värdet av punktligheten anses trafikläget vara bra. Om medelvärdet av punktligheten plus ett tillägg av en standardavvikelse av punktligheten är mindre än värdet av punktligheten anses trafikläget vara dåligt. Om värdet befinner sig där emellan anses trafikläget vara normalt.

$(\text{Medelpunktligheten} + 1 \text{ std. avvikelse}) > \text{Punktligheten} \rightarrow \text{BRA}$

$(\text{Medelpunktligheten} + 1 \text{ std. avvikelse}) < \text{Punktligheten} \rightarrow \text{DÅLIGT}$

$\text{BRA} > (\text{Medelpunktligheten} + 1 \text{ std. avvikelse}) > \text{DÅLIGT} \rightarrow \text{NORMALT}$

12.2. Datahantering

Själva skapandet av en prognos för tåg börjar i insamlingen av data, från olika källor. Det kan inkludera tidigare tidtabeller, historiska tågdata, information om eventuella förseningar och avvikelser, realtidsdata om tågtrafiken som sänds från järnvägsoperatörer, samt tidtabellsinformation om tidigare tågavgångar och ankomster. En betydande faktor för att skapa tillförlitliga prognoser är tillgången till högkvalitativa och uppdaterade data. Detta eftersom prognoserna inte blir tillförlitliga om data som ligger till grund är tvivelaktig. Därför är det avgörande att ha ett robust system och tillförlitliga datakällor som inkluderar korrekt tidtabellsdata och noggrann realtidsinformation om tågstatus och eventuella störningar.

Första steget var att importera all rå-data till SQL, vilket gjordes för att kunna strukturera upp den stora datamängden. För att visualisera hur precist de olika prediktionskällorna kan prediktera ankomsttiden med hjälp av rå-datan skrevs följande kod, se *Bilaga 1*. Ett utdrag ur SQL-outputen baserad på koden ses i *Figur 7* nedan, och har tagits fram för att illustrera strukturen i SQL. *Figur 7* visas de 13 första prediktionerna, vilka avser en framsynthet från 120-55 minuter innan tåget ankommer.

	Källa	FRAMSYNTHET	Precision_MEDEL	Precision_STDEV	Precision_MAE	Precision_RMSE	n	punktlighetsläge
1	AIRT	120	6	17	10	20	117	bra
2	AIRT	115	4	10	6	12	116	bra
3	AIRT	105	2	8	3	9	200	bra
4	AIRT	100	1	7	3	7	276	bra
5	AIRT	95	1	6	3	7	342	bra
6	AIRT	90	1	8	4	9	298	bra
7	AIRT	85	3	7	4	8	370	bra
8	AIRT	80	2	6	3	8	331	bra
9	AIRT	75	1	6	3	7	488	bra
10	AIRT	70	1	5	3	6	658	bra
11	AIRT	65	1	5	3	6	653	bra
12	AIRT	60	2	4	2	5	996	bra
13	AIRT	55	1	4	3	5	1107	bra

Figur 7 Output från SQL, baserad på Kod från Bilaga 1

Huvudrubriken 'Källa' i tabellen avser vilken prediktionskälla som använts, i detta fall användes 'AIRT'.

'FRAMSYNTHET' syftar på hur långt i förväg som prediktionen skapades och anges i minuter.

'Precision_MEDEL' är medelvärdet av differensen mellan den faktiska tid som tåget ankom och den tidpunkt som prognostiserats att tåget skall ankomma.

Precision_STDEV är standardavvikelsen för precisionen.

'Precision_MAE' är Mean Absolute Error och indikerar hur stort det absoluta medelfelet är. MAE önskas vara så litet som möjligt, och har enheten minuter.

'Precision_RMSE' är Root Mean Squared Error och är likt MAE ett utvärderingsmått för hur stort prediktionsfelet är.

Kolonnen 'n' avser antal, hur många datapunkter som medelvärdet genererats av. Koden skrevs så att 'n' åtminstone måste innehålla fler än 100 datapunkter.

'Punktlighetsläge' visar punktligheten, hur väl trafiken i driftledningsområdet har flutit på under förgående timme, och är definierad till 'bra', 'dåligt' eller 'normalt'. I detta fall var punktligheten förgående timme 'bra' för samtliga prediktioner.

Den första raden i *Figur 7* ovan avläses följande: Beräkningen är skapad med hjälp av modellen 'AIRT'. När det återstod 120 minuter tills tåget skulle ankomma var medelvärdet av precisionen 6 minuter. Denna prognostisering menar således att tåget i medeltal har en försening av 6 minuter, jämfört med vad som faktiskt skedde i verkligheten. MAE och RMSE har värdet 10 respektive 20 minuter och standardavvikelsen för precisionen är 17. Trafikläget har förgående timme varit 'bra', det vill säga att trafiken har flutit på. Medelvärdet av precisionen bygger, och är beräknad, på 117 datapunkter. Samma process och tankebana följs genom hela *Figur 7*.

SQL-koden genererade likartad typ av tabell för alla 5 källorna 'Amber', 'STEG', 'Manuell', 'Tågprognos' och 'AIRT'. Eftersom själva rådatan inte är någon specificerade data som är skapad för att göra extern analys på finns en viss oregelbundenhet. Exempelvis finns det olika mängder datapunkter för de olika beräkningskällorna. STEG och AIRT, som genereras automatiskt, innehåller avsevärt fler datapunkter än vad exempelvis Manuell gör. AIRT är den källa som har flest mätvärden och Tågprognos har minst. Vidare finns det avvikelser gällande de manuellt skapade datapunkterna. Inom en del driftledningsområden saknas det helt och hållet manuellt skapade data.

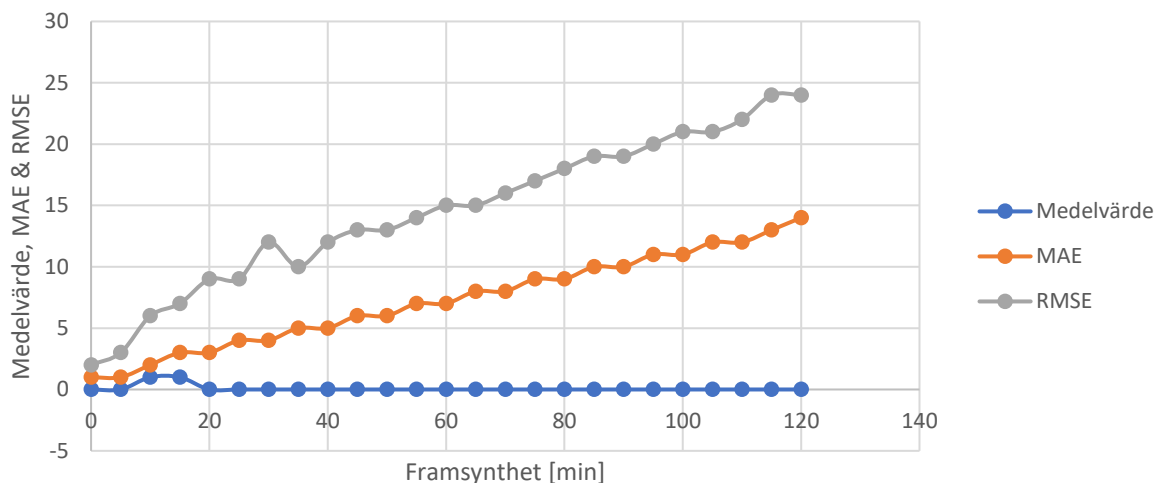
12.3. Vilket utvärderingsmått bör användas

För att besvara den första problemställningen och för att undvika att rapporten överväldigas av grafer var nästa steg att ta reda på vilket utvärderingsmått som tydligast genererar de mest tillförlitliga prediktionerna. Det bäst lämpade utvärderingsmålet kommer sedan uteslutande tillämpas genom resterande del av rapporten och när nästkommande frågeställningar besvaras. Om inte detta hade gjorts separat, utan var för sig, skulle det behöva skapas medelvärdes-, MAE- och RMSE-grafer för varje prediktionskälla, vilket i sig skulle medföra 15 grafer.

För att kunna ta reda på vilket utvärderingsmått som bör användas vid utvärdering av ankomstprognoser skrevs en kod i SQL, där tanken var att plocka fram beräknade medel-, MAE- och RMSE-värden för varje minut på framsynthet, mellan 0-120 minuter. Avsikten med denna kod var att se hur de olika utvärderingsmåten "straffar" prediktionerna, samt hur de ter och står sig mot varandra. Datapunkterna med tillhörande medel-, MAE- och RMSE-värden genererade av SQL-koden plottades i grafer och det framkom stora variationer med omfattande toppar och dalar. Vid närmare undersökning av SQL-koden framgick att i de avvikande

datapunkterna, topparna och dalarna, var antalet bakomliggande prediktioner markant lägre. Särskilt tydligt var detta för prediktioner som genererats tidigt, det vill säga med stora värden för framsynhet, långt innan tågets ankomst. För att undvika dessa toppar och dalar ändrades SQL-koden till att framsynheten delades in i femminuters-intervall mellan, 0-120 minuter. Denna ändring medförde att antalet prediktioner för de plottade datapunkterna nu i stället innehöll nästan 5 gånger fler prediktioner. Antalet ökade från ett lägsta värde $n=8620$ till ett nytt lägsta värde $n=42590$, vilket därmed genererade en avsevärt jämnare kurva i grafen. Vidare skrevs den nya koden med avsikt att inte ta hänsyn till vilken specifik prediktionskälla som låg bakom prediktionen, samt att inte heller ta hänsyn till det rådande trafikläget på banan. *Graf 1*, se nedan, är framställd genom att ha plockat fram Medel-, MAE- och RMSE-värde för samtliga 5 530 048 datapunkter, oberoende av prediktionskälla. Det vill säga att de tre plottade linjerna innefattar datapunkter innehållande prediktioner från samtliga prediktionskällor. Detta gjordes eftersom jag i detta skede endast var intresserad av hur utvärderingsmåttens står sig mot varandra och straffar datapunkterna. De framtagna värdena som genererats via den nya koden plottades på nytt och följande graf skapades, se *Graf 1* nedan.

Resultatet från vanligt förekommande utvärderingsmått



Graf 1 Resultat från de olika utvärderingsmått

Vad man kan se i *Graf 1* ovan är att medelvärdet inte är det bästa utvärderingsmålet att använda sig av. Via medelvärdemetoden ser man att precisionen av prediktionen då tåget ankommer är 0 minuters försening, vilket inte speglar verkligheten. Detta gäller för nästintill samtliga prediktioner 2 timmar innan tåget faktiskt ankommer. Vidare är medelvärdemetoden inte lämplig i denna studie eftersom felet inte tar ut varandra. Detta medför att det med medelvärdemetoden blir ett medelfel, vilket inte speglar verkligheten på ett tillförlitligt sätt. Om exempelvis de olika prediktionskällorna kontinuerligt överskattar ankomsttiderna skulle detta medföra att medelvärdet också blir överskattat, och vice versa. Om ett tåg ankommer 10 minuter för tidigt och ett annat ankommer 10 minuter för sent skulle detta medföra att medelvärdet av ankomsten blev 0 minuter, vilket således betyder att tågen enligt medelvärdemetoden ankommit i tid.

Både MAE och RMSE är utvärderingsmått som ”straffar” felprediktioner. Enligt definitionen av RMSE kvadreras prediktionerna innan kvadratroten tas, vilket gör att stora fel straffas extra hårt. Om man är intresserad av att jämföra storleken av förseningar kan det då vara en fördel att använda RMSE. Eftersom prediktionsvärdena i RMSE kvadreras blir eventuella negativa fel positiva, det vill säga att det här inte tas någon hänsyn till om tåget kommer x-antal minuter för

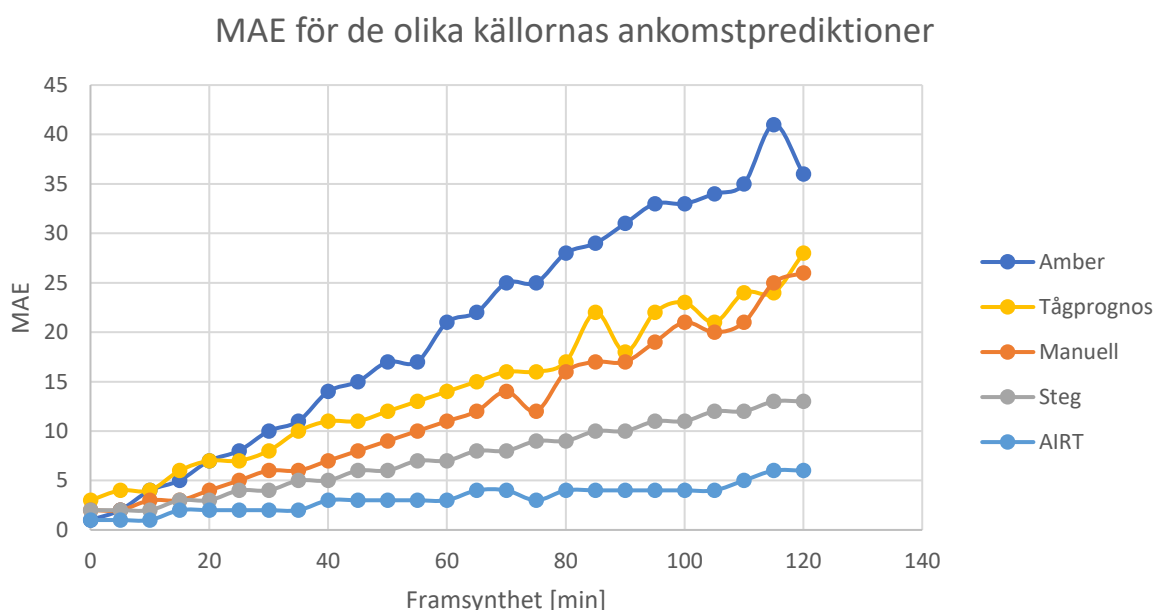
tidigt eller för sent, utan straffar överskattade eller underskattade prediktioner lika mycket. Enligt definitionen av MAE tas absolutbeloppet av den eventuella felprediktionen vilket även här gör att ingen hänsyn tas till om prediktionen är över- eller underskattad. Skillnaden mellan MAE och RMSE är att felprediktionerna i MAE inte kvadreras och att stora värden således inte straffas riktigt lika hårt. I denna studie tas ingen hänsyn till storleken av tågens ankomstförsening, att tågen ankommer för sent anses vara lika illa oavsett storleken av själva förseningen. Därav används MAE fortsättningsvis som utvärderingsmått genom rapporten.

12.3.1. Sammanfattning & rekommendation gällande lämpligast utvärderingsmått

Medelvärdesmetoden är inget bra mått att använda sig av vid utvärdering av ankomstprognoser, eftersom den inte speglar det verkliga utfallet. MAE och RMSE utgjorde snarlika resultat. Valet av att rapporten fortsatt skulle använda MAE som utvärderingsmått föll på att ingen hänsyn till förseningens storlek skulle vidtas. Om fallet varit annorlunda skulle RMSE per definition varit ett bättre utvärderingsmått. Rekommendationen är således att användning av MAE som utvärderingsmått bör användas. Detta tar vi med oss in i nästa problemställning.

12.4. Vilken prediktionskälla ger mest tillförlitlig ankomsttid

Nästa steg var att producera en ny kod som utifrån MAE utvärderade och straffade de olika prediktionskällornas ankomstprediktioner. Den nya koden skapades för att få fram MAE värden för samtliga prediktionskällor Amber, Tågprognos, Manuell, STEG och AIRT, och för att se källornas tillförlitlighet att spegla verkligheten. Denna kod skrevs snarlikt förgående SQL-kod och delade på nytt in framsynheten i 5-minuters intervall, som sträcker sig från 0-120 minuter innan tågets faktiska ankomst. Datapunkterna plottades och följande graf kunde nu produceras, se *Graf 2* nedan.



Graf 2 MAE-resultatet för samtliga prediktionskällors ankomstprediktioner

I *Graf 2* ovan syns MAE-resultaten från samtliga prediktionskällors ankomstprediktioner. Ur grafen kan utläsas att prediktionskällan Amber genererar markant sämre ankomstprediktioner för större värde på framsyntheten än vad de andra källorna gör. Inte förrän framsyntheten är ca 25 minuter sammanfaller Amber med en av de andra prediktionskällorna, Tågprognos. Källorna Tågprognos och Manuell visar likartade ankomstprediktioner för större värden på framsyntheten, även om Manuell näst intill för samtliga värden utgör aningen bättre prediktioner. Näst längst ner i grafen befinner sig STEG vilken ser ut att generera skapliga prediktioner utan några större fluktuationer. Längst ner i grafen hittar vi AIRT som är den källa som av alla prediktionskällor signifikant skapar mest tillförlitliga ankomstprediktioner. AIRT visar likt STEG en något rät linje utan större fluktuation. Skillnaden är att AIRT ger nästan dubbelt så tillförlitliga prediktioner.

För att skapa en grov översikt av hur stor skillnad som råder mellan prediktionskällornas MAE-värden togs medelvärdet fram från källorna i *Graf 2*, vilket framgår i *Tabell 1* nedan. Rankingen indikerar källornas resultat där 1 är den bästa och 5 den sämsta. Medelvärdet är detsamma som medelfelet, vilket förenklat kan tolkas som den felprecision i prediktionen som man i genomsnitt får räkna med.

Tabell 1 De olika källornas medelvärden och ranking.

	AIRT	Amber	Manuell	STEG	Tågprognos
Medelvärde:	3,2	20,16	11,84	7,28	14,24
Ranking:	1	5	3	2	4

I *Tabell 2*, nedan, framgår ett utdrag ur SQL-datan för några av de värden som ligger till grund för *Graf 2*. Här syns att framsyntheten är 120 minuter, det vill säga 2 timmar innan tåget faktiskt ankommer, har AIRT ett MAE värde på 6min, Amber ett värde på 36min, Manuell 26min, STEG 13min och Tågprognos ett MAE-värde på 28min. Om man kikar närmare i tabellen syns att AIRT har lägsta värden. Detta gäller för samtliga tider på Framsynthet, förutom vid Framsynthet 0 minuter då Amber utgör samma MAE-värde. AIRT är den källa som ger mest tillförlitliga prediktioner för tågens ankomster.

Tabell 2 MAE-värden för källorna vid olika tider på Framsynthet.

	AIRT	Amber	Manuell	STEG	Tågprognos
Framsynthet 120:	6	36	26	13	28
Framsynthet 100:	4	33	21	11	23
Framsynthet 80:	4	28	16	9	17
Framsynthet 60:	3	21	11	7	14
Framsynthet 40:	3	14	7	5	11
Framsynthet 20:	2	7	4	3	7
Framsynthet 0:	1	1	2	2	3

Tabell 3 nedan visar differensen mellan AIRT och övriga prediktionskällor (Amber-AIRT, Manuell-AIRT, STEG-AIRT samt Tågprognos-AIRT) för olika värden på Framsynthet. Ur tabellen kan utläsas att med minskande värde på Framsynthet, det vill säga ju närmare den faktiska ankomsten vi rör oss, desto lägre blir differenserna. Med detta sagt blir prediktionerna mer precisa ju närmare den faktiska ankomsten vi kommer. Tabellen utläses på följande vis: Då

framsyntheten är 120 minuter predikterar AIRT 30min bättre än vad Amber gör, 20min bättre än Manuell, 7min bättre än STEG och 22min bättre än Tågprognos. När framsyntheten är 100 minuter predikterar AIRT 29min bättre än Amber, 17min bättre än Manuell, 7min bättre än STEG och 19min bättre än Tågprognos.

Tabell 3 Differensen mellan Amber och AIRT, Manuell och AIRT, STEG och AIRT samt Tågprognos och AIRT.

	<i>Diff Amber-AIRT</i>	<i>Diff Manuell-AIRT</i>	<i>Diff STEG-AIRT</i>	<i>Diff Tågprognos-AIRT</i>
Framsynthet 120:	30	20	7	22
Framsynthet 100:	29	17	7	19
Framsynthet 80:	24	12	5	13
Framsynthet 60:	18	8	4	11
Framsynthet 40:	11	4	2	8
Framsynthet 20:	5	2	1	5
Framsynthet 0:	0	1	1	1

12.4.1. Sammanfattning & rekommendation gällande mest tillförlitlig ankomsttid

Det råder inga tvekan om att AIRT är den prediktionskälla som ger bäst resultat och mest tillförlitliga ankomstprognoser. Rekommendationen är således att AIRT vidare bör användas och övriga prediktionskällor kan avvecklas. Detta eftersom det inte finns någon anledning att ha flera olika prediktionskällor, speciellt inte när de skapar så pass mycket sämre prediktioner än vad AIRT gör. Vi tar med oss detta in i nästa problemställning som följer nedan.

12.5. Vilken framförhållning är väl avvägd

För att validera och se ifall rangordningen för prediktionskällorna var slumpmässiga eller står sig konstant utfördes ett test. Testet byggde på att skapa en ny kod i SQL där prediktionskällorna kontrollerades för olika trafiklägen ute på banan oavsett om det varit mycket förseningar eller om trafiken flutit på. För att ta reda på detta skrevs SQL-koden så att datapunkter framkom baserat på hur trafikläget inom driftledningsområdena sett ut förgående timme. Detta gjordes genom att kika på differensen mellan 'Planerad_tid' och 'Rapporterad_tid'. Om differensen mellan dessa var mindre än 6 minuter hade tåget per definition ankommit i tid och tilldelas då värdet 100. Detta värde utgör en procentsats och indikerar tågets punktlighetsgrad. Om differensen var större än 6 tilldelades värdet 0.

Därefter togs medelvärde- och standardavvikelse av punktligheten fram. Dessa adderades samman så att summan av medelvärde- plus standardavvikelsen av punktligheten tilldelades variabeln 'bra'.

$$avg(punktlighet) + round(stdev(punktlighet), 0) = bra$$

På liknande sätt skapades en subtraktion mellan medelvärdet- minus standardavvikelsen av punktligheten och tilldelades variabeln 'dåligt'.

$$avg(punktlighet) - round(stdev(punktlighet), 0) = dåligt$$

Om punktligheten var större än variabeln 'bra' tilldelades denna på nytt variabeln 'bra' och således antyddes att läget på banan förgående timme varit just 'bra'.

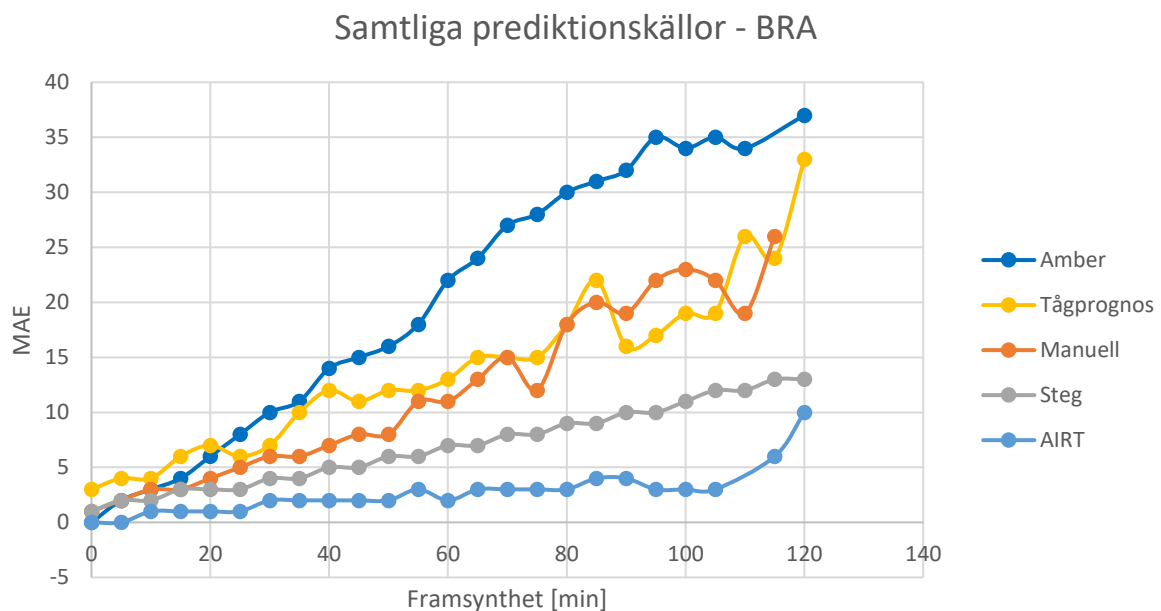
$$Punktlighet > bra = bra$$

Om punktligheten var mindre än variabeln 'dåligt' tilldelades denna på nytt variabeln 'dåligt' och trafiken förgående timme anses således inte ha flutit på som vanligt.

$$Punktlighet < dålig = dåligt$$

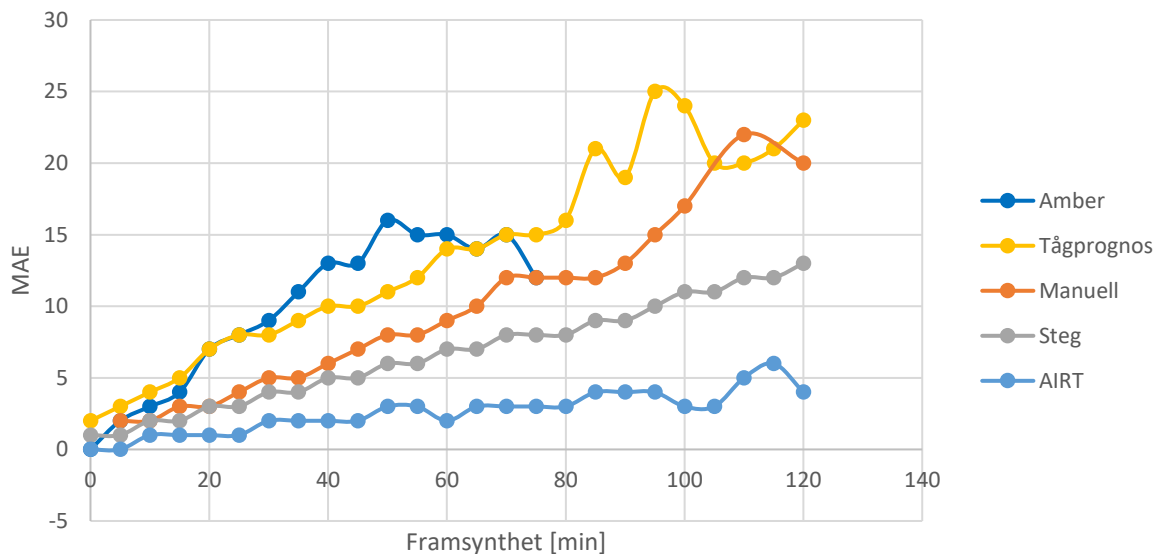
Värden som hamnar mellan bra och dåligt tilldelas variabeln 'normalt', vilket indikerar att trafiken förgående timme skett med 'normal' drift.

Med hjälp av denna SQL-kod kunde samtliga prediktionskällors datapunkter buntas ihop och sorterades utefter rådande trafikläge på banan: 'bra', 'normalt' eller 'dåligt'. Graf 3-5 nedan visar Framsynthet på x-axeln och MAE på y-axeln för samtliga prediktionskällor vid de olika trafiklägena.



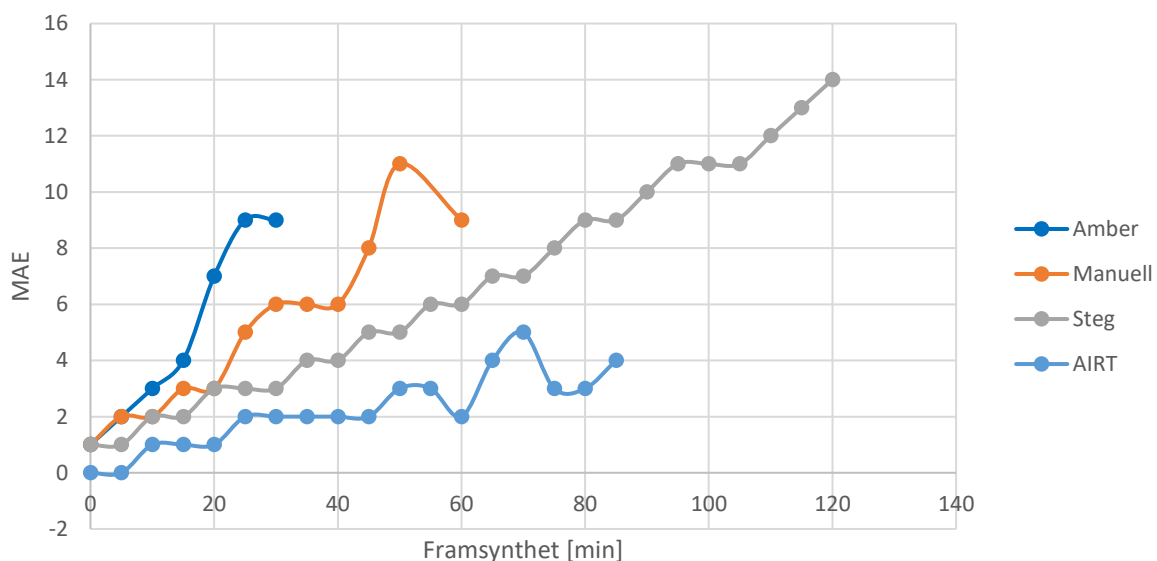
Graf 3 Samtliga prediktionskällors MAE-värden vid olika tider på framsynthet med 'BRA' trafikläge förgående timme

Samtliga prediktionskällor - NORMALT



Graf 4 Samtliga prediktionskällors MAE-värden vid olika tider på framsynthet med 'NORMALT' trafikläge förgående timme

Samtliga prediktionskällor - DÅLIGT



Graf 5 Samtliga prediktionskällors MAE-värden vid olika tider på framsynthet med 'DÅLIGT' trafikläge förgående timme

Ur graferna ovan kan vi utläsa att prediktionskällan Amber, marinblå linje, skapar sämst prediktioner med stora över- och underskattningar av tågens ankomster. Amber är den prediktionskälla som för samtliga trafiklägen gör sämst prediktioner. Detta visas genom att den marinblå linjen ligger högst upp i samtliga grafer. Vidare kan ur graferna avläsas att Amber, Tågprognos och manuella prognoser har, generellt, en lägre andel korrekta prognoser än de andra två källorna STEG och AIRT, men andelen korrekta ankomstprognoser ökar drastiskt ju

närmare tåget är sin ankomst. Prediktionskällorna Tågprognos, gul linje, och Manuell, orange linje, förhåller sig tämligen jämnt till varandra då trafikläget på banan förgående timme varit bra. För större värden på Framsynthet ser Tågprognos ut att skapa aningen bättre prediktioner. Med lägre värden på Framsynthet gör de Manuella tågprognoserna bättre prediktioner än Tågprognos. Detta blir särskilt tydligt då trafikläget varit normalt. Dessvärre finns det ingen data för Tågprognos vid dåligt trafikläge men eftersom prediktionskällorna i graferna för de olika trafiklägena haft samma rangordning är ett kvalificerat antagande att även Tågprognos skulle haft samma rangordning ifall datan varit fullständig. Prediktionskällan STEG, grå linje, utgör mycket jämna och icke-fluktuerande prediktioner för samtliga trafiklägen på banan. AIRT, ljusblå linje, är den prediktionskälla som i varje graf markant skapar bäst prediktioner. Likt STEG beräknar AIRT prediktioner som inte fluktuerar särskilt mycket. Skillnaden är att prediktionerna från AIRT är skarpare.

I graf 3-5 ovan syns att när Framsyntheten är runt 60 minuter så är prediktionerna nästan dubbelt så träffsäkra som de var när Framsyntheten var 120 minuter, med ett medelfel på ca 8 minuter, oavsett trafikförhållande. När Framsyntheten sänks till 40 minuter uppkommer ett medelfel på ca 6 minuter vilken är den tid då Trafikverket inte ger ut någon orsakskod vid försening. Således anses en framförhållning av 40 minuter vara rimlig att använda. Det skall dock tilläggas att detta värde på 6 minuter är ett snitt för samtliga prediktionskällor. Om man enbart använder sig av AIRT och slopar de andra prediktionskällorna skulle man utan problem kunna använda sig av prediktioner skapade med en framsynthet på 120 minuter och få ut ett bättre medelfel i prediktionen än vad de andra prediktionskällorna har vid 40 minuter.

12.5.1. Sammanfattning & rekommendation gällande framförhållning

Baserat på den data som tilldelats detta examensarbete påvisas att med hjälp av prediktionskällan AIRT kan ankomstprediktioner med god precision skapas med en framförhållning på 120 minuter. För övriga prediktionskällor speglar ankomstprediktionerna verkligheten avsevärt sämre med stora värden på Framsynthet. Inte för än omkring 40 minuter innan tågets faktiska ankomst börjar övriga prediktionskällor ge någorlunda relevanta prediktioner, dock inte i närheten av AIRT. Rekommendationen blir således fortsatt att slopa övriga prediktionskällor och i stället gå vidare med och lägga krut på att utveckla AIRT ytterligare.

13. Diskussion och slutsats

För att kunna ge en tydlig och rationell utvärdering av prediktionskällorna generellt och specifikt, är det viktigt att vara tydlig med vad de är till för och hur de presterar. Annars är det svårt att ge en tydlig rekommendation för framtida användning. Detta kapitel startar med en genomgång av de tre problemställningarna var och en för sig: vilket utvärderingsmått bör användas, vilken prediktionskälla ger mest tillförlitlig ankomsttid och vilken framförhållning är väl avvägd. Kapitlet avslutas med en allmän diskussion kring nyttan med prognoserna.

Målet med examensarbetet är att undersöka och utvärdera ankomstprognoser för tåg. Detta görs genom att analysera de bakomliggande prediktionskällornas förmåga att skapa tillförlitliga ankomstprognoser. Det huvudsakliga fokuset är att öka kunskapen om såväl styrkor som svagheter hos prediktionskällorna, samt att se om det finns situationer där den ena eller andra ter sig särskilt lämplig.

13.1. Vilket utvärderingsmått bör användas

För att uppnå målet med arbetet: undersöka och utvärdera ankomstprognoser för tåg, användes tre vanligt förekommande utvärderingsmått, medelvärdesmetoden, MAE och RMSE. Det visade sig att medelvärdesmetoden inte är något bra mått att använda sig av vid utvärdering av ankomstprognoser. Att så är fallet blev extra tydligt enligt *graf 1*. Medelvärdesmetodens utvärdering av alla datapunkter gav upphov till en näst intill vågrät linje, blå linje. Denna så gott som vågräta linje påvisar att tåget nästan inte har någon försening och att detta skulle gälla för samtliga värden på Framsynthet hela 2 timmar innan tåget faktiskt ankommer. Så är inte fallet i verkligheten och således är detta utvärderingsmått inte tillförlitligt vid utvärdering av ankomstprognoser.

Linjerna för MAE, orange linje, och RMSE, grå linje, skiljer sig distinkt från medelvärdeslinjen. Dessa linjer är relativt rätlinjiga och indikerar således att med ett större värde på Framsynthet, ju längre tid innan tåget ankommer, desto större värde och därmed sämre är prediktionen, vilket man kan förvänta sig. Detta eftersom det finns en stor mängd oförutsägbara händelser som kan göra att tågets predikterade ankomsttid påverkas när prediktionen är skapad tidigt. Valet att använda MAE som utvärderingsmått i rapporten gjordes då ingen hänsyn till förseningens storlek vidtogs. Om förseningens storlek är av intresse kan med fördel RMSE användas eftersom RMSE straffar extrema värden extra hårt. Självklart finns det en uppsjö andra utvärderingsmått som tänkbart skulle kunna användas. I detta arbete användes tre av de vanligast förekommande måtten vilka räckte för att få en indikation på prediktionernas tillförlitlighet. Troligen hade inte ett mer avancerat utvärderingsmått förändrat utfallet.

Resultatet av denna frågeställning är inte alls underlig, utan snarare väntad. Att medelvärdesmetoden skulle vara ett bra utvärderingsmått för ankomstprognoserna vore osannolikt. Detta eftersom medelvärdesmetoden presenterar ett medelfel utan att straffa eller ens ta felen i beaktande. Vidare straffar RMSE per definition prediktionernas extremvärden extra hård. RMSE förlorar sin främsta egenskap då avgränsningarna i detta arbete inte tar hänsyn till förseningens storlek. Därav anses MAE var det mest lämpade utvärderingsmålet att använda sig av här.

13.2. Vilken prediktionskälla ger mest tillförlitlig ankomsttid

För besvara den andra delen av examensarbetets mål: att öka kunskapen om såväl styrkor som svagheter med de olika prediktionskällorna visas följande: Ur *Graf 2* och *tabell 1* kan ses och utläsas att sämst tillförlitlighet utgörs av prediktionskällan Amber, vilken har ett medelfel på 20,16 minuter. Den näst sämsta prediktionskällan att spegla verkligheten utgörs av Tågprognos, med ett medelfel på 14,24 minuter. Tågprognoslinjen ligger tätt intill den manuella prediktionskällans linje, vilken har ett medelfel på 11,84 minuter. Det skall dock tilläggas att för större värden på Framsynthet, ju längre i förväg prognosen sätts, desto markant sämre presterar den manuella prognosen. Prediktionskällan STEG skapar näst bäst och tillförlitliga ankomstprognoser och ger ett medelfel på 7,38 minuter. STEG ger således nästintill dubbelt så bra prediktioner som Tågprognos. Den prediktionskälla som ger absolut mest tillförlitliga prognoser är AIRT, vilken enbart har ett medelfel på 3,2 minuter, vilket är nästan 2,5 gånger bättre medelfel än STEG.

I *Graf 2* ser vi att i fallet för Amber, Tågprognos, Manuell, och till viss del STEG, blir prediktionerna avsevärt bättre med ett lägre värde på Framsynthet. Det vill säga att ju närmare den faktiska ankomsten vi kommer desto markant bättre blir prediktionerna på att spegla verkligheten. För prediktionskällan STEG är denna skillnad inte speciellt stor och för AIRT är den enbart någon minut.

Vidare är det vid genomgång av resultatet i *Graf 2* anmärkningsvärt att de manuellt genererade prognoserna har så pass dålig prognosräffsäkerhet. Såklart finns en stor variation i data som är genererad av människor kontra den som är producerad automatiskt. Den mänskliga faktorn har stor påverkan och olika människor kommer alltid att arbeta på olika sätt. Dock vore det fördelaktigt att skapa tydliga regler och uppgifter för att minimera risken för ytterligare yttre påverkan. Särskilt viktigt att ha tydliga regler och riktlinjer är det då flera händelser inträffar på en och samma gång, samtidigt som personalen har stor arbetsbörda. I vissa fall finns det i datan manuella prediktioner som avslutats runt 24 timmar efter själva tågets ankomst, vilket gör att man kan undra hur pass tillförlitlig den informationen fortfarande är.

Att skapa tillförlitliga ankomstprognoser för tåg är av stort värde och fungerar i dagsläget, som syns i *Graf 2*, med varierande resultat. Genom användning av AI-modellen kan de prognostiserade ankomster förbättras och skapas till att bättre spegla verkligheten. I detta arbete kan vi konstatera att AI-prognosen, AIRT, genererar markant bättre och tillfredställande resultat än vad någon av de övriga fyra prediktionskällorna gör. Det fortsatta arbetet med att sätta prognoser för tågankomster är styrd av tillgången på realtidsdata. Allt efterhand som mer data finns att tillgå och fler prognoser kan skapas kommer algoritmen bakom AIRT att skarpas kunna lära sig känna igen de mönster och trender som råder, för att på så sätt skapa ytterligare precision i prognoserna. Vidare, är en rekommendation att se över kvaliteten på input-datan. Den data som Trafikverket samlar in i dagsläget är avsedd för de prognostiseringskällor som inte använder AI. Med andra ord fanns det förr andra krav på datakvalitet än vad det gör idag. Trafikverket hade vid starten av insamling av data förmodligen andra syften än för input till AI-modeller. Därför bör kvaliteten av input-datan ses över så att den möjliggör för AIRT att prestera på topp.

Rekommendationen är att uteslutande använda sig av AIRT. STEG utgör förvisso goda och jämna prediktioner, men eftersom AIRT skapar nästintill dubbelt så bra prediktioner för alla tider på Framsynthet är detta den prediktionskälla som uteslutande bör fortsatt användas framöver. Det finns ingen mening med att ha prognoser som ändå inte skapar prognoser med tillräckligt hög kvalitet. I stället bör man se till att lägga fokus på att utveckla algoritmerna och kontinuerligt mata dem med nya data för att kunna uppnå än bättre och vassare resultat. AIRT har endast varit verksamt under ca två månaders tid. Ändå utgör den redan en markant bättre

förmåga att skapa precisa prediktioner än vad någon av de andra prediktionskällorna gör och som varit verksamma under längre tid. Om man använder AIRT och matar algoritmen med rikligt med data bör man i framtiden kunna skapa träffsäkra och precisa prognoser.

13.3. Vilken framförhållning är väl avvägd

För att knyta an till examensarbetets sista delmål: i vilka situationer utvärderingsmått ger mest tillförlitliga ankomsttider för tågen samt vilken framförhållning som är väl avvägd, ges följande diskussion.

Framförhållningen för hur tidigt eller sent en prognos skall skapas bör vara densamma oavsett vilket trafikförhållande som råder på banan. Däremot kan prognoser som skapas en längre tid i förväg respektive tätt in på ankomsten, och med olika trafikförhållanden på banan såklart ge olika bra prediktioner. Oavsett hur trafikläget har sett ut på banan förgående timme är det AIRT som även denna gång skapar bäst prediktioner. Detta styrks närmare i *graf 3-5* där det inte råder någon tveksamhet om att AIRT är den prediktionskälla som på alla plan ger bäst resultat och mest tillförlitliga ankomstprognoser, oavsett förhållande på banan. Ordningen för hur de olika prediktionskällornas MAE-värden står sig mot varandra är den samma som i förgående frågeställning och i *graf 2*. Således kanske det inte är en slump att ordningen ser ut som den gör. Även om det inte finns någon data för hur källan Tågprognos står sig mot de övriga prediktionskällorna då förhållandena på banan är dåliga, så är ett rimligt antagande att även denna står sig konstant likt tidigare.

Utifrån den tillgängliga datan kan man med AIRT som prediktionskälla med fördel skapa tillförlitliga prognoser med en Framsynthet på 120 minuter. De övriga prediktionskällorna kräver att Framsyntheten är omkring 40 minuter för att tillförlitligheten skall börja bli trovärdig, vilket kan ses i *graf 3-5*.

Rekommendationen är att fortsättningsvis använda sig av AIRT. En notis är att prediktionskällan Amber, vilken genomgående haft sämst värden, ser ut att göra bättre prediktioner än både Tågprognos och Manuell då Framsyntheten blir tillräckligt liten. En tanke skulle kunna vara att använda sig av en kombination av prediktionskällor. Med detta menas att man hela tiden använder den bäst presterande källan som prediktionskälla. Exempelvis om det vid stora värden på Framsynthet är en källa som presterar särskilt bra används denna. När framsyntheten blir mindre och det är en annan källa som då presterar särskilt bra, så används i stället denna. Dock är detta inget alternativ i detta examensarbete eftersom AIRT genomgående, oavsett storlek på Framsynthet, presterar och ger bäst prediktioner. Sammanfattningsvis blir rekommendationen att prediktionskällorna Amber, Tågprognos, Manuell och STEG ska tas ur bruk och ersättas av AIRT.

13.4. Framtida utvecklingsmöjligheter

Då de olika prediktionskällorna presterar mycket ojämnt, speciellt då värdena på Framsynthet är stora, är det svårt att med ett gemensamt bestämt värde bestämma framförhållningen på prediktionen. Dock går det inte att ha olika framförhållning för olika typer av prediktionskällor, eftersom detta vore orättvist ur ett jämförande perspektiv. Vad jag menar är att man för en prediktionskälla inte kan ha en Framsynthet på 20 minuter medan en annan har en Framsynthet på 60 minuter.

Det går att på flera olika sätt beskriva hur väl en prognos ter sig i tågnätet. Ett sätt är att klargöra hur noga den predikerade prognosen för tågets ankomst till nästa stationer överensstämmer med tidtabellen. I stället för denna tidsgenererade prediktion kanske det kan vara intressant att prediktionen skapas då det är ett specifikt antal kilometer kvar till ankomst, eller med ett specifikt antal stationer kvar till ankomst? Men även genom en specifik kvarvarande distans sätta prognosen, det vill säga när tåget har ett specifikt antal km kvar till ankomst, så skall prognosen sättas. Ett annat exempel kan vara att prognosen genereras när tåget befinner sig ett visst antal stationer från prognosstationen. Ytterligare ett exempel kan vara att prognosen skapas en specifik tid innan planerad ankomst. Därför skulle en aspekt av framtida arbete kunna vara att ta fram var tåget borde befinna sig vid den aktuella tidpunkten då trafiksituationen faktiskt inträffar. Detta för att kunna skapa en trigger som aktiveras när en prognos behöver sättas. Om ett tåg enligt tidtabell ska befinna sig på driftplats A, men i självverket befinner sig någon annanstans, skulle detta kunna vara den utlösande faktorn till att prognoser för detta tåg påbörjas. Eftersom tåget i detta fall befinner sig på annan plats än vad den enligt tidtabell borde, blir tåget med största sannolikhet förtidigt/försenat och kommer därför behöva en ny prognos.

En uppmaning som är värd att flagga för är hur prioriteringen för tågen skall se ut. Med tanke på att prognoserna skapas i realtid för alla tåg i bruk kanske det är mest lämpligt att prioritera prognoser för de tåg som verkligen behöver det.

Varje tåg i Sveriges tågnät har specifika attribut som påverkar dess förutsättningar att anlända och avgå i tid samt dess förutsättningar att ta igen en försening. Dessa attribut tar form i exempelvis tågets tidigare förseningar, tågvikt, antal vagnar, drivmedel eller trafikutöware. Det absolut viktigaste attributet för prognostisering av tågets ankomst och avgång till station är dess tidigare förseningar över sträckan. Ett försenat tåg kommer förmodligen att bli mer försenat. Med andra ord är det större sannolikhet att ett tåg som är försenat vid station A också är försenat vid efterföljande, station B, än att det ankommer i tid. Därför bör framtida arbete kring tågprognoser till stor del baseras på hur den senaste förseningstrenden ser ut. Det är här AIRT, med sin förmåga att kunna finna mönster och trender, kommer att kunna leverera tillförlitliga prediktioner. Utöver tågets egna attribut spelar sträckan som tåget befinner sig på en stor roll. Sträckan i fråga omfattar även de tåg som befinner sig på sträckan eller på angränsande sträckor. Tågsystemet är ett komplext system och aktuella trafikmönster påverkar enskilda tågs förseningar. Därför är en annan viktig del i framtida arbete att även väga in och ta hänsyn till andra angränsande tågs prognoser vid prognostisering av det enskilda tåget.

Ankomstprognoser är inte enbart intressanta för nästkommande station, utan för samtliga framtida driftsplatser och stopp på sträckan. Detta gör att prognoserna i stor utsträckning är sammanlänkade och bygger på varandra. Därmed kan information från de tidigare prognoserna ge nyttig information för hur de framtida prognoserna kommer att vara. Ju tidigare i förväg, och därmed längre fram på sträckan, som en ankomstprognos efterfrågas desto vidare blir felmarginalerna i prognosen. De tidigare tågförseningarna och dess prognoser innehåller en stor mängd nyttig information. Därför är det av stor vikt att algoritmer använder sig av tidigare prognoser, och förseningar, som input vid prognostisering av framtida prognoser.

Självklart finns det även andra AI-modeller än AIRT som kan skapa och beräkna ankomstprognoser. Därför kan ytterligare framtida arbete med fördel testa andra maskininlärningsalgoritmer och jämföra dem med AIRT för att se hur de står sig gentemot

varandra. Hur komplext man utvecklar nätverkets attribut är direkt begränsat till den typ av maskininlärningsmodell man använder sig av.

För att spinna vidare på och fortsätta arbeta med utvärdering av tågprognoser vore det i ett framtida projekt lämpligt att ta hänsyn till storleken av förseningen och inte avgränsa bort den. Jag tror då att man med fördel hade kunnat använda RMSE i stället för MAE som utvärderingsmått, eftersom RMSE per definition straffar felprediktioner extra hårt. Intressant vore att undersöka och jämföra de olika prediktionskällorna vid stora respektive små förseningar. Skulle utvärderingsmättet RMSE lyckas straffa små förseningar bättre än vad MAE gör eller är det endast vid stora förseningar som RMSE ter sig särskilt lämpligt?

I ett framtida projekt skulle en annan metod än den som tillämpats i detta arbete kunna testas. Kanske vore det intressant att undersöka hur gränssnittet skulle se ut och hur informationen kring de nya predikterade förseningarna förmedlas. Hur mycket information är rimligt att presentera, hur tidigt skall det presenteras och vem ska den delges? Detta är frågor som ett framtida projekt skulle kunna undersöka.

13.5. Vem riktar sig prognoserna till

Tåg är ett ypperligt kollektivt färdmedel för att transportera stora mängder gods och människor. Resor med tåg kan vara mycket bekväma, speciellt när avstånden och restiden är stora. Under tiden som tåget kör kan passagerare bland annat sträcka på benen, röra sig fritt i tåget, gå på toaletten, äta mat och sova. En del utnyttjar restiden till avkoppling och andra utnyttjar den genom att vara produktiva. Att kunna arbeta när man sitter på tåget är en möjlighet och bidrar till att utnyttja dötiden effektivt. Ytterligare stäcker sig järnvägen ofta genom öppna landskap vilket gör att passagerarna får möjlighet att njuta av fin natur. Tågstationerna är nästan alltid anlagda i centrum av städer eller i anslutning till andra viktiga platser. Detta gör att resenärer snabbt och bekvämt kan nå sina destinationer som ofta ligger stadsnära. Av dessa praktiska skäl kan det vara bättre att resa med tåg inom regionens gränser än att använda exempelvis bilen. Många människor väljer att resa med tåg till stor del på grund av den låga miljöpåverkan som tågen har jämfört med många andra färdmedel. Eftersom tågen har stor kapacitet blir de också mer energieffektiva och avger avsevärt mindre koldioxid per personkilometer än vad andra transportalternativ gör, särskilt flyg och bil. Dock är det viktigt att komma ihåg att människors preferenser är olika och varierar beroende på individuella behov så som resmål och tillgänglighet till olika former av transportsystem.

Tågprognoser tillgodoser behovet hos flertalet olika grupper av människor och yrkesgrupper. Genom pålitlig och relevant information bidrar de till att alla inblandade parter kan hantera resandet, och ta beslut, på ett mer effektivt vis. Det finns oändliga behov att tillgodose inom tågtrafiken, och ett sätt att uppfylla en del av dem är genom att ha tillförlitliga tågprognoser. Resenärer har ett stort behov av att, i ett tidigt skede, kunna planera sina resor för att kunna välja den mest lämpliga avgången för sin resa. De vill kunna planera sina övriga aktiviteter, anslutningar och klockslag för ankomst till slutdestination, fatta beslut om alternativa resvägar och kunna uppleva tillit och vara nöjda. Detta kan exempelvis göras genom att man får en indikation om när man behöver befinna sig på stationen för att undvika onödig väntetid. För resenärer är prognoserna även användbara då information om förseningar eller inställda tåg möjliggör för att kunna vidta nödvändiga åtgärder i sin reseplanering, antingen genom att hitta alternativa rutter eller fortsätta vänta på det aktuella tåget. Att ha tillförlitliga tågprognoser underlättar och hjälper resenärerna att tillgodose dessa, och fler andra, behov. Det ökar den upplevda kundservicen och resenärer kan känna sig tryggare med informationen. Behovet av tillit och nöjdhet gentemot tågoperatörerna kan därmed ökas, vilket är av stor vikt med tågbranschens, i en del fall, bristfälliga rykte.

För tågoperatörer är prognoser ett värdefullt verktyg eftersom de hjälper till och möjliggör att optimera tågtrafiken och planera resurserna på ett effektivare sätt. En stor fördel med att operatörerna har tillgång till aktuella prognoser är att de då kan justera tidtabellerna och hantera personal och utrustning på ett mer effektivt sätt. På så vis kan riskerna för försening minskas. Vidare öppnar prognoserna upp portarna för operatörer att ha mer insikt om trafikmönster, vilket kan ge dem möjlighet till att övergripande förbättra kundnöjdheten och tågtrafikens prestationsförmåga.

Även för stationspersonalen är tågprognoser av största vikt eftersom de förlitar sig på och utgår från den när de ger information till resenärer. För att kunna ge korrekt och uppdaterad information är det av stor vikt att prognoserna är tillförlitliga. Baserat på prognoserna kan personalen på stationerna ge information vid uppkomna förseningar, eventuella plattformsförändringar eller liknande förändringar som kan påverka resenärerna. Att stationspersonalen har aktuella prognoser hjälper till att förbättra kundservicen och kommunikationen på stationerna.

Ytterligare en yrkesgrupp som kan dra nytta av prognoserna är infrastrukturförvaltare. Genom analyser av prognoser och annan trafikdata kan trafikförvaltarna optimera underhålls- och uppgraderingsåtgärder genom att exempelvis identifiera flaskhalsar på järnvägen. Prognoserna bidrar även till att infrastrukturförvaltarna med hjälp av prognoserna kan planera var störst effekt och nytta kan genereras, så att kapacitetsförändringar och den framtida efterfrågan hos resenärerna kan bemötas.

Turister som besöker nya städer, länder eller platser och reser med tåg behöver tillgång till behjälpliga och lättbegripliga prognoser. Detta för att de enkelt ska kunna planera sina utflykter och sevärdheter och kunna vidta eventuella åtgärder om så krävs. Prognoser ger turister en uppfattning om avvikelser från tidtabellen och kan hjälpa till att bidra till ett mer effektivt tidsutnyttjande under vistelsen. För turister som reser med tåg är det viktigt att enkelt kunna få tillgång till information om tågens avgångar, ankomster, plattformsinformation och anslutande kollektivtrafik för att planera sina resrutter till olika sevärdheter.

I dagens stressiga samhälle är det lätt att ta för givet att alla har samma förutsättningar. Personer med funktionsvariation kan ha specifika behov gällande tågprognoser. De kan behöva få särskild information angående stationernas tillgänglighet gällande exempelvis hissar eller trappor och om det finns assistansmöjligheter på de berörda tågen. Personer med nedsatt hörsel förmåga behöver på stationsområdena särskilt tillgång till visuella meddelanden där den viktiga informationen finns tillgänglig.

Det är inte enbart enskilda individer som drar nytta av tågprognoser utan även företag. Logistikföretag som bedriver godstransporter med tåg eller på annat vis nyttjar tåg som en del av sin verksamhet kan ha fördel av tillgång till tågprognoser. Detta för att kunna optimera sina leveranser och hålla sina kunder uppdaterade om eventuella förseningar i leveranserna. Genom att företagen är medvetna om eventuella förseningar eller andra ändringar i tidtabellen kan de själva vidta nödvändiga åtgärder i förväg för att minimera inverkan i sin verksamhet.

Järnvägsnätet är ett mycket komplext system och väldigt känsligt för störningar. Då det finns ett oändligt antal variabler som kan påverka tågen till försening är en fördel att involvera erfarna experter. Dessa kan bistå med hjälp för att förstå och hantera komplexiteten. Detta kan stärka prognoserna och göra dem mer tillförlitliga. Men för att definitivt öka tillförlitligheten för järnvägsnätet hos resenärerna är det viktigt att kommunicera eventuella förseningar eller ändringar snabbt och tydligt. Genom att tillhandahålla uppdaterad information samt använda flertalet kommunikationsplattformar ökar resenärernas möjligheter att tillgå information och på så sätt planera eller hantera förändringar i tidtabellen.

14. Källhänvisning

14.1 Källförteckning

Andersson, E. V. (2014). Assessment of robustness in railway traffic timetables. Licentiatavhandling. Norrköping: Linköpings universitet.

Angel, T. (2020). Operativt beslutsstöd i realtid – Pilot C-DAS. JBS Järnvägsbranschens sverkanforum.

Bårström, S och Granbom, P. (2012). *Den svenska järnvägen*. Borlänge: Trafikverket.

Caimi, G., Kroon, L. & Liebchen, C. (2016). 'Models for railway timetable optimization: Applicability and applications in practice', *Journal of Rail Transport Planning & Management*, 6(4), ss. 285-312.

Carlo Meloni, Marco Pranzo, Marcella Samà, Risk of delay evaluation in real-time train scheduling with uncertain dwell times, *Transportation Research Part E: Logistics and Transportation Review*, Volume 152, 2021, 102366, ISSN 1366-5545.

Gorman, M.F. Statistical estimation of railroad congestion delay. *Transp. Res. Part E Logist. Transp. Rev.* 2009, 45, 446–456.

Goverde R M P. Punctuality of railway operations and timetable stability analysis[D]. TU Delft, Delft University of Technology, 2005.

Gummeson, M. (2019). Tillsammans för tåg i tid Resultatrapport 2019. TTT

Hansen, I. A. (2010). 'Railway network timetabling and dynamic traffic management', *International Journal of Civil Engineering*, 8(1), ss. 19-32.

Hellström, P. (2014). Problems in the integration of timetabling and train traffic control. Uppsala: Uppsala University.

Hooghiemstra, J. S. & Teunisse, M. J. G. (1998). The use of simulation in the planning of the dutch railway services. I 1998 Winter Simulation Conference. Proceedings (Cat. No. 98CH36274), volym 2, ss. 1139- 1145. Washington D.C., USA, december 1998.

Honauer, U., & Ödeen, S. (2019). Underhållsplan 2019-2022. Trafikverket

Huang, Ping, Zhongcan Li et al. (2021). 'Modeling train timetables as images: A cost-sensitive deep learning framework for delay propagation pattern recognition'. In: *Expert Systems with Applications* 177, p. 114996. issn: 0957-4174. doi: <https://doi.org/10.1016/j.eswa.2021.114996>. url: <https://www.sciencedirect.com/science/article/pii/S0957417421004371>.

Ivina, D., Palmqvist, C-W., Olsson, N., & Winslott Hiselius, L. (2021). Train delays due to trackwork in Sweden. Paper presented at 9th International Conference on Railway Operations Modelling and Analysis (ICROMA) – RailBeijing, Beijing, China.

<https://portal.research.lu.se/en/publications/train-delays-due-to-trackwork-in-sweden>

Jensen, L. W. (2015). Robustness indicators and capacity models for railway networks. Doktorsavhandling. Kongens Lyngby: Danmarks Tekniske Universitet.

Järnvägslagen (SFS 2004:519)

Khoshniyat, F. & Peterson, A. (2017). 'Improving train service reliability applying an effective timetable robustness strategy', *Journal of Intelligent Transportation Systems*, 21(6), ss. 525-543.

Klabes, S. G. (2010). Algorithmic railway capacity allocation in a competitive European railway market. Doktorsavhandling. Aachen: RWTH Aachen University.

Kosolsombat, Somkiat and Wasit Limprasert. (2017). 'Arrival Time Prediction and Train Tracking Analysis'. In: *Trends in Artificial Intelligence: PRICAI 2016*, pp. 170–177.

Lee, W.-H.; Yen, L.-H.; Chou, C.-M. A delay root cause discovery and timetable adjustment model for enhancing the punctuality of railway services. *Transp. Res. Part C Emerg. Technol.* 2016, 73, 49–64.

Larsen Sørskår, A (2022). *Predicting Train Departure Delays*.

Lindner, T. (2000). Train schedule optimization in public rail transport. Doktorsavhandling. Braunschweig: Technische Universität Braunschweig.

Markovi´c, N.; Milinkovi´c, S.; Tikhonov, K.S.; Schonfeld, P. Analyzing passenger train arrival delays with support vector regression. *Transp. Res. Part C Emerg. Technol.* 2015, 56, 251–262.

Mattsson, Per. (2021). *Svenska startupen vill revolutionera världens kollektivtrafik – får draghjälp av Microsoft*.

Milinković, S.; Marković, M.; Vesković, S.; Ivić, M.; Pavlović, N. A fuzzy Petri net model to estimate train delays. *Simul. Model. Pract. Theory* 2013, 33, 144–157.

Naturvårdsverket. (den 13 10 2022). *Hur bidrar Sverige till Parisavtalet?*. Hämtat från Naturvårdsverket.se:

<https://www.naturvardsverket.se/amnesomraden/klimatomställningen/det-globala-klimatarbetet/parisavtalet/hur-bidrar-sverige-till-parisavtalet/>

Oneto, L.; Buselli, I.; Lulli, A.; Canepa, R.; Petralli, S.; Anguita, D. A dynamic, interpretable, and robust hybrid data analytics system for train movements in large-scale railway networks. *Int. J. Data Sci. Anal.* 2020, 9, 95–111.

Oneto, L.; Fumeo, E.; Clerico, G.; Canepa, R.; Papa, F.; Dambra, C.; Mazzino, N.; Anguita, D. Dynamic Delay Predictions for Large-Scale Railway Networks: Deep and Shallow Extreme Learning Machines Tuned via Thresholdout. *IEEE Trans. Syst. Man Cybern. Syst.* 2017, 47, 2754–2767.

Pachl, J. (2018). *Railway operation and control*. Edition, 4. Publisher, VTD Rail Publishing., 2018. Technische Universität Braunschweig. ISBN, 0971991596, 9780971991590.

Palmqvist, C-W. (2019). *Delays and Timetabling for Passenger Trains*. [Doctoral Thesis (compilation), Transport and Roads]. Lund University Faculty of Engineering, Technology and Society, Transport and Roads, Lund, Sweden.

Parbo, Jens, Otto Anker Nielsen, and Carlo Giacomo Prato. (2016). “Passenger Perspectives in Railway Timetabling: A Literature Review.” *Transport Reviews* 36 (4): 500–526. <https://doi.org/10.1080/01441647.2015.1113574>

Skjøren, C. (2022). *Evaluation of train delay estimates*.

Spanninger, Thomas et al. (2022). ‘A review of train delay prediction approaches’. In: *Journal of Rail Transport Planning Management* 22.

Stambanan. (2015). *Södra stambanan Malmö – Stockholm Sveriges viktigaste järnväg*

<http://stambanan.com/wp-content/uploads/2017/06/broschyr15juni.pdf>

Ödebrink, Carina. (2022). *Ny järnväg för Sverige*. Riksdagen.

https://www.riksdagen.se/sv/dokument-lagar/dokument/skriftlig-fraga/ny-jarnvag-for-sverige_HA1141

Trafikverket. (2016). *Tågplan – att skapa tidtabeller för tåg*. Trafikverket.

<https://bransch.trafikverket.se/for-dig-i-branschen/jarnvag/tagplan-att-skapa-tidtabeller-for-tag/>

Trafikverket. (2020a). *AIRT (AI-baserad Realtidsprognostisering av Trafikinformation)*.

<https://fudinfo.trafikverket.se/fudinfoexternWebb/pages/ProjektVisaNy.aspx?ProjektId=4586#:~:text=Projekt%20AIRT%2C%20AI%2Dbaserad%20Realtidsprognostisering,p%3%A5%20s%3%A5%20s%3%A4tt%20f%3%B6rb%3%A4ttra%20trafikinformationen.>

Trafikverket. (2020b). *Trafikverkets verksamhetsplan 2021-2023*. Trafikverket.

<https://bransch.trafikverket.se/contentassets/442e0ad2c9e749aa832ae0c0f612dd15/trafikverkets20verksamhetsplan2021-2023.pdf>

Trafikverket. (2022a). *Järnkoll – fakta om svensk järnväg*

[https://www.trafikverket.se/resa-och-trafik/jarnvag/jarnkoll--fakta-om-svensk-jarnvag/#:~:text=En%20vanlig%20dag%20\(2021\)&text=En%20vanlig%20dag%20reser%20290,k%3%B6rde%20var%20370%20000%20kilometer.](https://www.trafikverket.se/resa-och-trafik/jarnvag/jarnkoll--fakta-om-svensk-jarnvag/#:~:text=En%20vanlig%20dag%20(2021)&text=En%20vanlig%20dag%20reser%20290,k%3%B6rde%20var%20370%20000%20kilometer.)

Trafikverket. (2022b). *Kompetenscenter för C-DAS (Connected Driver Advisory System)*.

<https://bransch.trafikverket.se/for-dig-i-branschen/samarbete-med-branschen/jarnvagsbranschens-samverkansforum/Tillsammans-for-tag-i-tid/kompetenscenter-for-c-das-connected-driver-advisory-system/>

Trafikverket. (2022c). *Punktligheten tillbaka på rätt sida om 90-strecket*. Trafikverket.

<https://www.trafikverket.se/om-oss/nyheter/nationella-nyheter/2022/augusti/punktligheten-tillbaka-pa-ratt-sida-om-90-strecket/>

Trafikverket. (2023). *AIRT (AI-baserad Realtidsprognostisering av Trafikinformation)*.

<https://fudinfo.trafikverket.se/fudinfoexternWebb/pages/ProjektVisaNy.aspx?ProjektId=4586>

United Nations. (2021). *Fact sheet climate change*. United Nations.

Yaghini, M.; Khoshraftar, M.M.; Seyedabadi, M. Railway Passenger Train Delay Prediction via Neural Network Model. *J. Adv. Transp.* 2013, 47, 355–368.

Yuan, J., R.M.P. Goverde and I.A. Hansen (2002). 'Propagation of train delays in stations'. In: *Computers in Railways VIII* 61, p. 10. doi: 10.2495/CR020961. url: <https://www.witpress.com/elibrary/wit-transactions-on-the-built-environment/61/331>.

14.2. Figurförteckning

Trafikverket, 2016. Tågplan – att skapa tidtabeller för tåg. <https://www.trafikverket.se/for-dig-i-branschen/jarnvag/tagplan-att-skapa-tidtabellerfor-tag/> [Hämtad 2023-03-16]

Trafikverket, 2020. Första testet av C-DAS tillsammans med SJ genomfört. <https://bransch.trafikverket.se/om-oss/aktuellt-for-dig-i-branschen3/aktuellt-for-dig-i-branschen/2020-05/forsta-testet-av-c-das-tillsammans-med-sj-genomfort/> [Hämtad 2023-03-20]

15. Bilaga

`select`

Källa, FRAMSYNTHET,

Precision_MEDEL, Precision_STDEV, Precision_MAE, Precision_RMSE, n, punktlighetsläge

`from`

(`select` källa, punktlighetsläge,

`ROUND`(Framsynthet/5,0)*5 `AS` FRAMSYNTHET,

`avg`(precision) `as` Precision_MEDEL,

`avg`(`abs`(precision)) `as` Precision_MAE,

`round`(`sqrt`(`avg`(`power`(precision,2))),0) `as` Precision_RMSE,

`round`(`stdev`(`abs`(precision)),0) `as` Precision_STDEV, `count`(*) `as` n

`from` predictions..exjobbsdata_3

`where` framsynthet > 0

`group by` källa,punktlighetsläge, `ROUND`(Framsynthet/5,0)*5) a

`where` n > 100 and FRAMSYNTHET <= 120

`order by` Källa, punktlighetsläge, FRAMSYNTHET `DESC`