

Punktlighet för pendeltågstrafik i Stockholm

- Analys av förseningsdata och fordonsdata



LUNDS
UNIVERSITET

Lunds Tekniska Högskola

LTH Ingenjörshögskolan vid Campus Helsingborg
Teknik och samhälle

Examensarbete:
Hendrik Co

© Copyright Hendrik Co

LTH Ingenjörshögskolan vid Campus Helsingborg
Lunds universitet
Box 882
251 08 Helsingborg

LTH School of Engineering
Lund University
Box 882
SE-251 08 Helsingborg
Sweden

Tryckt i Sverige
Media-Tryck
Biblioteksdirektionen
Lunds universitet
Lund 2021

Sammanfattning

Arbetet ämnar till att undersöka eventuella samband mellan fordonsdata och förseningsdata för att fastställa om fordonsdata kan användas för att förutse förseningar. Fordonsdata som genererats till följd av behov för kontrollundersökning, i.e. felkoder, har kombinerats med förseningsdata från Stockholms pendeltåg mellan 2017-01-01 och 2018-12-08. Statistiska beräkningsmetoder har använts för att avgöra om det finns ett samband. Oddskvot samt konfidensintervall beräknades för att vidare möjliggöra en analys. Analysen delades in i 2 delar varav första delen undersöker om felkoder generellt, oavsett typ, har något samband med förseningar. Andra delen fokuserar på specifika felkoder för att försöka identifiera felkoder som möjligtvis har ett samband med förseningar. Under den studerade tidsperioden genererades 811 olika felkoder. Detta innebär att samtliga 811 felkoder analyseras individuellt.

Resultatet för felkoder generellt visar ett signifikant samband mellan felkoder och gångtidförseningar men icke-signifikant samband mellan felkoder och uppehålltidförseningar.

Analysens andra del, specifika felkoder, uppvisade likartade resultat som den första delen. För specifika felkoder visar resultaten att 78 olika felkoder har statistiskt signifikanta samband med gångtidförseningar. Upphållsförseningar visade sig ha färre signifikanta felkoder med endast 9 olika felkoder samt att sambanden för samtliga 9 felkoder var svaga jämfört med gångtidförseningar.

Eftersom resultaten för analysens första del visade signifikant samband mellan felkoder och gångtidförseningar men icke-signifikant samband mellan felkoder och uppehålltidförseningar var resultaten för analysens andra del föga förvånande. Gångtidförseningar visade sig ha en tydligt starkare koppling med felkoder än uppehålltidförseningar. Anledningen till att uppehålltidförseningar har en svag till ingen koppling med felkoder kan möjligen förklaras av att uppehålltidförseningar ofta sker i stationer och därmed har en större koppling med den mänskliga faktorn, bland annat tidtabellens planering, tågpersonalens beteende, ombord- och avstigning av passagerare. Sambandet mellan felkoder och gångtidförseningar kan möjligen bero på bland annat att delar av tåget i högre grad utsätts för slitage och stress under gång från infrastruktur samt väder och temperatur som i sin tur ger upphov till uppkomsten av felkoder. Det innebär att gångtidförseningarna möjligen beror mindre på den mänskliga faktorn och till en större del på ett problem med fordonet, något som kan leda till att en felkod genereras.

Nyckelord: Förseningsdata, fordonsdata, felkod, järnväg, försening, tåg

Abstract

This thesis intends to investigate possible connections between vehicle data and delay data to determine whether vehicle data can be used to predict delays. Vehicle data generated as result of the need for inspection, that is, event codes, has been combined with delay data from Stockholm's commuter trains between 2017 -01-01 and 2018-12-08. Statistical calculation methods were used to determine possible correlations. Odds ratio and confidence intervals were calculated to further enable an analysis. The analysis was divided in two parts of which the first part examines whether event codes in general, regardless of type, have any correlation with delays. The second part focuses on specific event codes to identify event codes that may be related to delays. During the studied period 811 different error codes were generated. This means that all 811 error codes are analysed individually.

The result for event codes generally shows a significant correlation between event codes and runtime delays but a non-significant correlation between event codes and dwell time delays.

The second part of the analysis, specific event codes, displayed similar results of that of the first part of the analysis. For specific event codes, the results show that 78 different event codes have statistically significant correlations with runtime delays. Dwell time delays were found to have fewer significant event codes with only 9 different event codes as well as weak correlations for all 9 event codes compared to those of run time delays.

As the results of the first part of the analysis showed a significant correlation between event codes and runtime delays but a non-significant correlation between event codes and dwell time delays, the subsequent results of the second part of the analysis were not unexpected. Run time delays were found to have a clearly stronger correlation with event codes than dwell time delays. The reason why dwell time delays have a weak or no correlation with event codes can possibly be explained by the fact that dwell time delays often occur in stations and thus have a greater correlation with the human factor including timetable planning, train staff behaviour, boarding and disembarking passengers. The correlation between event codes and run time delays may possibly be due to parts of the train being exposed to a greater degree to wear and tear during run time from infrastructure, weather and temperature, which in turn generate event codes. This means that the run time delays may be due less to the human factor and to a greater extent to a problem with the vehicle, which can lead to an event code being generated.

Keywords: Delay data, vehicle data, event code, railway, delay, train

Förord

Detta examensarbete är den avslutande delen av utbildningen Högskoleingenjör i byggteknik med inriktning järnvägsteknik vid Lunds Teknisk Högskola, Campus Helsingborg. Utbildningen omfattar 180 högskolepoäng, varav examensarbetet omfattar 22,5 högskolepoäng.

Jag vill tacka min examinator Andreas Persson för att ha introducerat arbetet och hänvisat mig till Carl-William Palmqvist.

Ett stort tack till min handledare Carl-William Palmqvist som med sin kompetens, tid och råd hjälpt mig genom flertalet utmanande moment under arbetets gång.

Slutligen vill jag även tacka mina hjälpsamma vänner. Stort tack till Hawoay Tao som med sina programmeringskunskaper gav mig råd och hjälp som i sin tur förenklade inlärningsprocessen för SQL. Ett särskilt tack till Ali Haghparast för hjälp med korrekturläsning och goda råd genom hela arbetsprocessen.

Innehållsförteckning

1	Bakgrund	2
1.1	Syfte	4
1.1	Problemformulering	4
1.2	Avgränsningar	4
2	Teknisk bakgrund	5
2.1	Fakta om pendeltågstrafiken i Stockholm län	5
2.2	Förseningsdata och fordonsdata	6
3	Metod	8
3.1	Hantering och filtrering av data	8
3.2	Sammankoppling av data	8
3.3	Oddsquot och konfidensintervall	9
4	Förseningar och fordonsfel i pendeltågstrafiken	13
4.1	Förseningar i pendeltågstrafiken	13
4.2	Felkoder i pendeltågstrafiken	13
5	Resultat	16
5.1	Samband mellan fordonsfel och förseningar	16
5.1.1	Felkoder, generellt	16
5.1.2	Uppehålltidsförsening och Gångtidsförsening	16
5.2	Felkoder, specifika	16
5.2.1	Uppehålltidsförseningar	16
5.2.2	Gångtidsförsening	18
6	Diskussion	20
6.1	Metoddiskussion	20
6.2	Felkoder, generellt	21
6.3	Felkoder, specifika	21
7	Slutsats	24
8	Felkällor och framtida rekommendationer	25
9	Referenser	26
10	Bilagor	28

Inledning

Med den globala och nationella populationens ökade intresse och medvetenhet kring miljöfrågor har miljövänliga transportmedel samt kollektivtrafik blivit en alltmer dagsaktuell fråga. Fler människor väljer, om möjligt, miljövänligare alternativ även när dessa innebär justeringar inom livsstilen (SJ klimatrappport, 2019). För att förbli en attraktiv konkurrent till fordonstrafiken föreligger behov för pendeltågen att kunna passa in i resenärernas mest essentiella delar av livsstilen bland annat förutsägbarhet gällande tid. Kunden måste kunna betrakta tåg som likvärdigt alternativ till bil gällande punktlighet. Sverige har en befolkning på ca 10,3 miljoner, varav Stockholms län utgör ca 2,3 miljoner (SCB, 2020). Stockholmsregionen omfattar hela Stockholms län och delar av Södermanlands län och Uppsala län. Med ökande befolkning över tid sätter arbets- och utbildningspendling högre krav på tågens punktlighet. Även ökad distans på resor resulterar i större krav på hållbarhet för transportmedel. I en rapport från Transportstyrelsen från 2017 om resenärers syn på järnvägsmarknaden framkom punktlighet som en av de viktigaste faktorerna som påverkar resenärens nöjdhet (Transportstyrelsen, 2017). Punktlighet utgör en ytterst relevant faktor samt konstant utmaning inom den svenska järnvägen. Bättre förståelse av förseningar inom pendeltågstrafiken i storstäder underlättar troligen identifiering av åtgärder för förbättring av järnvägen. Järnvägstrafikens punktlighet påverkas av flertal olika faktorer. Förseningar kan uppstå i stationen, exempelvis i samband med ombord- och avstigning av passagerare eller till följd av passagerarbelastning på tåget eller ute i spåret på grund av infrastrukturfel, obehöriga i spår, olyckor och andra incidenter. Genom studier och planering kan förseningars orsaker och incidens kartläggas och därmed genom adekvata åtgärder förebyggas för ett effektivare järnvägssystem. Flertalet studier har gjorts inom tågförsening och befintligt arbete ämnar till att använda statistiska metoder och en uppsättning fordonsdata från pendeltågen i Stockholm för att undersöka korrelationer mellan förseningar och felkoder och därmed utröna orsaker till förseningar. Aktuell metod är inte beprövad i tidigare studier och utgör därmed en ny infallsvinkel inom ämnet som förhoppningsvis resulterar i ökad förståelse kring punktlighet.

1 Bakgrund

För att höja resenärernas intresse för pendeltåg finns det några faktorer som behöver hålla en hög standard. I en marknadsundersökning gjord av Transportstyrelsen fann man att två av de viktigaste faktorerna är säkerhet och punktlighet (Transportstyrelsen, 2017), varav i det här projektet ligger punktlighet i fokus. Det har gjorts ett flertal studier inom området och för att förbättra förståelsen för förseningar är det nödvändigt med fortsatta studier.

Branschsamarbetet TTT (Tillsammans för tåg i tid) startades år 2013 för att uppnå djupare insikt inom frågor om punktlighet. Målet är att öka punktlighet och förtroende för tåg som transportmedel. Ett av målen är att 95% av alla tåg ska ankomma till slutstation inom en marginal på 5 minuter efter utsatt tid. Från TTTs resultatrapport (2019) syns det att punktligheten i pendeltågstrafiken är idag relativt hög. Den genomsnittliga punktligheten från 2013 till 2018 är låg på 94,5%, vilket fortfarande ligger en bit under målet 95% och visar att fortsatt förbättring fortfarande behövs för att inte tappa resenärer till andra transportmedel då nöjdare resenärer kan leda till ett större intresse för järnvägen som ett transportalternativ.

En rapport från 2019 rörande samarbetet mellan MTR, Trafikverket och Trafikförvaltningen går vidare med målet att identifiera åtgärder som fort kan ge en positiv effekt på punktligheten för pendeltågsresenärerna. De tre områden som var i fokus var förbättrad infrastruktur, förbättrad hantering av trafiken under och efter störningar samt en utvecklad och mer tillförlitlig trafikinformation (Trafikverket, Trafikförvaltningen, MTR, 2019). Bland åtgärderna var att utöka förebyggande underhåll på spårväxlar och spårledningar, utnyttja scenarioplaner för att effektivare hantera störningar när de väl dyker upp, snabbare återhämtning vid mindre störningar med hjälp av nya operativa riktlinjer för att vid störningar vända tågen på bästa sätt. Åtgärderna visade sig ge en positiv effekt på punktligheten.

Orsaker som kan påverka förseningstiden är många och de påverkar punktligheten i olika grader. Exempel på dessa är infrastrukturfel (spårväxelfel, elfel, spårfel etcetera), obehöriga i spår, olyckor och tillbud, yttre faktorer som exempelvis snö och is. Palmqvist, Olsson och Hiselius (2017) analyserade olika faktorer som påverkar punktligheten och fann att höga och låga temperaturer kan påverka punktligheten signifikant samt att frekventa tåginteraktioner även påverkar punktligheten negativt.

Ett annat område där förseningar ofta uppstår är på stationer. Palmqvist, Tomii och Ochiai (2020) studerade uppehållstidsförseningar i Stockholm och Tokyo med hjälp av passagerardata. Studien nämner att små förseningar som är 5 min

och under står för 96% respektive 97% av alla förseningar i Stockholm och Tokyo. 91% respektive 88% av den totala förseningstiden i Stockholm och Tokyo uppstår i stationer där orsaken till stor del beror på förlängda uppehållstider i stationerna. Studien visar att ombord- och avstigning av passagerare är en viktig faktor som bidrar till förlängda uppehållstider. Olsson och Haugland (2004) använder empiriska data från studier i Norge för att undersöka faktorer som påverkar punktlighet. I studien finner de att punktlighet i lokala och regionala tåg i överbelastade områden har ett samband med ombord- och avstigning av passagerare samt hanteringen av dessa. På enkelspåriga linjer i mindre belastade områden verkar hanteringen av plankorsningar vara en faktor som kan påverka punktligheten. Gysin (2018) använder passagerar- och uppehållstidsdata för en station i Zurich, Schweiz för att studera faktorer och möjliga orsaker som sammanhänger med uppehållstider i stationer. I studien finner man bland annat att långa uppehållstider, och antalet passagerare påverkar förseningar.

Ceder och Hassold (2015) gjorde en analys på järnvägsverksamheten i Auckland, Nya Zeeland och fann att tågpersonalens beteende vid schemaanläggningsproblem kunde leda till ökade väntetider för passagerare som i sin tur resulterade i signifikant påverkan av järnvägens tillförlitlighet, dock kan detta effektiviseras genom förbättringar i besättningsschemat. Nie och Hansen (2005) studerade förhållandet mellan planerad och faktisk förfarande genom att analysera tågdetekteringsdata och dess påverkan på bland annat punktlighet. De fann att tågpersonalens beteende samt tåghinder som orsakas av ruttkonflikter som till följd av schemaanläggning påverkar uppehållstidsförseningar på plattformsspår.

Palmqvist, Olsson och Hiselius (2017) gjorde en studie med syftet att studera och kvantifiera olika variabler som påverkar punktligheten i tågtrafiken. Variablerna var väder-, tidtabell, operativ- och infrastrukturelaterade. Studien nämner att väldigt höga och låga temperaturer kan ha en stor påverkan på punktligheten. Studien nämner även att mindre ändringar på tidtabellen kan leda till ökad punktlighet. Exempel på ändringar är minimering av interaktioner mellan tåg och reducering av antal tåg med negativa marginaler. Palmqvist, Olsson och Hiselius (2017) gjorde ytterligare en studie samma år vars fokus låg på tidtabellens planering. Studien visar att tidtabellen överlag inte är den största bidragande faktorn till förseningar men har en signifikant påverkan på förseningar. Buchmueller, Weidmann och Nash (2008) visade att det är flera delprocesser bland annat öppning och stängning av dörrar, ombord- och avstigning, avsändning som påverkar uppehållstidsförseningar. Vidare kunde informationen nyttjas för att skapa ett beräkningsverktyg för att utforma mer precisa tidtabeller.

1.1 Syfte

Syftet är att använda statistiska metoder och en uppsättning fordonsdata samt förseningsdata från pendeltågen i Stockholm för att undersöka eventuella samband mellan förseningar och felkoder och därmed öka kunskapen kring förseningar.

1.1 Problemformulering

- Finns det något samband mellan felkoder och förseningar vid generell analys av felkoder?
- Finns det något samband mellan felkoder och förseningar vid analys av individuella/specifika felkoder?
- Är det möjligt att genom analys av resultaten förklara påvisade samband?

1.2 Avgränsningar

Arbetet är begränsat till pendeltågen inom Stockholm från tidsperioden 2017-01-01 till 2018-12-08.

2 Teknisk bakgrund

2.1 Fakta om pendeltågstrafiken i Stockholm län

Stockholms län har en befolkning på ca 2,3 miljoner invånare och är Sveriges mest folkrika samt folktäta län (SCB, 2020). Pendeltågssystemet har en sammansatt längd på 241 km (mtrnordic.se) och består av åtta linjer och 55 stationer från Nynäshamn och Gnesta i söder till Bålsta och Uppsala i norr (tunnelbanakarta.se). Pendeltågen i Stockholm har ca 500 avgångar varje dag (SLL, 2019) och antalet påstigande människor är ca 370 000 varje dag (SL, 2018).

Wigren och Abelin (2019) analyserade störningsorsaker och punktlighet i Stockholmsregionen och studien visar att pendeltågstrafiken i Stockholmsregionen utgör 40% av alla tågkilometer inom pendeltågstrafik i Sverige. År 2019 hade pendeltåget i Stockholm, likt tidigare år, fortsatt det lägsta betyget för nöjdhet av alla transportalternativ gällande punktlighet. Endast 54% av alla pendelresenärer i Stockholm län var nöjda med tidhållningen jämfört med 81% och 80%, tunnelbana respektive lokalbana. (Trafikförvaltningen Stockholm Läns Landsting, 2019).



Figur 1 Linjenätskarta för Stockholms pendeltåg

2.2 Förseningsdata och fordonsdata

Data för arbetet kommer från Trafikverkets uppföljningssystem Lupp som används för att ta fram statistik om punktlighet och störningar. Datat är i form av flertalet tabeller och innehåller en stor mängd data som omfattar information om hela Sveriges tågssystem. Förseningsdata innehåller 15 882 801 observationer av tågrörelser från hela Sverige från 2017 – 2018 - 12 – 08 och fordonsdata innehåller 4 346 102 registrerade felkoder från tåg inom Stockholm från 2017 – 2019.

Förseningsdata

Förseningsdata innehåller 15 882 801 observationer av tågrörelser från hela Sverige. Varje rad i tabellen innehåller information om en registrerad observation av en tågrörelse på en delsträcka. Informationen är uppdelad i kolumner för bland annat tågnummer, datum, tågkategori, avgång- och ankomstation, tidsstämplar i form av planerad och faktiska avgång, ankomst-, uppehåll- och gångtid från en station till en annan. Det finns även kolumner med uträknade gångtidsförseningar och uppehålltidsförseningar för varje observation som anges i minuter. Notera att detta är registrerade *merförseningar*.

Gångtidsförsening sker när tåget inte håller den planerade gångtiden.

Exempel: Ett tåg ska köra från station A till B. När tåget har lämnat station A så har den en planerad gångtid på 2 minuter för att anlända till station B. Men den faktiska gångtiden är 3 minuter. Detta innebär att tåget har en gångtidsförsening på 1 minut.

Upphålltidsförsening sker när tåget drar över den planerade uppehållstiden.

Exempel: Ett tåg har ett uppehåll i station A. Den planerade uppehållstiden är 1 minut. Men den faktiska uppehållstiden är 3 minuter. Detta innebär att tåget har en uppehålltidsförsening på 2 minuter. Tabell 1 visar ett exempel på en rad i tabellen, det vill säga en registrerad observation av en tågrörelse.

Tabell 1 Exempel på en rad i tabellen för förseningsdata

Tågnummer	Datum	Avgångsstation	Ankomstation	Planerad uppehållstid	Faktisk uppehållstid	Planerad gångtid	Faktisk gångtid	Gångtidsförsening	Upphålltidsförsening
2872	2018-11-22	Trångsund	Farstastrand	2018-11-22 13:17:00.000	2018-11-22 13:17:00.000	2018-11-22 13:18:00.000	2018-11-22 13:18:00.000	0	0

Fordonsdata

När ett fel uppstår i ett tåg genereras och registreras en felkod. Informationen kan avläsas i tabellformat som kallas för fordonsdata. Under den avgränsade tidsperioden genererades 811 olika typer av felkoder. Felkoder genereras av

olika skäl i fordonen och har olika innebörd. De är klassade efter allvarlighetsgraderna A, B, C och M varav A är allvarligast följ av B, C och sist M. Att en felkod genereras kan bero på flera skäl. Exempelvis att en specifik del i fordonet som felkoden är anknuten till är i behov av en kontrollundersökning eller information om ändringar i fordonet.

Fordonsdata innehåller 4 346 102 registrerade felkoder från tåg inom Stockholm. Varje rad i tabellen innehåller information om en registrerad felkod bland annat typ av felkod, tågnummer, tidsstämpel, GPS position var fordonet befann sig när felkoden genererades. Tabell 2 visar ett exempel på en rad i tabellen, det vill säga en registrerad felkod.

Tabell 2 Exempel på en rad i tabellen för fordonsdata

Felkod	Tågnummer	Tidsstämpel	GPS position latitud	GPS position longitud
56022	2872	2018-04-21 20:35:19.000	59.47607	17.74629

3 Metod

För att svara på arbetets frågeställning gjordes uppdelning i följande steg:

1. Behandling av all data genom sortering, filtrering
2. Sammankoppling av data
3. Beräkning av data med statistiska beräkningsmetoder, oddskvot och konfidensintervall.

3.1 Hantering och filtrering av data

All data hanterades i Microsoft SQL Management Studio. För att analysera all data krävdes förberedelse i form av sortering och filtrering. Först isolerades data från Stockholms pendeltåg från tidsperioden 2017-01-01 till 2018-12-08. Den tillgängliga fordonsdata sträcker sig från 2017 t.o.m. 2019 men eftersom den tillgängliga förseningsdata var begränsad till tidsperioden 2017-01-01 till 2018-12-08 så anpassades fordonsdata till samma period. Därefter filtrerades rader med felregistrerade eller saknade värden bort och resultatet var att 2 003 216 rader av förseningsdata och 1 703 766 rader av fordonsdata kvarstod. Vidare delades förseningarna upp i gång- och uppehållstidsförsening. Resultatet var 1 848 817 rader av gångtider och 1 767 311 rader av uppehållstider.

Aggregation

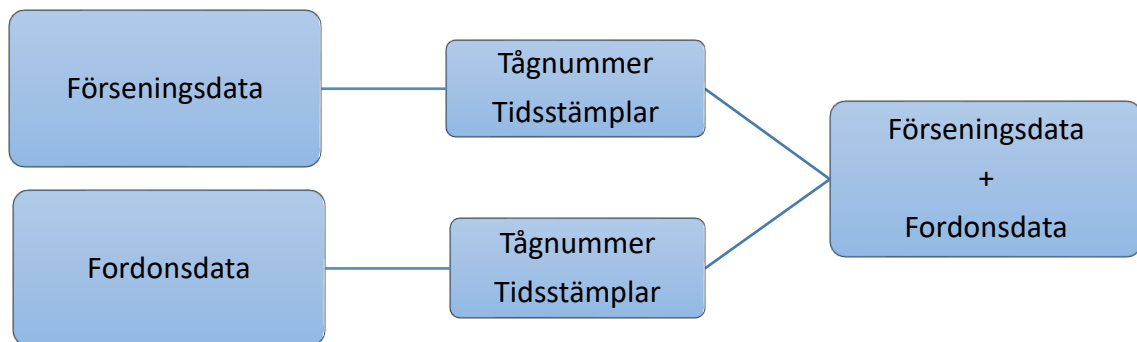
Syftet med arbetet är att undersöka om felkoder har någon korrelation med förseningar. I detta arbete ligger intresset i att veta om en försening har skett eller inte. Det innebär att storleken av en försening inte är i fokus. Oavsett storleken av en försening så räknas det som en försening i samma skala som alla andra förseningar. Alla punktliga och tidiga tåg gavs värdet 0 och icke punktliga tåg gavs värdet 1. D.v.s. både en försening på 4 minuter och en försening på 1 minut ges värdet 1 och anses vara en försening av samma skala.

3.2 Sammankoppling av data

Syftet med arbetet är att undersöka om det finns något samband mellan felkoder och förseningar. För att vidare kunna analysera all data behövs det därför vetskap om vilka tågrörelser som har felkoder. Detta görs genom att koppla samman tabellerna för förseningsdata och fordonsdata. På så sätt kan tågrörelser som har en eller flera felkoder identifieras.

Genom tågnummer samt tidsstämplar kopplades förseningsdata ihop med fordonsdata. Resultatet är att 240 811 observationer av gångtider lyckades matcha med 520 236 felkoder och 217 084 observationer av uppehållstider

lyckades matcha med 468 412 felkoder. Detta innebär att 240 811 gångtider och 217 084 uppehållstider har en eller flera felkoder.



Figur 2 Process för att sammankoppla data i SQL

3.3 Oddskvot och konfidensintervall

Efter matchning av data kan man med resultatet utnyttja statistiska beräkningsmetoder för att analysera om fordonsdata har något samband med förseningsdata. Med gång- och uppehållstidsförsening uppdelade i två olika sektioner kan man räkna ut värden för vardera. Den statistiska beräkningsmetoden som valdes för att undersöka om det finns någon samband mellan faktorerna är oddskvot, OK. Genom att skapa en 2x2 matris med den data som ska undersökas (se figur 3) kan eventuella samband mellan felkoder och förseningar undersökas. Detta avgörs genom att se hur värdet av OK ser ut.

Analysen delas in i två delar. Första delen av analysen undersöker hur oddsen för förseningar ökar om en felkod förekommer. I denna sektion står inte en specifik felkod i fokus. Fokus ligger på att undersöka om oddsen för förseningar ökar när en felkod generellt förekommer. Om tågobservationen har flera eller en specifik typ av felkod är inte av intresse eftersom syftet är endast att veta om det finns en felkod eller inte för att undersöka om det finns något samband mellan felkod och försening. Tabell 6 visar en rad i tabellen som berättar om tågrörelsen har en felkod eller ej. Andra delen av analysen undersöker varje specifik felkod för sig för att ta reda på om oddsen för förseningar ökar beroende på vilken specifik felkod som förekommer. Detta innebär att samtliga 811 felkoder analyseras individuellt. Tabell 7 och 8 visar hur tabellerna ser ut när det finns av en eller flera specifika felkoder.

	Felkoder (I)	Inga felkoder (C)
Förseningar (E)	Förseningar med felkoder (IE)	Förseningar utan felkoder (CE)
Icke förseningar (N)	Icke - förseningar med felkoder (IN)	Icke - förseningar utan felkoder (CN)

Figur 3 2x2 matris för att räkna ut oddskvot

OK är en statistisk beräkningsmetod som kan användas för att undersöka om det finns en koppling mellan två variabler. I detta fall försening och felkod. Värdet för OK berättar hur mycket högre oddsen är att förseningar sker när en felkod genereras under en tågrörelse jämfört med om en felkod inte hade genererats.

OK är kvoten av två odds, i detta fall oddsen mellan försening och icke försening när det finns en felkod och oddsen mellan försening och icke försening när det inte finns någon felkod. När OK är lika med 1 innebär det att en felkod inte har något samband med en försening, nollhypotes. När värdet av OK är mindre än 1 innebär det att en felkod inte har något samband till en försening. När OK är större än 1 så innebär det att en felkod har en samband med en försening. Ju större värdet är desto större är oddsen att det sker en försening när en felkod genereras jämfört med om felkoden inte hade genererats (Szumilas, M. 2010). För uträkning av OK används följande ekvation:

$$OK = \frac{\frac{IE}{IN}}{\frac{CE}{CN}} = \frac{IE * CN}{IN * CE}$$

För att avgöra om värdet för OK är statistisk signifikant så kontrolleras att värdet ligger inom ett konfidensintervall. Vidare får konfidensintervallet inte innehålla värdet för nollhypotes, dvs där värdet är lika med 1 (Tenny S; Hoffman, MR. 2020). Det kritiska värdet, $Z_{critical}$, väljs efter storleken på konfidensintervallet. Eftersom urvalet av koder är stort och för att vara säkra på att resultaten är statistiskt signifikanta så valdes användning av ett brett konfidensintervall på 99,9%. Jämfört med ett smalare konfidensintervall, exempelvis 95%, då det finns 5% chans att ett värde som hamnar inom intervallet är där på grund av slumpen så kan man med ett bredare konfidensintervall på 99,9% vara säkrare på att värdena inom intervallet faktiskt är statistiskt signifikanta och inte där av slump. För uträkning av konfidensintervallet för OK används följande ekvation:

$$\text{Konfidensintervall} = OK \pm e^{Z_{critical} SF(\log(OK))}$$

För uträkning av standardfelet, SF, används följande ekvation:

$$SF(\log(\text{OK})) = \sqrt{\frac{1}{IE} + \frac{1}{IN} + \frac{1}{CE} + \frac{1}{CN}}$$

3.4 Jämförelsevärde

Då de signifikanta felkoderna kan vara många i antalet samt att vissa felkoder uppträder oftare men har låga oddskvoter medan andra har högre oddskvoter men uppträder mer sällan så kan det vara svårt att identifiera vilka felkoder som har en större påverkan på förseningar i helhet. Därför har ett värde skapats för att göra det lättare att identifiera felkoder som möjligen har en större påverkan på förseningar i helhet och som bör prioriteras först. Värdet gör det möjligt att jämföra felkoder som uppträder sällan men har stor oddskvot med felkoder som uppträder ofta men har en lägre oddskvot och visar ”*antal extra förseningar*” som orsakas när en viss felkod genereras jämfört om felkoden inte hade genererats. Tabellerna för resultaten kommer att rangordnas från störst till minst efter detta jämförelsevärde, *antal extra förseningar*.

De två variabler som behöver jämföras är *antal gånger* en specifik felkod har genererats och *storleken på oddskvoten* på samma felkod. Dessa variabler behöver därför vara med i ekvationen. För att kunna jämföra variablerna behövs ytterligare en variabel som är gemensam för alla felkoder, vilket är *oddsen för de totala förseningarna*. Med hjälp av dessa tre variabler kan antalet förseningar räknas ut när de blir exponerade jämfört med när de inte blir exponerade för en felkod. För uträkning av värdet används följande variabler:

$$\text{Odds för de totala förseningarna} = \frac{\text{Antal Förseningar}}{\text{Antal icke-förseningar}} = \frac{IE+CE}{IN+CN} = X$$

Totala antalet av en specifik felkod = F

Oddskvot för en specifik felkod = OK

Genom att multiplicera oddsens för de totala förseningarna, X, med det totala antalet av en specifik felkod, F, som har genererats så är det möjligt att räkna ut antalet förseningar som felkoden har genererats i samband med. Med andra ord, detta är förseningar som inte har påverkats av någon specifik felkod. För uträkning av värdet används följande ekvation:

$$X * F$$

Om oddskvoten är statistiskt signifikant är värdet alltid större än 1. Det betyder att när en felkod genereras i samband med en försening så är oddsen för att en försening sker, alltid större jämfört med när en felkod inte hade genererats. Genom att multiplicera tidigare ekvation med oddskvoten, OK , så exponeras förseningarna för en viss felkod och då visar värdet hur mycket en specifik felkod har ökat antalet förseningar med. En statistisk signifikant oddskvot kommer alltid att vara större än 1 och därmed öka antalet förseningar. Med andra ord, detta är förseningar som har påverkats av en specifik felkod. För uträkning av värdet används följande ekvation:

$$X * F * OK$$

Genom att subtrahera förseningarna som har påverkas av en felkod med förseningarna utan påverkan av en felkod så beräknas ett nytt värde som visar antalet extra förseningar som felkoden har orsakat jämfört om felkoden inte hade varit där. För uträkning av värdet används följande ekvation:

$$X * F * OK - X * F = \text{Antal extra förseningar}$$

4 Förseningar och fordonsfel i pendeltågstrafiken

4.1 Förseningar i pendeltågstrafiken

Efter sortering och filtrering är förseningsdata redo för att kopplas ihop med fordonsdata. Tabell 3 visar de filtrerade siffrorna för förseningsdata uppdelad i antal tågobservationer och antalet av dessa som är i tid och antalet som är förseningar för uppehåll- resp. gångtider samt procentandelen av tågobservationerna som är förseningar. Trots färre antal tågobservationer är det tydligt att uppehållstidsförseningar totalt sätt uppstår mycket oftare än gångtidsförseningar och står för 91,5% av de totala förseningarna medan gångtidsförseningarna står för 8,5%.

Tabell 3 Totala antalet tågobservationer och antal förseningar för uppehåll- resp. gångtider

Försening	Total antal observationer	Antal icke-förseningar	Antal förseningar	Andel förseningar	Andel av totala förseningar
Uppehåll	1 767 311	1 297 307	470 004	26,5%	91,5%
Gång	1 848 817	1 805 256	43 561	2,4%	8,5%

4.2 Felkoder i pendeltågstrafiken

Efter sammankoppling av förseningsdata och fordonsdata kan tågrörelser som har en eller flera felkoder identifieras. När ett fel uppstår i ett tåg registreras det i form av en felkod och samlas i en tabell som kallas för fordonsdata. Under den avgränsade tidsperioden genererades 811 olika typer av felkoder. Felkoder genereras av olika skäl i fordonen och har olika innebörd. De är klassade efter olika allvarlighetsgrader. Att en felkod genereras innebär att den specifika delen i fordonet som felkoden är anknuten till är i behov av en kontrollundersökning.

Tabell 4 Samtliga observationer med identifierade felkoder uppdelade i uppehåll- resp. gångtider.

		Uppehållstider		Gångtider	
Total antal observationer		1 767 311		1 848 817	
Förseningar utan felkoder	utan	406 384	1 550 227	32 487	1 608 006
Icke-förseningar utan felkoder	utan	1 143 843		1 575 519	
Förseningar med felkoder	med	63 620	217 084	11 074	240 811
Icke-förseningar med felkoder	med	153 464		229 737	

Tabell 4 visar samtliga tågobservationer uppdelade i uppehåll- resp. gångtider samt fördelningen av dessa i förseningar och icke-förseningar med och utan felkoder. Av 1 767 311 observationer av uppehållstider har 217 084 av dessa en

eller flera felkoder och av 1 848 817 observationer av gångtider har 240 811 av dessa en eller flera felkoder.

Notera att en tågobservation är, som tidigare nämnt, en registrerad observation av en tågrörelse som går från en station till en annan. Under en tågobservation kan flertalet felkoder genereras. Det betyder att varje tågobservation kan ha en eller flera felkoder. Efter uträkning i Microsoft SQL är det känt att 217 084 observationer av uppehållstider har sammanlagt 468 412 genererade felkoder och 240 811 observationer av gångtider har sammanlagt 520 236 genererade felkoder.

Med de identifierade tågobservationerna som har felkoder kan andelen felkoder som har genererats under förseningar räknas ut. Av 470 004 uppehålltidsförseningar har 63 620 av dessa felkoder, vilket innebär att 13,5% av uppehålltidsförseningarna har felkoder. Av 43 561 gångtidsförseningar har 11 074 av dessa felkoder, vilket innebär att 25,4% av gångtidsförseningarna har felkoder. Siffrorna visar att gångtidsförseningar endast står för 8,5% av de totala förseningarna. Däremot är andelen förseningar med felkoder större för gångtidsförseningar (25,6%) jämfört med uppehålltidsförseningar (13,5%).

Tabell 5 visar antalet tågobservationer med felkoder för uppehåll- och gångtider, 217 084 resp. 240 811. Av 217 084 uppehållstider med felkoder är 63 620 av dessa förseningar, vilket innebär att 29% av uppehållstider med felkoder är i form av förseningar. Av 240 811 gångtider med felkoder är 11 074 av dessa förseningar, vilket innebär att 4,6% av gångtider med felkoder är i form av förseningar.

Tabell 5 Antal tågobservationer med felkoder uppdelade i uppehåll- och gångtider

Upphållstider	217 084	100%
uppehålltidsförseningar	63 620	29%
Icke-förseningar	153 464	71%
Gångtider	240 811	100%
Gångtidsförseningar	11 074	4,6%
Icke-förseningar	229 737	95,4%

Tabell 6 Exempel på om en tågrörelse har någon felkod, generellt sett.

Tågnummer	Tidsstämpel	Avgångsstation	Ankomststation	Felkod
2877	2018- 11-22 - 17.00	Farsta strand	Trångsund	Ja eller Nej

Tabell 7 Exempel på en tågrörelse som har genererat en felkod samt vilken specifik felkod

Tågnummer	Tidsstämpel	Avgångsstation	Ankomststation	Felkod
2877	2018- 11-22 - 17.00	Farsta strand	Trångsund	14860

Tabell 8 *Exempel på en och samma tågrörelse som har fler än en genererad felkod samt vilka specifika felkoder*

Tågnummer	Tidsstämpel	Avgångsstation	Ankomststation	Felkod
2877	2018- 11-22 - 17.00	Farsta strand	Trångsund	14860
2877	2018- 11-22 - 17.04	Farsta strand	Trångsund	3909

5 Resultat

5.1 Samband mellan fordonsfel och förseningar

5.1.1 Felkoder, generellt

Denna sektion behandlar resultaten från analysens första del. Fokus ligger på felkoder generellt och ämnar identifiera eventuella samband mellan felkoder generellt och förseningar.

5.1.2 Uppehålltidsförsening och Gångtidsförsening

I tabell 9 nedan visas oddskvoter samt lägsta och högsta värde för konfidensintervall för uppehåll- resp. gångtidsförsening. Oddskvoten för felkoder generellt under uppehållstider är 1,17 och har ett konfidensintervall från 0,15 till 2,18. Resultatet innebär att när en felkod genereras under uppehållstid är oddsen att en försening sker 17% högre jämfört med om en felkod inte hade genererats. Eftersom konfidensintervallet innehåller värdet 1 innebär det att resultatet inte är statistiskt signifikant och att det inte finns något samband mellan felkod och uppehålltidsförsening.

Oddskvoten för felkoder generellt under gångtider är 2,34 och har ett konfidensintervall från 1,30 till 3,38. Resultatet innebär att när en felkod genereras under gångtid så är oddsen att en försening sker 2,34 gånger högre jämfört med om en felkod inte hade genererats. Eftersom det lägsta värdet för konfidensintervallet inte innehåller värdet 1 så innebär det att resultatet är statistiskt signifikant och att det finns samband mellan felkod och gångtidsförsening.

Tabell 9 Oddskvot och konfidensintervall för uppehåll resp. gångtidsförsening.

	OK	99,9% konfidensintervall	
		Lägsta värde	Högsta värde
Uppehållstid	1,17	0,15	2,18
Gångtid	2,34	1,30	3,38

5.2 Felkoder, specifika

Denna sektion behandlar resultaten för analysens andra del. Fokus ligger på specifika felkoder och ämnar identifiera eventuella samband mellan specifika felkoder och förseningar. Detta innebär att samtliga 811 felkoder analyseras individuellt.

5.2.1 Uppehålltidsförseningar

Tabell 10 visar nio felkoder som är statistiskt signifikanta. Antal gånger som felkoderna har genererats är 14 till 402. Oddskvoten för felkoderna sträcker sig

från 2,4 till 35,5. Tabell 11 visar samtliga nio felkoder och dess felmeddelanden. Båda tabellerna är rangordnade efter jämförelsevärdet ”antal extra förseningar”.

Tabell 10 Statistisk signifikanta felkoder för uppehålltidsförseningar.

Felkod	Antal förseningar med felkoder	Antal icke-förseningar med felkoder	Total antal felkoder	Oddsquot	Antal extra förseningar
11803	189	213	402	2,4	209,88
55469	13	1	14	35,6	176,88
46139	51	21	72	6,6	148,56
45212	114	117	231	2,7	140,69
11805	108	111	219	2,7	133,08
51814	35	15	50	6,4	98,40
51838	30	18	48	4,6	62,46
51839	28	17	45	4,5	57,67
16454	23	15	38	4,2	44,38

Tabell 11 Statistisk signifikanta felkoder och dess felmeddelanden.

Felkod	Oddsquot	Felmeddelande	Allvarlighetsgrad
11803	2,4	PIS driver's desk A2: Error PIS-display 2	B
55469	35,6	Traction A2: Traction motor controlled by inverter #2 is faulty	M
46139	6,6	Doors M3 door 41: Door obstructed	A
45212	2,7	Doors A1 door 12: Continuous passenger button signal	B
11805	2,7	PIS cab A2: Error coach controller 2	B
51814	6,4	PIS A2: Wheelchairspace - Error emergency speech unit	B
51838	4,6	PIS A2 door 61: Error emergency speech unit	B
51839	4,5	PIS A2 door 64: Error emergency speech unit	B
16454	4,2	TCMS: MMI lost communication to the AUX1	M

Felkod 55469 sticker ut med den högsta oddsquoten på tabellen. Med en oddsquot på 35,6 innebär det att när felkoden genereras under uppehållstid så är oddsen att en försening sker 35,6 gånger högre än om felkoden inte hade genererats, med andra ord ett starkt samband. Antal gånger felkoden har genererats är 14, vilket är betydligt färre gånger jämfört med de mer frekventa felkoderna i tabellen. Trots att felkoden har genererats få antal gånger resulterar dess höga oddsquot i ett starkt samband med förseningar och därmed dess placering i tabellen. Felkoden innebär att det finns ett problem med en särskild traktionsmotor.

Felkod 11803 ligger högst upp på tabellen trots en oddsquot på 2,4, vilket är den lägsta för uppehålltidsförseningar. Antalet gånger felkoden har genererats är däremot överlägset flest, 402 gånger. Denna höga frekvens resulterar i att felkoden, trots en låg oddsquot, i helhet har den största påverkan på

uppehålltidsförseningar av alla felkoder i tabellen. Felkoden innebär att det finns ett problem med visaren för passagerarinformation i förarpanelen.

Felkod 46139 har genererats 72 gånger och har en oddskvot på 6,6. Felkoden innebär att en särskild tågdörr är blockerad.

Felkod 45212 har genererats 231 gånger och har en oddskvot på 2,7. Felkoden innebär att det kan finnas ett problem med en särskild passagerarknapp.

Felkod 11805 har genererats 219 gånger och har en oddskvot på 2,7. Felkoden innebär att det finns ett problem med passagerarinformationssystemet på en särskild del av tåget.

Felkod 51814 har genererats 50 gånger och har en oddskvot på 6,4. Felkoden innebär att det finns ett problem med en nödtalenhet för en särskild rullstolssektion.

Felkod 51838 och *51839* innebär att det finns ett problem med en nödtalenhet för en särskild dörr.

Felkod 16464 ligger längst ner på tabellen med en oddskvot på 4,2 och har genererats 38 gånger. Det innebär att när felkoden genereras under uppehållstid så är oddsen 4,2 gånger högre att en uppehållsförsening sker jämfört med om felkoden inte hade genererats. Felkoden innebär att det finns ett problem med kontrollsystemet.

5.2.2 Gångtidsförsening

Bilaga 1 visar gångtidsförseningar där 78 felkoder är statistiskt signifikanta. Oddskvoten för felkoderna sträcker sig från 2,2 till 624,8. Bilaga 2 visar samtliga 78 felkoder och dess felmeddelanden. Eftersom felkoderna är många i antalet visar tabell 13 tio felkoder som ligger högst upp i listan exkluderat felkoder med allvarlighetsgrad M. Tabell 14 visar samma tio felkoder och dess felmeddelanden. Båda tabellerna är rangordnade efter jämförelsevärdet ”antal extra förseningar”.

Tabell 13 Statistisk signifikanta felkoder för uppehålltidsförseningar.

Felkod	Antal förseningar med felkoder	Antal icke-förseningar med felkoder	Total antal felkoder	Oddsquot	Antal extra förseningar
42263	714	3768	4482	6.7	729.8
2709	18	1	19	624.8	341.6
57300	562	8615	9177	2.3	337.7
56007	107	75	182	49.6	254.9
57318	202	2268	2470	3.1	149.4
51820	212	3047	3259	2.4	133.4
56012	55	201	256	9.5	62.7
56016	39	63	102	21.5	60.2
56009	55	234	289	8.2	59.7
56005	41	101	142	14	53.6

Tabell 14 Statistisk signifikanta felkoder och dess felmeddelanden.

Felkod	Oddsquot	Felmeddelande	Allvarlighetsgrad
42263	6.7	AUX M1: Battery charger defect	B
2709	624.8	Brakes A1 bogie 2: Transducer fault (minor)	C
57300	2.3	ERTMS: Loss of the equipment	A
56007	49.6	STM ATC2 - EVC communication faulty	A
57318	3.1	TCMS RIOM car A1: Analog output faulty	B
51820	2.4	PIS: UIC active, fallback Mode	C
56012	9.5	STM ATC2 recorder faulty or STM system failure	B
56016	21.5	maintenance - STM ATC2 emergency brake relay faulty	B
56009	8.2	EVC is in system failure (emergency brake)	A
56005	14	maintenance - ETCS emergency brake relay faulty	B

Felkod 42263 har genererats 4482 gånger under den studerade tidsperioden och har en oddsquot på 6,7. Felkoden innebär att batteriladdaren är defekt.

Felkod 2709 har den högsta oddsquoten av alla felkoder, 624,8, och har därmed det starkaste sambandet med förseningar. Felkoden har genererats i samband med en försening i 18 av 19 gånger. Men på grund av att den endast genererats 19 gånger totalt under den studerade perioden hamnar den lite längre ner på listan. Felkoden innebär att det finns ett problem med bromsen på en särskild bogie.

Felkod 57300 har en oddsquot på 2,3 och genererats 9177 antal gånger varav 562 i samband med gångtidsförseningar. Felkoden är både totalt och i samband med gångtidsförseningar bland de mer förekommande felkoderna. Felkoden innebär att det finns ett problem med signalsystemet eftersom en del av utrustningen har förlorats.

Felkod 56007 har en oddsquot på 49,6 och har genererats 182 gånger varav majoriteten av gångerna har varit i samband med gångtidsförsening. Felkoden innebär att EVC (European Vital Computer) har ett kommunikationsproblem.

Felkod 57318 har genererats 2268 gånger och har en oddskvot på 3,1. Felkoden innebär att det finns ett problem med TCMS (Train Control and Management System).

Felkod 51820 har genererats 3259 gånger och har en oddskvot på 2,4. Felkoden innebär att det finns ett problem med PIS (Passenger information system).

Felkod 56012 har genererats 256 gånger och har en oddskvot på 9,5. Felkoden innebär att det finns ett problem med signalsystemet.

Felkod 56016 har genererats 102 gånger och har en oddskvot på 21,5. Felkoden innebär att det finns ett problem med ett relä till en särskild nödbroms och är i behov av underhåll.

Felkod 56009 har genererats 230 gånger och har en oddskvot på 8,2. Felkoden innebär att det har dykt upp ett systemfel i EVC.

Felkod 56005 har genererats 142 gånger och har en oddskvot på 14. Felkoden innebär att det finns ett problem med ett relä till en särskild nödbroms och är i behov av underhåll.

6 Diskussion

6.1 Metoddiskussion

Oddskvot och relativ risk är två vanliga statistiska beräkningsmetoder varav relativ risk är den mer intuitiva metoden men eftersom detta är en retrospektiv fallkontrollstudie är det inte möjligt att räkna ut relativ risk. Därmed används oddskvot i arbetet för att undersöka om det finns något samband mellan felkod och förseningar.

Det valda konfidensintervallet i detta arbete är 99,9% på grund av det stora urvalet av felkoder. Uträkning med 95% utfördes även för att kontrollera potentiella skillnader. Vid analys av felkoder generellt hade resultaten varit oförändrade d.v.s. samband mellan felkoder och gångtidsförseningar men ingen samband mellan felkoder och uppehålltidsförseningar. Vid analys av specifika felkoder hade fler felkoder varit statistiskt signifikanta men eftersom resultatet visade att 78 felkoder var statistisk signifikanta för gångtider med ett konfidensintervall på 99,9% ansågs resultatet vara tillräckligt för att dra en slutsats. Dessutom är det säkrare att värdena inom intervallet faktiskt är statistiskt signifikanta och inte där av slump.

6.2 Felkoder, generellt

Vid analys av felkoder generellt visar resultaten för uppehållstidsförseningar att det inte finns någon koppling mellan felkoder och förseningar då konfidensintervallet innehåller värdet 1 och innebär att resultatet inte är statistiskt signifikant. Gällande gångtidsförseningar skiljer sig resultatet från uppehållstidsförseningar då resultatet för gångtidsförseningar är statistiskt signifikant och visar att det finns ett samband mellan felkoder och förseningar baserat på oddskvot.

Resultaten är intressanta då uppehållstidsförseningar står för den stora majoriteten av totala förseningar i pendeltågstrafiken (91,5%). Antal genererade fordonsfel är även många fler under uppehållstidsförsening (29%) jämfört med gångtidsförsening (4,6%). Däremot är andelen förseningar med felkoder större för gångtidsförseningar (25,6%) jämfört med uppehållstidsförseningar (13,5%).

Resultatet visar att det inte finns någon signifikant koppling mellan felkoder och uppehållstidsförseningar. Detta kan möjligen förklaras med hjälp av tidigare studier om uppehållstidsförseningar såsom Palmqvist et al., (2020), Olsson & Haugland (2004), Palmqvist et al. (2017) där studierna nämner multipla faktorer som påverkar uppehållstidsförseningar. Bland dessa finner man att den mänskliga faktorn såsom ombord- och avstigning av passagerare och tågpersonalens beteende har en signifikant påverkan på uppehållstidsförseningar. Ytterligare en viktig faktor är tidtabellens planering, vilket tidigare studier visat kan ha en stor inverkan på uppehållstidsförseningar.

6.3 Felkoder, specifika

Vid analys av varje specifik felkod för sig visar det sig att resultaten påvisar goda möjligheter att finna korrelationer mellan specifika felkoder och gångtidsförseningar baserat på oddskvot. Korrelationer mellan specifika felkoder och uppehållstidsförseningar påvisas också men är däremot inte lika många i antal.

Bilaga 1 visar 78 statistiskt signifikanta felkoder genererade som gångtidsförseningar enligt given beräkningsmetod där oddskvoterna varierar från 2,3 till 624,8. Antal gånger som respektive felkod genererats varierar från 5 till 41 689 gånger. Då samtliga 78 felkoder har statistiskt signifikanta samband med gångtidsförseningar bör de undersökas vidare. Av 78 felkoder har 48 en allvarlighetsgrad på A (18), B (24) och C (6).

Resultaten från första sektionen som behandlar felkoder generellt för gångtidsförseningar visade statistisk signifikans mellan felkoder och gångtidsförseningar. Därför var det väntat att även specifika felkoder uppvisade

likartad resultat för gångtidsförseningar. Antalet statistiskt signifikanta felkoder för gångtidsförseningar är utmärkande större än för uppehållstidsförseningar. Detta kan bland annat bero på att delar av fordonet i högre grad utsätts för slitage och stress under gång från infrastruktur, väder och temperatur som i sin tur ger upphov till uppkomsten av felkoder. Det innebär att gångtidsförseningarna möjligen beror mindre på den mänskliga faktorn och till en större del på ett problem med fordonet, något som kan leda till att en felkod genereras.

Tabell 10 visar nio statistisk signifikanta felkoder genererade under uppehållstidsförseningar. Jämfört med resultaten för gångtidsförseningar är felkoderna mycket färre i antalet, antal gånger genererad för varje felkod är lägre och värdet för oddskvot är i de flesta fall låga. Detta leder att felkoderna har svagare samband med uppehållsförseningar jämfört med gångtidsförseningar. Trots att antalet statistisk signifikanta felkoder som har samband med uppehållstidsförseningar är färre tenderar en större andel av det totala antalet gånger som felkoderna har genererats vara under uppehållstidsförseningar jämfört med motsvarande för gångtidsförseningar. Det innebär att felkoderna inte genereras särskilt ofta. Däremot tenderar de att genereras i samband med förseningar när de väl genereras. Efter närmare undersökning märks att flera av felkoderna som har samband med uppehållstidsförseningar har en anknytning till dörrar. Felkoder som har en anknytning till dörrar är inte lika vanliga under gångtid. Detta beror troligen på att dörrar öppnas och stängs under uppehållstid samt hålls stängda och låsta under gångtid. Även om de exakta felkoderna inte är vanliga under gångtid, är felkoder av detta slag är vanliga. I grunden har felkoder ett samband med delar av fordonet som påverkar förseningar på grund av stress och slitage, vilket är vanligare under gångtid. Detta förklarar varför sambandet är mycket starkare mellan felkoder och gångtidsförseningar jämfört med uppehållsförseningar. Fordonsfel är inte den primära anledningen till att uppehållstidsförseningar sker men trots att sambanden mellan felkoder och uppehållstidsförseningar inte är lika många så förekommer de fortfarande. Då även under uppehållstider, om än inte i samma grad, är delar av fordonet aktiva, så som tågdörrar. Om en fordonsdel aktivt används så utsätts den för slitage och med tiden löper större risk att påverka förseningar.

Resultatet för specifika uppehållsförseningar är inte oväntat då resultatet för felkoder generellt visat att det inte finns någon korrelation mellan felkoder och uppehållstidsförseningar. Av samma skäl kan resultaten för specifika felkoder likt felkoder generellt för uppehållstidsförseningar möjligen förklaras med hjälp av tidigare studier såsom Nie och Hansen (2005), Buchmueller et al., (2007), Palmqvist et al., (2020), Olsson och Haugland (2004), Palmqvist et al., (2017). Som tidigare nämnts uppger studierna multipla faktorer som påverkar uppehållstidsförseningar. Bland dessa finner vi att den mänskliga faktorn såsom

ombord- och avstigning av passagerare, tågpersonalens beteende och tidtabellens planering har en signifikant påverkan på uppehållsförseeningar.

7 Slutsats

Syftet med arbetet var att undersöka om det finns något samband mellan felkod och förseningar med hjälp av statistiska metoder, oddskvot samt konfidensintervall vilket i tur bekräftades med statistisk signifikans.

Resultaten för första delen av analysen, felkoder generellt, visar ett signifikant samband mellan felkoder och gångtidsförseningar men icke-signifikant samband mellan felkoder och uppehålltidsförseningar.

Resultaten för andra delen av analysen, specifika felkoder, visar att 78 olika felkoder har statistiskt signifikanta samband med gångtidsförseningar. Uppehållsförseningar visades ha färre signifikanta felkoder med endast 9 olika felkoder. Förklaringen är att den mänskliga faktorn så som ombord- och avstigning av passagerare, tågpersonalens beteende och tidtabellensplanering möjligen är den primära anledningen till att uppehålltidsförseningar sker. Gångtidsförseningar beror till en större del på ett problem med fordonet, något som kan leda till att en felkod av signifikans genereras, vilket förklarar det starka sambandet mellan felkoder och gångtidsförseningar.

Analys av förseningsdata och fordonsdata med hjälp av statistiska metoder är inte beprövad i tidigare studier och utgör därmed en ny infallsvinkel inom ämnet som förhoppningsvis resulterar i ökad förståelse kring punktlighet.

8 Felkällor och framtida rekommendationer

Det möjligt att resultaten hade sett annorlunda ut om en större mängd data hade kunnat förenats. För att stärka resultaten bör tillgång till mer lättkombinerad data från grunden möjliggöras. För att detta ska fungera är det viktigt att minimera felregistrerade fordonsdata och förseningsdata. I detta arbete har fordonsdata och förseningsdata kombinerats genom matchning av tågnummer samt tidsstämplar. På grund av felregistrering samt avsaknad av data för vissa observationer kunde inte all data matchas. Detta har lett till att resultaten i detta arbete utgörs av beräkningar på enbart en bråkdel av den totala data som fanns tillgänglig i början. Om man i framtiden vill använda liknande metod som i detta arbete så vore det en fördel att ha data som går lättare att kombinera. För att detta ska fungera så måste så mycket data som möjligt vara korrekt registrerade från start. Ju mer data man lyckas kombinera desto tillförlitligare resultat kan erhållas.

Som tidigare nämnt i metodsektionen har förseningsminuterna aggregerats i detta arbete. Oavsett storleken av en försening så räknas det som en försening i samma skala som alla andra förseningar. Alla punktliga och tidiga tåg gavs värdet 0 och icke punktliga tåg gavs värdet 1. Ett förslag på fortsatta studier är att göra en analys som tar hänsyn till storleken av förseningarna för att se hur resultaten påverkas.

9 Referenser

Buchmueller, S; Weidmann, U; Nash, A. (2008). Development of a dwell time calculation model for timetable planning.

Ceder, A; Hassold, S. (29 Jun 2015). Applied analysis for improving rail-network operations.

Gummeson, M. (2019). TTT- Tillsammans för Tåg i Tid Resultatrapport 2019.

Gysin, K. (2018). An Investigation of the Influences on Train Dwell Time. Swiss Federal Institute of Technology, ETH

Nie, L; Hansen, I.A. (April 2005). System analysis of train operations and track occupancy at railway stations.

Olsson, N; Haugland, H. (20 Jul 2004). Influencing factors on train punctuality—results from some Norwegian studies.

Palmqvist, C-W; Olsson, N; Hiselius, L. (2017) Delays For Passenger Trains On A Regional Railway Line In Southern Swede.

Palmqvist, C-W; Olsson, Nils; Hiselius, L. (2017) Some Influencing Factors For Passenger Train Punctuality In Sweden.

Palmqvist, C-W; Tomii, N; Ochiai, Y. (2020) Explaining dwell time delays with passenger counts for some commuter trains in Stockholm and Tokyo.

SJ. (Maj 2019). Rapport om klimatsmart resande. Hämtat från SJs hemsida: <https://www.sj.se/content/dam/SJ/pdf/sj-klimatrapport-maj-2019.pdf> [2021-06-27]

Statistiska centralbyrån. (Dec 2020) Sveriges befolkning. Hämtat från SCBs hemsida: <https://www.scb.se/hitta-statistik/sverige-i-siffror/manniskorna-i-sverige/sveriges-befolkning/> [2021-06-27]

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 19(3), 227–229. Hämtat från NCBI's hemsida: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/#idm139779227050832title> [2021-06-27]

Tenny S; Hoffman, MR. (Uppdaterad 2020 Jul 10) Odds Ratio. StatPearls [Internet], Treasure Island, FL. StatPearls Publishing; 2021 Jan. Hämtat från NCBI:s hemsida: <https://www.ncbi.nlm.nih.gov/books/NBK431098/> [2021-06-27]

Trafikförvaltningen och Waxholmsbolaget. Årsrapport. (2019). Hämtad från SLL:s hemsida: <https://www.sll.se/globalassets/2.-kollektivtrafik/upplevd-kvalitet-i-sl-trafiken/upplevd-kvalitet-arsrapport-2019.pdf> [2021-06-27]

Trafikverket, Trafikförvaltningen, MTR (12 Dec 2019). Rapport Senior samverkan: Från ord till handling – åtgärder för punktligare pendeltåg.

Transportstyrelsen (14 Maj 2017). Resenärers syn på järnvägsmarknaden. Hämtat från: <https://transportstyrelsen.se/sv/publikationer-och-rapporter/rapporter/marknadsovervakning/tsj-2017-1405/> [2021-06-27]

Wigren, A; Abelin, K. (Sep 2019). Störningsorsaker och punktlighet i SL:s pendeltåg

10 Bilagor

Bilaga 1 Statistisk signifikanta felkoder för gångtidföräningar.

Felkod	Antal föräningar med felkoder	Antal icke-föräningar med felkoder	Total antal felkoder	Oddsquot	Antal extra föräningar
56022	3296	38 273	41 569	3.1	2536.6
14860	1232	12 888	14 120	3.4	964.8
3908	620	1185	1805	18.4	902.7
15012	695	2340	3035	10.4	824.6
43908	747	3136	3883	8.4	824.4
42263	714	3768	4482	6.7	729.8
15212	630	2497	3127	8.9	707.1
15412	556	2118	2674	9.2	631.4
3909	400	760	1160	18.4	581.4
14815	361	813	1174	15.5	490.7
46812	472	4954	5426	3.3	363.9
16446	301	977	1278	10.7	358.9
2709	18	1	19	624.8	341.6
57300	562	8615	9177	2.3	337.7
56007	107	75	182	49.6	254.9
16312	171	371	542	16.0	234.9
16403	207	871	1078	8.3	226
57318	202	2268	2470	3.1	149.4
51820	212	3047	3259	2.4	133.4
14816	108	840	948	4.5	94.8
3914	142	1892	2034	2.6	94.3
56018	27	13	40	72.1	82
16212	68	288	356	8.2	73.9
56012	55	201	256	9.5	62.7
56016	39	63	102	21.5	60.2
56009	55	234	289	8.2	59.7
14802	54	276	330	6.8	55.1
56005	41	101	142	14	53.6
14841	23	21	44	38	46.9
51802	60	670	730	3.1	44.4
43104	32	155	187	7.1	33.2
42504	32	198	230	5.6	30.6
42526	32	204	236	5.4	30.2
56004	18	32	50	19.5	26.7
43704	29	209	238	4.8	26.2
56013	33	344	377	3.3	25.3
6726	12	12	24	34.7	23.3
43304	21	96	117	7.6	22.2
16059	28	297	325	3.3	21.3
42726	24	194	218	4.3	20.7
14820	13	21	34	21.5	20.1
7102	4	1	5	138.8	19.9
42704	21	134	155	5.4	19.8
56015	10	10	20	34.7	19.4
51801	20	141	161	4.9	18.2
45212	21	211	232	3.5	16.4

56023	13	37	50	12.2	16.1
42926	16	132	148	4.2	13.7
11804	16	142	158	3.9	13.3
42904	14	96	110	5.1	12.9
42260	3	1	4	104.1	11.0
45230	14	134	148	3.6	11.2
2525	5	4	9	43.4	11
55486	7	17	24	14.3	9.2
55040	9	64	73	4.9	8.2
16700	8	51	59	5.4	7.6
3102	6	18	24	11.6	7.3
55311	8	56	64	4.9	7.3
55086	7	42	49	5.8	6.7
9701	7	43	50	5.6	6.7
5221	2	1	3	69.4	5.9
14202	2	1	3	69.4	5.9
10711	2	1	3	69.4	5.9
7305	2	1	3	69.4	5.9
56008	3	3	6	34.7	5.8
56010	3	3	6	34.7	5.8
14808	4	10	14	13.9	5.2
14218	4	10	14	13.9	5.2
1642	4	12	16	11.6	4.9
16033	4	19	23	7.3	4.2
8706	2	2	4	34.7	3.9
10702	2	2	4	34.7	3.9
10710	2	2	4	34.7	3.9
46146	2	2	4	34.7	3.9
14118	3	9	12	11.6	3.7
45838	2	3	5	23.1	3.2
55305	2	3	5	23.1	3.2
45240	2	3	5	23.1	3.2

Bilaga 2 Statistisk signifikanta felkoder för gångtidförseening och dess felmeddelanden.

Felkod	Oddsquot	Felmeddelande	Allvarlighetsgrad
56022	3.1	Inhibition of the Traction by ETCS	M
14860	3.4	TB 60: Dead man button not pressed when starting	M
3908	18.4	Brakes: MTB applied at speed > 20 kph	M
15012	10.4	Traction A1: ED-brake effort reduced	M
43908	8.4	TB 29: Train Safety Loop opened	M
42263	6.7	AUX M1: Battery charger defect	B
15212	8.9	Traction M: ED-brake effort reduced	M
15412	9.2	Traction A2: ED-brake effort reduced	M
3909	18.4	Brakes: Emergency brake at speed > 20 kph	M
14815	15.5	TB 15: HL pressure below 4.5 bars	M
46812	3.3	TB 01: Door release not in 0 position while v < 3 km/h	M
16446	10.7	TCMS: MMI changed local level of access	M
2709	624.8	Brakes A1 bogie 2: Transducer fault (minor)	C
57300	2.3	ERTMS: Loss of the equipment	A
56007	49.6	STM ATC2 - EVC communication faulty	A
16312	16.0	TCMS: MMI2 user is logged in as driver	M
16403	8.3	TCMS: Train configuration was confirmed	M
57318	3.1	TCMS RIOM car A1: Analog output faulty	B
51820	2.4	PIS: UIC active, fallback Mode	C
14816	4.5	TB 16: Park Brake not released	M
3914	2.6	Brakes: Change cab test: OK - Change cab brake test fulfilled	M
56018	72.1	ATC: ATC is isolated. Speed is limited to 80 kph	M
16212	8.2	TCMS: MMI1 user is logged in as driver	M
56012	9.5	STM ATC2 recorder faulty or STM system failure	B
56016	21.5	maintenance - STM ATC2 emergency brake relay faulty	B
56009	8.2	EVC is in system failure (emergency brake)	A
14802	6.8	TB 02: WTB Trainbus inauguration in progress	M
56005	14	maintenance - ETCS emergency brake relay faulty	B
14841	38	TB 41: Running direction change while driving	M
51802	3.1	PIS A2: Error in Coupler-Backbone	B
43104	7.1	Brakes M3 bogie 4: Train Line and PWM or MVB are incoherent on EPAC 4	A
42504	5.6	Brakes A1 bogie 1: Train Line and PWM or MVB are incoherent on EPAC 1	A
42526	5.4	Brakes A1 bogie 1: MTB not available	B
56004	19.5	ETCS emergency brake relay faulty	A

43704	4.8	Brakes A2 bogie 7: Train Line and PWM or MVB are incoherent on EPAC 7	A
56013	3.3	ETCS - Loss of redundancy	B
6726	34.7	Doors A2 door 64: Door parameters not acquired	M
43304	7.6	Brakes M3 bogie 5: Train Line and PWM or MVB are incoherent on EPAC 5	A
16059	3.3	Cab A2: Screen heating faulty or temperature out of range	A
42726	4.3	Brakes M1 bogie 3: MTB not available	B
14820	21.5	TB 20: Hold Brake not released	M
7102	138.8	DBC: A desk is armed as well as partly armed concurrently	B
42704	5.4	Brakes A1 bogie 2: Train Line and PWM or MVB are incoherent on EPAC 2	A
56015	34.7	STM ATC2 emergency brake relay faulty	A
51801	4.9	PIS A1: Error in Coupler-Backbone	B
45212	3.5	Doors A1 door 12: Continuous passenger button signal	B
56023	12.2	ETCS is isolated. Speed is limited to 80 kph	M
42926	4.2	Brakes A2 bogie 7: MTB not available	B
11804	3.9	PIS cab A1: Error coach controller 1	B
42904	5.1	Brakes M1 bogie 3: Train Line and PWM or MVB are incoherent on EPAC 3	A
42260	104.1	AUX M1: Battery charger battery overcurrent	C
45230	3.6	Doors A1 door 12: Limited usability	B
2525	43.4	Brakes A1 bogie 1: Main pipe pressure out of range (minor)	C
55486	14.3	Traction A2: PMCF is faulty	M
55040	4.9	Traction A1: Contactor Fault - VCB - Lockout	B
16700	5.4	Maintenance period ON	M
3102	11.6	Brakes M3 bogie 4: Pneumatically brake stop cock closed on bogie n°4	A
55311	4.9	Traction M3: Both traction inverter inhibited	B
55086	5.8	Traction A1: PMCF is faulty	M
9701	5.6	HVAC8 Cab A2: Communication lost	B
5221	69.4	Doors A1 door 01: Door parameters not acquired	M
14202	69.4	Propulsion Case A2: Failure in the Coolant Circuit	B
10711	69.4	High Voltage: Pantograph M3 not raised	M
7305	69.4	DSD/DVS, cab A2: DSD disabled, DVS enabled instead	M
56008	34.7	STM ATC2 - EVC communication faulty	A
56010	34.7	ETCS fan faulty	B
14808	13.9	TB 08: Coupler state fault	M

14218	13.9	Traction Motor Cooler A2: General Cooler Fault	B
1642	11.6	AUX M1: Battery Charger, failure in internal voltage supply	C
16033	7.3	Fire detection, M4: Fire detection faulty	B
8706	34.7	HVAC3 Saloon M1: Exhaust air fan or contactor failure	C
10702	34.7	High voltage system: Major fault in the HV system	A
10710	34.7	High Voltage: Pantograph M2 not raised	M
46146	34.7	Doors M3 door 44: Safety failure of door!	A
14118	11.6	Traction Motor Cooler M: General Cooler Fault	B
45838	23.1	Doors M2 door 34: Door malfunction	A
55305	23.1	Traction M: Over Temperature Traction Motor Level 1	B
45240	23.1	Doors A1 door 02: Door obstructed	A

Bilaga 3 Kritiska värden för uträkning av konfidensintervall

Table 1 Critical (z) values used in the calculation of confidence intervals

Confidence level	Critical (z) value to be used in confidence interval calculation
50%	0.67449
75%	1.15035
90%	1.64485
95%	1.95996
97%	2.17009
99%	2.57583
99.9%	3.29053