

Seminarieuppsatser nr 190

Kvalitet och osäkerhet i geografisk analys-

En studie om kvalitetsaspekter med fokus på osäkerhetsanalys av rumslig prognosmodell för trafikolyckor

Edvin Eriksson

2010
Centrum för Geobiosfärvetenskap
Naturgeografi och Ekosystemanalys
Lunds Universitet
Sölvegatan 12
223 62 Lund



Centrum för Geobiosfärvetenskap
Naturgeografi och Ekosystemanalys
Lunds Universitet

Kvalitet och osäkerhet i geografisk analys

En studie om kvalitetsaspekter med fokus på
felfortplantning i rumslig prognosmodell för
fordonsolyckor

Masteruppsats i Geomatik

Edvin Eriksson

Handledare: Lars Harrie
Kursanvarig: Harry Lankrejer

Förord

Detta examensarbetet på 30 hp är det sista obligatoriska momentet i min master utbildning i Geomatik. Därmed avslutar även examensarbetet utbildningen i Lund. Examensarbetet har varit en intressant resa med många nya lärdomar som jag inte hade kunskap om tidigare. Det har krävts att jag satt mig in i tekniska beskrivningar av olyckor på vägar, trafikflöden o.s.v.

Förhoppningsvis kommer examensarbetet ingå i den officiella beskrivningen av planerings verktyget ROPA som har premiär nu i vår. I vilket fall så vill jag passa på att tacka alla berörda inom BeRädd projektet för att jag fått möjlighet att vara del av projektet. Ett extra tack vill jag rikta till Åsa Forsman på VTI som bistått med outhärlig hjälp vid felfortplantningsanalysen.

På hemmaplan så vill jag rikta ett stort tack till min handledare Lars Harrie på Lunds universitet för råd, nästan alla tider på dygnet (även på tjänsteresa). Dessutom vill jag tacka min sambo för moraliskt stöd under många långa arbetsdagar.

Abstract

Geographical analysis is increasingly used in decision making processes. It is therefore very important that such analysis is performed with a satisfying quality, so that the results can be used as a representative model of reality. Quality descriptions of spatial data are common and several standards have been developed that describes how spatial data should be gathered and quality assured. Less common, however, is error propagation analysis. Error propagation within geographical analysis is a mean of study how uncertainty in spatial data propagates through a geographic model. The results from an error propagation analysis can be used to estimate the quality of a geographical analysis. This paper describes the general concept of quality and uncertainty in geographical analysis. An error propagation analysis is carried out to examine a forecasting model of vehicle accidents. The error propagation method is analytical and is based on a first degree Taylor expansion. The result gives that the uncertainty of the forecasting model based on error propagation analysis is highly varied within the test area. As a complement to the error propagation analysis, an evaluation study is conducted to compare the forecasting model output with accident data from STRADA (Swedish Traffic Accident Data Acquisition). The evaluation result indicates that the forecasting model has a good agreement in areas consisting predominantly of larger roads but less agreement in larger built up areas.

Key words: Geographical analysis, Quality, Uncertainty, Error propagation, Forecasting model, GIS

Sammanfattning

I dagens samhälle används geografiska analyser i stor uträkning som verktyg för att hjälpa till vid beslutsfattning. Det är viktigt att sådana geografiska analyser håller en god kvalitet för att resultatet ska kunna representera verkliga förhållanden. Kvalitetsgranskning av geografiska data är vanligt och flera standarder har tagits fram som beskriver hur geografiska data ska insamlas och kvalitetssäkras. Mindre vanligt är felfortplantningsanalyser. Felfortplantningsanalyser används för att studera hur osäkerheten i geografiska data fortplantas i en geografisk analys. Uppsatsen beskriver generella begrepp om kvalitet i geografisk analys. En felfortplantningsstudie genomförs på en prognosmodell för fordonsolyckor. Metoden för felfortplantningen som används är en analytisk bestämning med första gradens Taylorutveckling. Resultatet visar att osäkerheten baserat på felfortplantningen är varierande inom testområdet. Som komplement till felfortplantningsanalysen så utvärderas även prognoskiktet mot olycksstatistik från STRADA (Swedish Traffic Accident Data Acquisition). Resultatet från utvärderingen indikerar på bättre överensstämmelse på områden innehållande större vägar än områden i större tätorter.

Nyckelord: Geografisk analys, Kvalitet, Osäkerhet, Felfortplantning, Prognosmodell, GIS

Innehållsförteckning

Inledning	4
1. Kvalitetsaspekter i rumslig analys	10
1.1. Osäkerhet.....	10
1.2. Kvalitet	11
1.2.1. Kvalitet av geografiska data	11
1.3 Metoder för kvalitetsutvärdering av geografisk analys.....	16
1.3.1. Felfortplantning	17
2.2.4. Test av rumslig modell och färdig produkt	21
2. Rumsliga prognosmodeller och planering av räddningstjänst	24
2.1. Allmänt om rumsliga prognosmodeller.....	24
2.2. Prognosmodell för planering av räddningstjänst.....	24
2.2.1. BeRädd och ROPA.....	25
3. Fallstudie: Kvalitetsanalys av prognosskikt för vägolyckor	32
3.1. Områdesbeskrivning.....	32
3.2 Data	34
3.3 Metod	35
3.3.1 Metod för beräkning av prognosskiktet	35
3.3.2 Implementering av prognosskiktet	36
3.3.3 Metod för felfortplantningen	38
3.3.4. Implementering av felfortplantningen.....	45
3.3.5 Statistisk utvärdering mot utfall	46
3.4. Resultat.....	49
3.4.1 Resultat från felfortplantningsanalysen.....	49
3.4.2 Resultat från utvärdering mot utfallsdata	51
4 Diskussion	55
4.1 Diskussion om felfortplantningen	55
4.1.1 Resultaten	55
4.1.2 Övrig diskussion.....	55
4.1.3 Diskussion om osäkerhetsval hos parametrar	57
4.2 Diskussion om utvärdering mot polisrapporterade olyckor från STRADA.....	57
4.2.1 Utvärdering inom intresseområden	57
4.2.2 Utvärdering hela prognosskiktet	58
4.3 Diskussion om fortsatta studier och egna reflektioner	59
4.3.1 Fortsatta studier	59
4.3.2 Egen reflektion	60
5. Slutsats	63
5.1 Specifika slutsater	63
5.1 Allmänna slutsater	63
Referenser	64
BILAGA 1 - Vägverkets effektkatalog	69
BILAGA 2 - Flödescheman för implementeringen	68
BILAGA 3 - MatLab script	69

Inledning

Bakgrund

Geografiska analyser ligger till grund för många beslut som fattas i samhället där GIS (Geografiska Informations System) används som verktyg i analysprocessen. Tack vare GIS kan flera verksamhetsområden underlättas inom planering och analys. Ett urval av processer där geografiska analyser är en förutsättning är planering vid nybyggnation, underhåll av naturskyddsområde och kartering av miljöföroreningar (Burrough 1998). Användningsområdena är minst sagt flertaliga och ytterst varierande, variationen beror delvis beror på att ungefär 80 procent av all data som används i olika beslutsprocesser kan knytas an geografiskt (Rönnbäck 2004).

Geografiska analyser baseras på *modeller*. Det är viktigt att producera modeller som återspeglar verkligheten till så stor del som möjligt. Resultatet av en geografisk modell kan dock aldrig bli bättre än den data som bygger upp modellen (Heuvelink 1998). Geografiska data, såväl som alla andra data, är alltid behäftad med någon typ av osäkerhet, som har sitt ursprung från flera källor t.ex. oprecis inmätning och uppskattningar.

Användbarheten utav en geografisk modell avgörs genom dess förmåga att svara mot en kravspecifikation d.v.s. vilken kvalitet modellen håller för sitt tänkta syfte. En sådan undersökning benämns ofta som en *kvalitetsundersökning* och utförs dels för att studera kvaliteten i indata och modellen och dels för att se vilken påverkan kvaliteten får på resultatet av modellen. Kvalitetsundersökningar är förhållandevis kostsamma men en negligering av en sådan kan innebära en större kostnad än utförandet (Rönnbäck 2004). Exempelvis är det viktigt då en bro konstrueras att inte vattenflödet överskrider brottgränsen för brons fundament; om en dålig vattenflödesmodell använts och modellerat flöde är för lågt jämfört med det faktiska flödet så kan det få allvarliga konsekvenser (Frank 2004).

Ett GIS innehåller i regel flera funktioner och verktyg. Ofta behöver en användare enbart nyttja ett GIS med avseende på ett fåtal funktioner, t.ex. användare som är intresserade av trafikflöden behöver förmodligen inte verktyg för hydrologisk modellering. Av den orsaken så finns det flera system som är anpassade till ett användningsområde. Exempelvis hitta.se (Hitta.se 2010) har funktioner för navigering d.v.s. närmaste vägen mellan två platser samt sökning av personer och företag. Däremot har tjänsten inte några verktyg för att planera kollektivtrafik vilket skånetrafikens karttjänst har (Skanetrafiken 2010).

Exempel på geografiska analyser är planering av räddningstjänst. MSB (Myndigheten för Samhällsskydd och Beredskap) har som uppdrag att stödja kommuner vid planering och organisering av räddningstjänst. Planering räddningstjänst är svårt och kräver erfarenhet. För att underlätta planeringen och ge beslutsfattare en objektiv grund så har en projektgrupp

med MSB som uppdragsgivare startats med uppdrag att utveckla ett GIS verktyg. Verktöget syftar till att effektivisera räddningsinsatser genom att beräkna tillgänglighet till räddningsresurser d.v.s. optimal allokering av räddningsresurser baserat på prognostiserade riskområden. Applikationen är ett skraddarsytt GIS där enbart planering av räddningstjänst är möjligt. Analysen baseras på sannolik olycksfrekvens där andel olyckor skattas med olika prognoskikt, vilka utgörs av ett antal rasterskikt. Varje prognoskikt är baserat på en modell som i sin tur är uppbyggd av geografiska data (Axelsson et al. 2008).

Givetvis är det av yttersta vikt att ett verktyg för planering av räddningstjänst är behäftad med så liten osäkerhet som möjligt. Stor vikt läggs vid att prognoskikten representerar verklig riskfördelning. Av den orsaken är en kvalitetsutvärdering utav prognoskikten nödvändig för att tillgänglighetsanalyserna som bygger på prognoskikten ska generera så tillförlitliga resultat som möjligt, och framförallt att kvaliteten är känd. I denna studie undersöks olika metoder för att utvärdera kvaliteten utav prognoskikten.

Syfte

1. Utföra en litteraturstudie om kvalitetsaspekter av geografisk analys.
2. Beskriva prognosmodeller för olycksrisker i allmänhet och prognosmodellerna som används i vid ett planeringsprogram för räddningstjänst i synnerhet.
3. Genomföra en kvalitetsundersökning av prognoskiktet för fordonsolyckor.
 - Undersökningen sker i ett testområde och syftar till att skatta osäkerheten i prognoskiktet baserat på osäkerhet i indata samt överensstämmelse mellan prognoskiktet och rapporterad olycksstatistik.

Metod

Teoretisk del

Den teoretiska delen utgörs av två studier. Den första delen är en litteraturstudie av kvalitetsaspekter där ämnen som standarder, datakvalitet och felfortplantning inom geografisk analys har studerats. Informationskälla till litteraturstudien har varit gängse akademiska artiklar och böcker.

Den andra delen inleder med en allmän beskrivning av geografiska prognosmodeller och användningsområden. Vidare beskrivs planering av räddningstjänst och prognosmodeller som används i planeringsprogrammet ROPA. Beskrivning av ROPA i sin helhet och de ingående prognoskikten med modeller har hämtats uteslutande från rapporter och utkast från de olika aktörerna i BeRädds projektgrupp.

Praktisk del

Fallstudien utgörs av en kvalitetsutvärdering av VTI:s prognosmodell för trafikolyckor samt prognosskikt genererade från samma modell. Testområdet för kvalitetsutvärderingen ligger i Östergötlands län med merparten i Linköping kommun. Prognosskikten har skapats i ESRI ArcGIS Desktop 9.3 med MathWorks MatLab 7.9 som stödprogram. Utvärderingen är indelad i två delar:

1. Felfortplantningsstudie av prognosmodellerna. Ett metod för att se hur osäkerhet i indata fortplantas i modellerna. Metoden som använts är analytisk felfortplantning baserat på första ordningens Taylorutveckling. Felfortplantningen utförs i MatLab.
2. Statistisk utvärdering utav prognosskiktets överensstämmelse med "verkligheten" genom jämförelse med polisrapporterad insatsdata under åren 2003-2008. Jämförelsen utförs dels som en områdesmässig jämförelse för att se hur olika områden motsvarar den faktiska olycksstatistiken och dels som en utvärdering för hela prognosskiktet inom testområdet där utvärderingen sker på cell nivå.

Avgränsningar

Studien koncentrerar sig på kvalitetsutvärdering av prognosskiktet för fordonsolyckor. De andra prognosskikten som används i BeRädd , som bostadsbränder, utvärderas inte.

Osäkerheten som studeras är enbart den som uppkommer p.g.a. osäkerhet i indata. Ingen hänsyn tas till osäkerhet av modellernas uppbyggnad t.e.x. om ekvationen för olyckor på enskilda vägar representerar faktisk olycksrisk på enskilda vägar.

Den geografiska begränsning för studien är ett område i Östergötland som b.l.a. inkluderar Linköping tätort.

Målgrupp

Examensarbetet kan med fördel läsas av fackpersoner och studenter med intresse för geografiska analyser. Exempel på målgrupper kan vara tekniska lantmätare, geografer med inriktning mot GIS samt personer från räddningstjänsten med intresse för planeringsfrågor. Ingen nämnvärd bakgrundskunskap krävs i området gällande fordonstrafik och räddningstjänst. Givetvis förutses en grundförståelse av statistik, matematik och GIS.

Definitioner och begrepp

Axelpar: Ett hjulpar på ett fordon. Trafikflöde uttrycks b.l.a. i axelpar då olika fordon har olika antal axelpar.

BeRädd: Projektgrupp med MSB som uppdragsgivare, projektet går ut på att ta fram ett verktyg för planering av räddningsresurser.

Georeferering: Metoder för att definiera och transformera en geometrisk figur eller bild till ett geodetiskt referenssystem.

GIS: Geografiska informationssystem är IT lösningar som innehåller metoder för insamling, behandling, analys och presentation av lägesbundna data.

Enskild väg En enskild väg är en väg som inte ägs av staten. Ofta karakteriseras de med en gul vägskylt med röd bård. De kan delas upp i olika kategorier:

- Enskilda vägar med statsbidrag. Dessa är tillgängliga för allmän trafik
- Enskilda privata vägar till bebodda fastigheter
- Enskilda vägar som används i skogsnäringen

Felfortplantning: Metoder inom vetenskap och teknik för att uppskatta hur fel i indata fortplantas i en analys/modell.

Första ordningens Taylor utveckling: En linjär polynomapproximation av en funktion.

Kommunal väg: En väg där kommunen står som väghållare d.v.s. kommunen äger vägen och sköter underhållet.

MSB: Myndigheten för samhällsskydd och beredskap, f.d. Räddningsverket och Krisberedskapsmyndigheten.

Monte Carlo simulering: En iterativ simuleringsmetod. En modell körs flera gånger där indata till modellen ändras för varje gång, värdena samplas från den statistiska fördelningen av indata. Resultatet är en fördelning om modellresultatets sannolika värden för olika indata.

Normalfördelning: Gauss klockformade populationsfördelning. I särklass den vanligaste fördelningen vid statistiska beräkningar.

Normalvärde: Ett normeringsvärde som används vid beräkning av olyckor på staliga vägar.

NVDB: Nationella vägdatan, består av linjeskikt som symboliserar det Svenska vägnätet. Innehåller attribut om t.e.x. hastighetsgräns och vägklass. Insamling sker som ett

samarbete mellan: vägverket, skogsstyrelsen, lantmäteriet, kommunerna, försvaret och privata bidragsgivare. Ansvarig producent är vägverket.

Poissionfördelning: En diskret statistisk fördelning som förklarar händelsers sannolika utfall i rummet eller i tiden. Fördelningen kan t.e.x. förklara antal trafikolyckor i rummet.

Primärväg: Vägbenpar i korsning med högst flöde.

Prognosmodell: En modell som används för att beräkna förutsägelser om framtiden. Inom naturvetenskap och teknik baseras sådana modeller på samband erhållna från tidigare observationer.

Relativ olycksrisk: Ett värde som används vid beräkning av olyckor på kommunala vägar och korsningar.

ROPA: Räddningstjänst, Optimering, Planering och Allokering. Datorverktyg som tagits fram utav BeRädd. Syfte att hjälpa räddningstjänsterna att planera allokering av tillgängliga räddningsresurser.

Räddningsresurs: Resurser som används vid räddningstjänst t.ex. släckbilar och stegbilar.

Sekundärväg: De vägben i en korsning som inte identifieras som *primärvägar*.

Statlig väg: En väg som ägs och underhålls av staten.

STRADA: Swedish Traffic Accident Data Acquisition. Informationsdatabas vars mål är att innehålla information om alla polisrapporterade trafikolyckor i Sverige från 2003.

Trafikplats: Planskild korsning. Kan vara över 1 km stora i diameter.

VTI: Statens Väg- och TransportforskningsInstitut. Forskar om b.la. trafiksäkerhet.

Vägben: I en korsning finns det 3 eller flera vägben som ansluter. Om vägfilerna är separerade i NVDB kan de räknas som olika vägben. Generellt har en 3 vägs korsning 3 vägben som ansluter, 4 vägs har 4 vägben osv.

Väglänk: Ett vägdela som symboliseras av en linje i en geometrisk databas.

Osäkerhet: Konceptuell term som redogör för avvikelserna från den "sanna" representationen av verkligheten.

ÅDT: Trafikflöde (ÅrmedelDygnst Trafik). Beräknas av vägverket och är en kombination av uppmätta värden och modellvärden.

1. Kvalitetsaspekter i rumslig analys

1.1. Osäkerhet

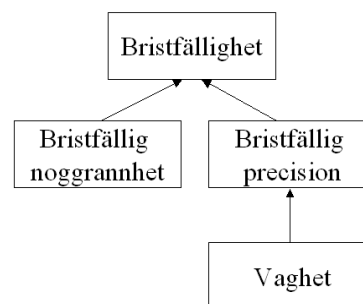
Alla geografiska analyser är förknippade med någon typ av *osäkerhet*. Förmodligen förknippar de flesta begreppet osäkerhet med ett mänskligt sinnestillstånd, vilket influerar hur vi fattar beslut och agerar i vår vardag. Ofta kan en sådan sinnesstämning beskrivas med adjektiv såsom tveksamhet, tvetydighet och vaghet. Adjektiv, som även vid geografiska analyser är högst relevanta. Relevansen uppkommer för att tvetydiga och vaga beskrivningar av data orsakar en tveksamhet vid skapandet av geografiska modeller (Worboys 2004). Utöver den allmänna beskrivningen ovan så kan osäkerhet också handla om fysiska/konkreta företeelser, särskilt intressant i geografiska analys-sammanhang. Exempel på sådana faktorer är osäkerhet vid inmätning och osäkerhet vid digitalisering. Företeelser som kan kvantifieras numeriskt snarare än en tvetydig specifikation av metodiken (Zhang 2002).

Begreppet osäkerhet beskrivs lättast genom att ställa upp en konceptuell figur. I figur 1 illustreras de fundamentala osäkerhets-element som beskrivs av Worboys (2004). Vår uppfattning om verkligheten är baserad på observationer och representationer av densamma. Huvudklassen i modellen, *bristfällighet*, hänvisar till den oundvikliga avvikelserna mellan en perfekt återgivning av verkligheten och vår uppfattning. Orsakerna till avvikelserna beror på *bristfällig noggrannhet* och *bristfällig precision*. *Bristfällig noggrannhet* beskriver avsaknaden av korrelation mellan observationer av verkligheten och en representation av densamma. *Bristfällig precision* beskriver avsaknaden av detaljnivå för att fullgott beskriva verkligheten. En viktig del i precisionen är *vaghet*, som är en typ av *bristfällig precision* (Worboys 2004).

Vaghet hänvisar till otydliga semantiska specifikationer bidrar till osäkerhet, och därmed inte detaljnivån i sig (NE *vaghet* 2009). Betrakta följande påstående: "Skogen är en barrskog." Detta är givetvis sant om skogen består till 100 % av träd med barr. Men om skogen innehåller 1 % lövträd, är det fortfarande sant att skogen är en barrskog? Var går gränsen? Gränsen går givetvis inte att bestämma entydigt eftersom orsaken till problemet är en språklig beskrivning som inte har någon numerisk definition (Worboys 2004).

Flera geografiska analyser utnyttjar kännedom om *osäkerhet* för att skapa bättre modeller, där koncept som *oskarp logik* används frekvent. För en genomgående beskrivning av *oskarp logik* hänvisas till Bourroughs (1998); exempel på analyser som nyttjat *oskarp logik* är McBratney & Odeh (1997) och Eastman & Hong (2000).

Figur 1: Hierarkisk konceptuell modell som visar de olika elementen i begreppet osäkerhet (Omarbetad från Worboys (2004).



Sammanfattningsvis så förklarar osäkerheten de brister som uppstår då vi försöker skapa en representation av verkligheten d.v.s. en modell. Måttet som beskriver osäkerheten i en geografisk modell och dess ingående data benämns *kvalitet*.

1.2. Kvalitet

Kvalitet är ett brett begrepp. Traditionellt kan det delas upp i: primär-, sekundär- och tertiär kvalitet. Primär kvalitet beskriver fysiska ting som materia, volym o.s.v. Detta skiljer sig från sekundär kvalitet som beskriver t.ex. färger och lukt; ting som har med våra sinnen att göra. Tertiär kvalitet har att göra med estetiska värderingar och tillämpas i kombination med de övriga indelningarna t.ex. två maträtter med samma primära och sekundära kvalitet skulle kunna skilja sig i hänseende på tertiär kvalitet (NE *kvalitet* 2009).

De tre beskrivningarna ovan är ytterst filosofiska och hänvisar mer till allmän kvalitet. En bättre benämning då kvalitet ska beskrivas för geografiska analyser och data är: "*Kvalitet är egenskaper hos produkt eller tjänst som är av betydelse för dess förmåga att tillgodose kända krav*" (Stanli 1996). Alltså, kvalitet är ett relativt begrepp som måste kopplas till användningsområde.

Stanlis definition för kvalitet är specifik och fungerar bra för detta examensarbete eftersom kraven för prognosskikten som ska utvärderas är kopplade till ett specifikt användningsområde, nämligen räddningstjänst.

1.2.1. Kvalitet av geografiska data

Att geografiska data har en viss kvalitet är av yttersta vikt då de kan ligga till grund för flera olika beslut (Crosetto et al. 2000). Om inte kvaliteten på indata är känd så kan inte heller en kvalitetsutvärdering utföras på de analyser som använder samma data (Rönnbäck 2004). Av olika anledningar så är emellertid kvaliteten på indata ofta förbisedd i många geografiska analyser (Veregin & Lanter 1995, Goodchild 2000).

En geografisk datamängds kvalitet beror på en mängd olika faktorer. Data samlas in med olika metoder. Geometrin mäts in med fotogrammetriska och/eller geodetiska mätmetoder

och knyts till attribut med information om de icke spatiala egenskaperna. Dessa processer sker ofta inte samtidigt och inte sällan av olika arbetsgrupper. Utöver detta så pågår insamlingsprocessen för en datamängd dessutom ofta över en längre tidsperiod, vilket innebär att det inte säkert att samma metoder har använts i datainsamlingen eller att samma aktörer varit inblandade. Kvaliteten på geografiska data är alltså beroende av olika typer av mätfel, osäkerheter gällande attributen hos data och dessutom är den även förändlig över tiden (Goodchild 2000).

Vanligtvis delas kvalitet för geografiska data in i två subkategorier: intern (teknisk) och extern (funktionell) kvalitet. Den interna kvaliteten förklaras med kvalitetstema och kvalitetsparametrar (vilket beskrivs senare) och kan t.ex. uttryckas som: fullständigheten för byggnader med hänseende på angiven fastighetsägare är 95 %. Den externa kvaliteten syftar till användbarheten av data och är mer lösryckt i sin beskrivning eftersom dataset lämpar sig olika bra till olika användningsområden. Exempelvis så behöver inte data till en turistkarta erhålla samma kvalitet som till en kommunplan. (Lantmäteriet *Kvalitet* 2010) , därmed innebär hög extern kvalitet för en tillämpning inte innebära en hög kvalitet för en annan.

1.2.1.1 Fel i geografiska data

Fel definieras olika i skilda discipliner. Termen återfinns både inom juridiskt språkbruk som inom matematiken. I juridiska termer så anger ett fel en olaglig handling och har lite med kvalitet att göra. Ett matematiskt fel definieras som skillnaden mellan ett exakt värde av en storhet (b) och en approximation av densamma (a) (NE *fel* 2009 ; Heuvelink 1998), allmänt kan felet (v) av en mätning formuleras enligt:

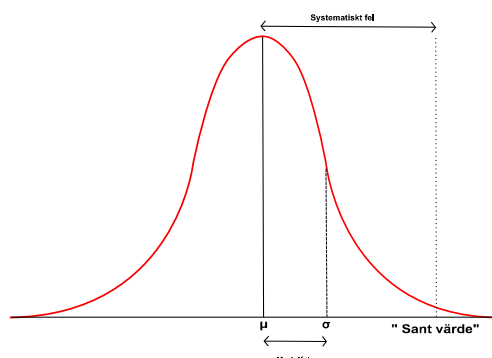
$$v = a - b \quad (1)$$

Ekvationen ovan är emellertid inte särskilt realistisk i praktiska tillämpningar. Det exakta värdet är aldrig känt. Om det exakta värdet hade varit känt så hade även en perfekt approximation varit möjlig. Felet går alltså inte att bestämma entydigt. (Heuvelink 1998)

Med utgångspunkt från ovanstående resonemang så är det bättre att beskriva de ingående variablerna som stokastiska variabler. Det går inte att mäta in en företeelse exakt. Om ett upprepat antal mätningar gör så kommer de att skiljas åt. Det värdet som brukar användas är väntvärdet d.v.s. medelvärdet av de uppmätta värdena. Felen i data är då avvikelsen mellan väntvärdet och de faktiska mätningarna (Heuvelink 1998).

Fel i geografiska data kan delas upp i slumpmässiga och systematiska fel och kan ofta beskrivas med en *normalfördelning*. Påståendet förklaras med *den centrala gränsvärdessatsen* d.v.s. summan av många slumpvarierande företeelser med samma fördelning går mot en normalfördelning (Olsson 2005). Med utgångspunkt från en normalfördelning så är det systematiska felet lika med avvikelsen mellan det "sanna värdet" och väntvärdet hos den uppmätta datamängden. De slumpmässiga felens fördelning kring väntvärdet ger upphov

till storleken för medelfelet eller osäkerheten hos fördelningen som är likvärdigt med standardavvikelsen σ , se figur 2.



Figur 2: Den röda fördelningen visar den uppmätta fördelningen av data. Det systematiska felet är skillnaden i mellan det uppmätta medelvärdet och värdet det borde ha. Standardavvikelsen anger medelfelet utav de slumpmässiga (stokastiska) felen (Omarbetad från Wasström et al. 2009).

Systematiska fel uppkommer generellt p.g.a. brister i mätutrustning eller metodik. Sådana fel går att korrigera för genom att använda bättre utrustning och metodik. Däremot är slumpmässiga fel mer oberäknliga och beror på andra orsaker och har därmed större inverkan på en analys. Generellt brukar andelen positiva och negativa slumpmässiga fel vara lika kring inmätningarnas väntvärde (Ollvik 2002). Heuvelink (1998) beskriver fördelningen utav möjliga värden med hänseende på väntvärdet och osäkerheten som en osäkerhetsfördelning. Desto större osäkerhet eller medelfel för data desto större är spridningen på fördelningen och då minskar användbarheten..

1.2.1.2 Kvalitetsteman och metadata

Kvaliteten på en geografisk datamängd beskrivs i medföljande dokumentation, s.k. *metadata*. Metadata beskriver t.e.x. kvaliteten av data, koordinatsystem och förvaltare. Metadata kan beskrivas som information som gör data användbar (Goodchild 2000). Rekommendationer med hänseende på innehåll och utformning om metadata beskrivs av ISO 19115 (Kresse & Fadai 2004). Det finns även specifika rekommendationer om hur ISO 19115 ska tillämpas i Sverige (SIS 2004).

Metadata innehåller flera kvalitetsteman (d.v.s. kvalitetsattribut). Det är rekommenderat att använda samtliga kvalitetsteman i sin kvalitetsbeskrivning, men dock inte obligatoriskt. Kvalitetsteman som beskrivs av ISO 19113 i 19000 serien är *användbarhet*, *syfte*, *ursprung*, *fullständighet*, *lägesnoggrannhet*, *tematisk noggrannhet*, *logisk konsistens* och *temporal noggrannhet* (jfr. Rönnbäck 2004; Wasström et al. 2008). Till varje kvalitetstema finns ett antal mått som benämns *kvalitetsparametrar* t.ex. så beskrivs lägesnoggrannhet av parametrarna absolut- och relativ noggrannhet.

Vanligtvis delas kvalitetsteman upp i två grupper beroende på om innehållet är kvantitativt mätbart. Användbarhet, syfte och ursprung refererar till språklig beskrivning av data och är

därmed kvalitativa till sin natur snarare än kvantitativa (SIS 2004). Sådana kvalitetsteman brukar benämnas översiktliga kvalitetsteman (eng. "Overview quality elements"). De övriga benämns helt enkelt kvalitetsteman (eng. "Quality elements") (jfr. Krasse & Fadai 2004).

Nedan beskrivs samtliga kvalitetsteman enligt ISO 19113 (jfr. Kresse & Fadai 2004):

- *Syfte* beskriver syftet med datamängden och beskriver kort användningsområdet.
- *Användbarheten* hänvisar till tidigare projekt där datamängden har använts.
- *Ursprung* är mest omfattande av samtliga kvalitetsteman då den redogör för livscykeln av datamängden från insamling och bearbetning till dess nuvarande form.
- *Fullständighet* hänvisar till mängden instanser av ett objekt som finns i databasen relativt mängden instanser i verkligheten. Temat beskriver även mängden attributinformation som är angivet förhållande till verkligheten.
- *Lägesnoggrannhet* beskriver objektets relativa och absoluta fel. D.v.s. felet relativt ett annat objekt respektive relativt ett geodetiskt referenssystem.
- *Tematisk noggrannhet* är ett noggrannhetsmått som hänvisar till hur rätt angiven attributinformationen för ett objekt är. Vanligt kvalitetsmått för olika klassningsobjekt t.ex. landskapsklassningsskikt. Ej att blanda ihop med *fullständighet*.
- *Logisk konsistens* refererar till regler som gäller för objekt i datamängden. Exempelvis huruvida två instanser av objekt har korrekt *topologisk* relation till varandra.
- *Temporal noggrannhet* beskriver temporal konsistens samt noggrannhet i metodik för bokföring av tidsuppgifter.

Ett ytterligare kvalitetsmått som ofta används är *aktualitet*. Aktualitet beskrivs inte i ISO 19113 som ett kvalitetstema. Däremot så förknippas det ofta med temporal noggrannhet. Aktualitet beskriver när och hur ofta data/datamängden uppdateras (när data senast konstaterades vara korrekt) och anses därmed av många vara det viktigaste kvalitetsmättet som ska anges i metadata (Rönnbäck 2004).

Det är viktigt att notera att kravet på kvalitet som återges i de olika kvalitetsteman ställs relativt en specifikation från en tänkt användare (Wasström et al. 2009, Jacobsson 2006). Exempelvis ställer en turistkarta som visar vandringsleder inom ett naturreservat inte samma krav på lägesnoggrannhet som en detaljkarta på en byggarbetsplats (Heuvelink 1998).

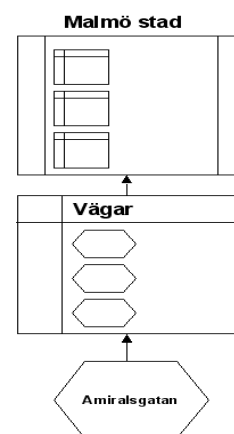
Beskrivning av datakvalitet i metadata är ofta negligerad av producenter. Många producenter anser att metadatastandarder är krångliga eller att kostnaden för kvalitetsdokumentation överväger nyttan för de stora datamängder som produceras (Wayne 2005). Konceptuella ersättningsmetoder har föreslagits istället för kvalitetsteman; dessa syftar till att vara ett dynamiskt verktyg för kommunikation av datakvalitet tillgängligt via webben, exempel på en sådan metod och vidare beskrivning finns i Devillers et al. (2005).

1.2.1.3 Hierarkiska nivåer av data

För att förstå hur kvalitetsteman ska tillämpas, behövs förståelse om hierarkiska nivåer för geografiska data (Devillers et al. 2005). Enklast förklaras nivåerna genom att studera ett praktiskt exempel, vilket även illustreras i figur 3. I data över Malmö kommun finns olika objekttyper som vägar, hus, parker, stationer o.s.v. Varje objektstyp har flera instanser inom kommunen t.ex. föreningsgatan, Amiralsgatan, Slottsparken, Centralstationen. I detta exempel utgör kommunen med samtliga data den högsta hierarkiska nivån, objektstgrupperna såsom vägar utgör mellersta nivån och instanser t.ex. Amiralsgatan den lägsta hierarkiska nivån.

Det blir uppenbart att inte alla kvalitetsteman kan tillämpas på hela hierarkiska skalan t.ex. är inte fullständighet meningsfullt för objekt på instansnivå men högst relevant på objektstypsnivån (Wasström et al. 2008).

Figur 3: Illustration av hierarkiska nivåer av data. Högst upp i hierarkin ses datamängden, som i exemplet, är all tillgängliga data i Malmö kommun. Datamängden består av olika objekttyper t.ex. vägar. Längst ner i hierarkin, befinner sig instanser av objektstyperna t.ex. Amiralsgatan. (Omarbetad från Wasström et al. 2008)



1.2.1.4 Kvalitetsutvärdering av geografiska data

Det finns flera utvärderingsmetoder för att testa en datamängd (Devillers et al. 2005). Enligt Rönnbäck (2004) är det kraven från en användare som styr utvärderingsprocessen d.v.s. godtagbara avvikelser för datamängden jämfört med verkligheten.

ISO 19114 är standarden i ISO 19100-serien som beskriver hur data bör utvärderas samt presenteras. Dock används standarden sällan i utvärderingssammanhang eftersom den kan vara svår att använda (Krasse & Fadai 2004). En tidig svensk studie som använde standarden är Brage (2004).

De ställda kraven från användaren avgör vilka kvalitetstema som ska utvärderas (Krasse & Fadai 2004) t.ex. ett ortofoto, d.v.s. en geometrisk korrigerad bild från någon typ av sensor, bör vara detaljerad i sin geometriska överensstämmelse med verkligheten. Ett ortofoto bör då utvärderas med hänseende på lägesnoggrannhet eftersom det finns höga krav på positionen (Thompson 2000).

Enligt ISO114 ska data utvärderas antingen direkt eller indirekt. Direkt utvärdering innebär att data utvärderas mot verkligheten jämfört med indirekt utvärdering där utvärderingen sker mot behandlad referensdata. Alternativen bedöms med utgångspunkt för hur mycket resurser som finns till förfogande i tid och kapital. Dock är direkt utvärdering med fördel den mest rekommenderade (Tastan & Altan 2002).

Beroende på vilken typ av kvalitet som utvärderas så finns det olika metoder samt rekommendationer, beträffande huruvida data ska jämföras totalt mot hela referensmängden eller om stickprovsundersökning ska utföras (Jacobsson 2003). Det finns flera sätt att välja ut lämpliga stickprov, där antalet referenspunkter väljs godtyckligt eller utifrån en rekommendation. För vidare beskrivning av de olika metoderna både gällande sampling av referenspunkter samt metoder för utvärdering av kvalitetsteman hänvisas till Wasström et al. (2008) och Stehman & Czaplewski (1998).

Slutligen presenteras resultatet från utvärderingen i metadata, dels numeriskt och dels som ett värde för varje kvalitetstema som ingår i utvärderingen (Krasse & Fadai 2004). ISO 19114 beskriver en rekommenderad uppställning på redovisningen av kvalitet. Enklast möjligt är en tabell med uppställning enligt tabell 1. Tabellen visar vilka kvalitetstema som undersökts, med vilket test temat har undersökts, kravet från användarna för att det ska bli godkänt samt resultatet från utvärderingen. (jfr Wasström et al. 2008).

Tabell 1. Resultatet från en kvalitetsutvärdering enligt ISO 19114. Raden högst upp visar kvalitetselementen, metoden, kravet och resultaten. I exemplet har objekt utvärderats med avseende på lägesnoggrannhet, både absolut och relativ. Kvalitetsmättet är medelfelet, vilket är samma sak som standardavvikelsen i en normalfördelad datamängd. Kravet är specificerat från användaren Resultatet talar för att data blev godkänd i hänseende på absolut noggrannhet men inte läges noggrannhet. (Omarbetad från Wasström et al. 2008)

Kvalitetstema	Kvalitetsparameter	Kvalitetsmått	Krav	Utvärderingsmetod	Utvärderingsresultat	Kvalitetsvärde
Lägesnoggrannhet	Absolut noggrannhet	Medelfel stickprov	<3m	Slumpmässig	Medelfel 2 m	Godkänt
Lägesnoggrannhet	Relativ noggrannhet	Medelfel stickprov	<1 m	Fullständig kontroll	Medelfel 3 m	Underkänd

1.3 Metoder för kvalitetsutvärdering av geografisk analys

I föregående avsnitt diskuterades kvalitet av geografiska data, hur det definieras och används. Detta avsnitt redogör för hur fel som uppkommer fortplantas i en geografisk analys och hur en kvalitetsutvärdering kan utföras utav analysen.

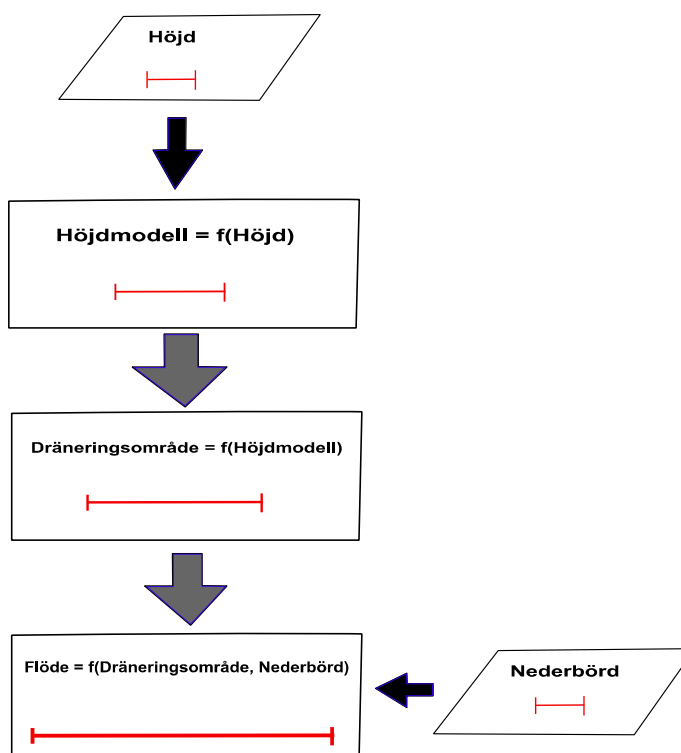
1.3.1. Felfortplantning

Den osäkerhet som data är behäftad med kommer att påverka kvaliteten på resultatet av analyserna och därmed även tillförlitligheten. Analysens totala osäkerhet kommer till stor del vara beroende av osäkerhet i indata och modellernas känslighet för osäkerheten (Zhang & Goodchild 2002, Heuvelink 1998).

En modell som beräknar vattenflödet i ett vattendrag är ett utmärkt exempel för att konkretisera hur en felfortplantning fungerar i praktiken, exemplet illustreras även i figur 4. Trots att flödesmodeller kan vara väldigt avancerade så är de viktigaste faktorerna nederbörden och topografin, där topografin i sin tur både påverkar dräneringsområdets storlek samt lutningen. Topografin representeras vanligtvis av digitala höjdmodeller (*DEM:er*) som interpolerats utifrån höjdmätningar i punkter. Höjdmätningarna är förenade med ett mättningsfel som är beroende av kvaliteten på instrumenten samt metodiken för inmätningen. I interpolationsmodellen modelleras ett raster fram. Höjdmodellen har nu ett fel som är beroende av inmätningens fel samt fel som har uppstått i interpolationsmodellen. Utifrån den digitala höjdmodellen ska dräneringsområdet skattas; där det finns olika metoder som är mer eller mindre avancerade. Dräneringsområdet erhåller alltså en osäkerhet som är beroende av förenklingar i modellen samt felet i höjd-modellen, som i sin tur är beroende av förenklingar i interpolationsmodellen och mättnings-felen. Vattenflödet skattas slutligen utifrån en modell där nederbörden och dränerings-området storlek är ingående faktorer. Givetvis är även nederbördsdata förenat med någon typ av mätfel. Flödet är även beroende av hur mycket av vattnet som infiltreras av marken, vilket ytterliggare ökar osäkerheten i analysen (Oksanen & Sarjakoski 2005).

Figur 4 :

Figuren visar ett flödesschema för beräkning av vattenflöde i ett vattendrag. Data representeras av romber, modeller av rektanglar. Osäkerheten i de olika stegen symboliseras med röda linjer. (Omgjord från Leung et al. 2004)



Osäkerheten i analysen kan bli väldigt stor vilket i värsta fall kan göra resultaten så pass osäkra att de inte går att använda. Användbarheten avgörs utifrån syftet och kraven (Rönnbäck 2004). Precis som vid kvalitet av data så ställs ofta olika krav på kvalitet av geografiska analyser beroende på användningsområdet. I GIS-analyser finns ibland en kravspecifikation som måste uppfyllas. Utifrån kravspecifikationen kan projektledaren bedöma hur pass detaljerade dataset och modeller som behövs för att genomföra projektet och ändå hålla kostnaden så låg som möjligt (Frank 2008).

Vid felfortplantningsanalys är alltså granskaren intresserad av att studera osäkerheten i den geografiska analysen baserat på osäkerhet i indata. Av särskilt intresse är variansen och medelfelet (Zhang & Goodchild 2002). Antalet möjligheter att angripa felfortplantningen är flera men det finns i princip två huvudkategorier: *analytisk bestämning* och *simuleringar* (Heuvelink 1998).

2.3.1.1 Analytisk bestämning

Enligt Heuvelink (1998) är en fullständigt analytisk bestämning i princip omöjlig om inte modellerna som nyttjar data är helt linjära. I praktiken kan dock approximationer göras. Analytiska lösningar är ofta att föredra framför simuleringar eftersom de är mindre beräkningsintensiva. Problemet med analytiska bestämningar är om sambanden är väldigt komplexa.

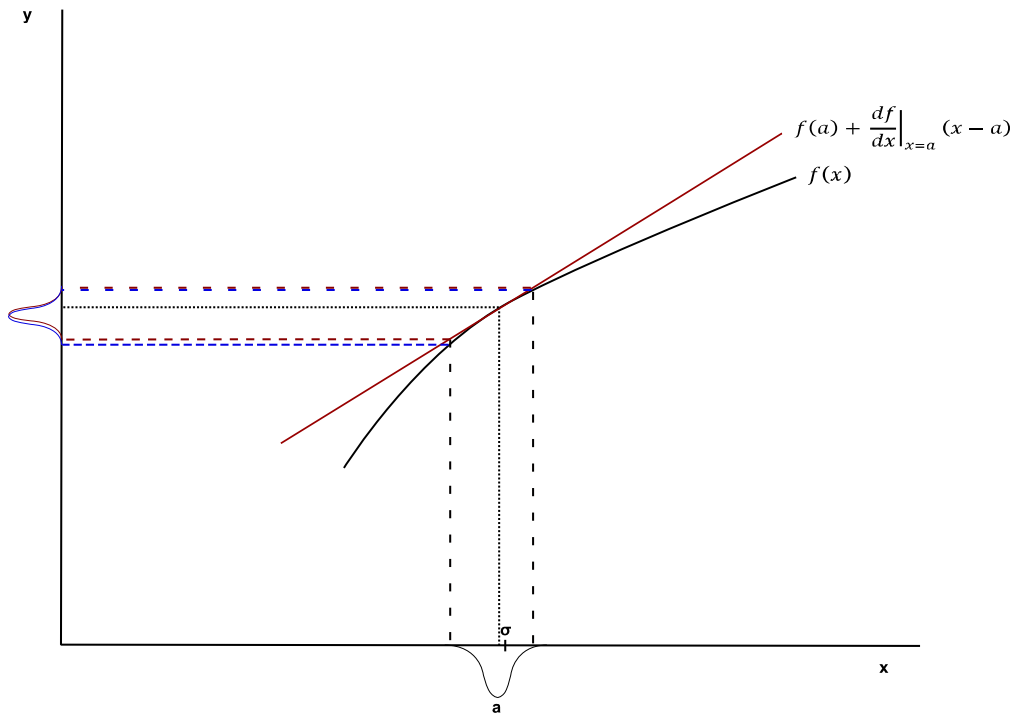
Heuvelink (1998) rekommenderar tre analytiska metoder för att genomföra en felfortplantningsanalys: *första ordningens Taylorutveckling (FOT)*, *andra ordningens Taylorutveckling (AOT)* och *Rosenbleuth metoden*. Rosenbleuth metoden används stort sett aldrig i praktiska tillämpningar, men kan användas då modellen inte är kontinuerligt deriverbar för alla värden x (Heuvelink 1998).

Taylorutveckling är en metod för att dela upp en matematisk funktion $f(x)$ nära en punkt a polynom av olika grad n . Desto högre grad på polynomet ju mer efterliknar polynomet funktionen kring punkten a .

$$f(x) = f(a) + \left. \frac{dF}{dx} \right|_{x=a} (x - a) + \frac{1}{2!} \left. \frac{d^2F}{dx^2} \right|_{x=a} (x - a)^2 + \dots + \frac{1}{n!} \left. \frac{d^n F}{dx^n} \right|_{x=a} (x - a)^n \quad (2)$$

I en strikt utveckling av Taylorpolynom medföljer alltid en restterm av högre ordning. Vid felfortplantningsanalys kan dock resttermen negligeras (Heuvelink 1998).

Fortplantningen av osäkerhet i indata genom en modell $f(x)$ kan skattas genom Taylorpolynom. Grafiskt visar det sig att normalfördelningen av sannolika värden för $f(x)$ i punkten a står i proportion till intervallet på Taylorpolynomet som motsvarar variabeln x :s fördelning i punkten a . Detta illustreras i figur 5, där första ordningens Taylorpolynom (FOT) $f(a) + \left. \frac{df}{dx} \right|_{x=a} (x - a)$ anpassats till funktionen $f(x)$.



Figur 5: Illustration av första ordningens Taylor utveckling till en funktion $f(x)$ i punkten a . Den röda fördelningen visar den skattade felfördelningen, den blåa visar den "sanna" d.v.s. den fördelning som felet i $f(x)$ skulle ha haft om det varit möjligt att beräkna osäkerheten direkt från funktionen. Omarbetad från Heuvelink (1998).

Skattningen av osäkerheten i $f(x)$ skiljer sig mer från den "sanna" osäkerheten om $f(x)$ är konvex växande uppåt än konkavt växande neråt.

Utan att gå in på rigorösa matematiska bevis (intresserade hänvisas till Heuvelink 1998) hur $\sigma_{f(x)|_{x=a}}$ beräknas numeriskt, så visar det sig att felet är proportionellt till derivatan av funktionen $f(x)$, med hänseende på variabeln x värde i punkten a , multiplicerat med medelfelet för variabel x . (Heuvelink et al 1998):

$$\sigma_{f(x)|_{x=a}} = \sqrt{(\sigma_{x|x=a})^2 \left(\frac{df}{dx}\bigg|_{x=a}\right)^2} \quad (3)$$

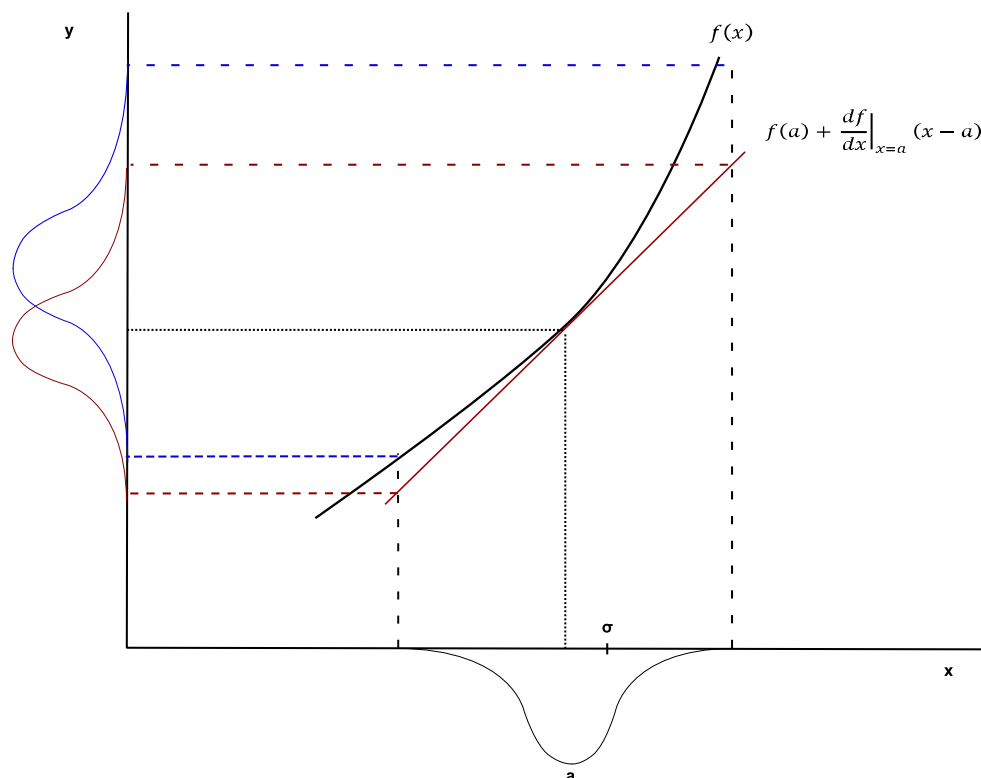
Om modellen F är beroende av fler variabler x_1, x_2, \dots, x_n summeras felet, förutsatt att variablerna är oberoende av varandra. Därmed erhålls ett allmänt uttryck för osäkerheten i analysen σ_F beroende av de ingående felet av variablerna i punkten a .

$$\sigma_F = \sqrt{(\sigma_{x_1|x=a})^2 \left(\frac{df}{dx_1}\bigg|_{x=a}\right)^2 + (\sigma_{x_2|x=a})^2 \left(\frac{df}{dx_2}\bigg|_{x=a}\right)^2 + \dots + (\sigma_{x_n|x=a})^2 \left(\frac{df}{dx_n}\bigg|_{x=a}\right)^2} \quad (4)$$

Genom att skriva om ekvation 4, så kan enklare anpassningar göras som förenklar implementeringen av felfortplantningen (Wasström et al. 2008).

Att uppskatta felens fortplantning med en *FOT* är i särklass den vanligaste metoden inom naturvetenskaper och teknik. Metoden benämns vanligtvis *den speciella felfortplantningslagen* (Ollvik 2002). Exempelvis beskriver Frank (2008) en aktuell fallstudie, där metoden används för att studera huruvida en bro klarar de tänkta kraven gällande nederbörd och flöden. Metoden har även använts vid storskaliga GIS analyser som vid uppskattning av osäkerheten hos en luftkvalitetsmodell (Hwang 1997).

FOT har klara fördelar gällande beräkningsintensitet och enkelhet. Trots detta uppkommer en del svårigheter. Som beskrivet tidigare, så uppskattar *FOT* första derivatan av funktionen i punkten a och därefter beräknas felfördelningen i $y = f(a)$ från intervallet på tangenten som motsvarar felfördelningen i a . Om felet i x är stort och modellen är olinjär så kommer den uppskattade felfördelningen för $y = f(x)$ skilja sig avsevärt från den sanna (figur 6). Av den anledningen anpassas ibland högre gradens Taylor polynom (Heuvelink 1998), som t.ex. andra ordningens Taylorutveckling (*AOT*).



Figur 6 : Samma resonemang som i figur 6, skillnaden är att felet i x är större vilket gör att skattningen av felet i $A(x)$ avviker mer från det "sanna" värdet. Omarbetad från Heuvelink (1998).

Andra ordningens Taylor (AOT) kan användas om felet i variabeln är för stora för att använda *FOT*. Metoden ger dessutom ett exakt värde om funktionskurvan är en andragradsekvation. Problemet med *AOT* är att metoden är betydligt svårare att beräkna och implementera. Matematiska härledningar och beräkningsmetoder för *AOT* finns i Forsling & Neymark (2004); anpassning av densamma till geografiska analyser beskrivs i Heuvelink (1998).

Givetvis så finns det ingen teoretisk gräns för ordningstalet av Taylorutveckling som kan användas. Svårigheten ligger i att komplexiteten ökar markant för varje ordning, vilket till slut leder till en nästintill omöjlig analytisk bestämning. Om de ingående stegen i analysen är väldigt komplexa så är det bättre att simulera fram en felfördelning istället för att uppskatta den analytiskt (Heuvelink 1998, Rönnbäck 2004, Wasström 2008).

2.3.1.2 Simulering

I början av kapitlet beskrevs en konceptuell modell för att beräkna vattenflöde utifrån dräneringsområde och nederbördsdata. Framtagandet av ett dräneringsområde baseras på komplicerade analyser av en höjdmödel; analyser så pass komplicerad att en analytisk bestämning ej är lämpligt att utföra (Oksanen & Sarjakoski 2005). Vid sådana analyser tillämpas med fördel istället simuleringmetoder (Zhang & Goodchild 2002).

Den vanligaste typen av felsimulering inom geografiska analyser är *Monte Carlo simulering (MCs)* (Zhang & Goodchild 2002; Heuvelink 1998). Principen för en MCs är att ett antal itereringar av analysen genomförs, där ett slumpmässigt fel adderas för varje iterering till varje variabel och det slumpmässiga felet extraheras från felets normalfördelning.

Resultat utav en MCs ger inte ett uppskattat medelfel som de analytiska metoderna gör. Istället produceras ett stort antal analysresultat med olika värden. Dessa resultat används för att skatta en resultatfördelning d.v.s. med vilka värden resultatet sannolikt kommer att anta. Fördelningens standardavvikelse kommer då att motsvara medelfelet.

Styrkan med MCs är att oavsett hur komplicerade sambanden är, så kan felet uppskattas relativt enkelt (Oksanen & Sarjakoski 2005), förutsatt då att granskaren har tillgång till en dator med kapacitet att utföra simuleringarna. Nackdelen är att ett stort antal itereringar måste utföras för att få bra resultat, rekommenderat mellan 500-1000 (Zhang & Goodchild 2002), men med nuvarande datorutveckling bör det vara ett smärre problem.

Nackdelen med MCs är att felet inte beräknas direkt utan istället får skattas från en fördelning av osäkerhetsvärden. Ett vanligt missförstånd är att metoden kan användas även om de ingående variablerna är beroende av varandra, om beroende föreligger måste även ett beroende modelleras i indata till modellen (Harrie 2010).

Om sambanden är enkla så är en simulering onödig. Då är analytiska metoder att föredra framför MCs (Heuvelink 1998).

2.2.4. Test av rumslig modell och färdig produkt

Bra kvalitet på geografiska data är viktig för att modellerna ska kunna producera pålitliga resultat. Lika viktigt är det att modellerna representerar de fenomen som ska modelleras. Om inte modellen/modellerna är tillräckligt representativa så påverkar datakvaliteten föga i

analysen d.v.s. då blir resultatet otillräckligt oavsett kvaliteten på data (Wasström et al. 2008).

I en strikt osäkerhetsanalys bör osäkerheten i modellen tas i beaktande d.v.s. hur väl modellen svarar mot att representera det verkliga fenomenet. Ofta så finns inga sådana uppgifter (Heuvelink et al. 1998), vilket givetvis är en stor brist.

Istället för att studera modellen/modellernas osäkerhet så kan en kvalitetsgranskare utvärdera analysens resultats överensstämmelse med verkligheten och därmed utvärdera något om tillförlitligheten. Inom naturvetenskap och teknik tillämpas ofta *kvantitativa* undersökningar mot referensdata, där statistiska test används för att pröva om frammodellerade värden stämmer överens mot de faktiska värdena i samma område.

2.2.4.1 Kvantitativ undersökning

Kvantitativa undersökningar är samlingsnamn som beskriver metoder att systematiskt samla in data, jämföra mot andra data och utföra statistiska tester för att pröva hypoteser av olika slag.

Ibland benämns kvantitativa undersökningar som objektiva d.v.s. att forskaren sitter inne med objektiva data som granskas utan subjektiva bedömningar. Istället undersöks om sambanden är statistiskt signifikanta (Nationalencyklopedin *kvantitativa metoder* 2009). För att sammanfatta, en kvantitativ undersökning kan formuleras med beskrivande statistik såsom väntevärde, standardavvikelse, korrelation o.s.v. (Nationalencyklopedin *statistik* 2009).

Det finns ett flertal olika metoder för att testa hur väl en modell utvärderar mot faktiskt utfall. Här beskrivs tre vanliga metoder: korrelationsanalys, MDA (Mean Deviation analysis) och RMSE (Root Mean Square Evaluation).

Korrelationsanalys är en vanlig statistisk metod inom naturvetenskap och teknik för att undersöka samband mellan två populationer (NE 2009 *korrelation*). A (Di Lucio et al. 2005, Hildebrand et al. 2008). Metoden är besläktad med linjär regression med skillnaden att det inte antas något observerat beroende av data. Måttet på korrelationen anges med korrelationskoefficienten r . Ett värde på nära 0 visar på liten korrelation medan ett värde nära 1 indikerar på en hög.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (5)$$

Hildebrand et al. (2008) rekommenderar en annan statistisk metod för att testa sambandet mellan en modells skattade värde och utfallsdata nämligen MDA (Mean Deviation Analysis). Felet, för varje lokal x , beräknas som medelvärdet av den absoluta skillnaden mellan modellens värde \hat{y} och det faktiska värdet y . Metoden är enkel och fungerar oavsett

sambandets utformning. Intuitivt påvisar ett *MDA* värde nära 0 ett litet fel, omvänt ett högt *MDA* ett stort fel.

$$MDA = \frac{\sum_{i=x}^n |\hat{y}_i - y_i|}{n} \quad (6)$$

Ett annat vanligt mått i GIS för att jämföra rumsliga ytor är RMSE (Root Mean Square Evaluation). Metoden ger ett mått på medelfelet i modellen, problemet är att den är känslig för uteliggare eftersom avvikelsen i varje jämförelse kvadreras (Zhang & Goodchild 2002)

$$RMSE = \sqrt{\frac{\sum (\hat{y}_i - y_i)^2}{n}} \quad (7)$$

2. Rumsliga prognosmodeller och planering av räddningstjänst

2.1. Allmänt om rumsliga prognosmodeller

Prognosmodeller återfinns inom flera discipliner och vetenskaper. Det kan röra sig om ekonomiska prognosmodeller, industriella hållbarhetsmodeller, populationsmodeller, väderprognosmodeller o.s.v. Intressant i den här rapporten är rumsliga prognosmodeller, där det modellerade värdet kan knytas till en position i ett geodetiskt referenssystem. Exempel på rumsliga prognosmodeller är SMHI:s (Sveriges Meteorologiska och Hydrologiska Institut) hydrologiska sannolikhetsprognoser (SMHI 2010) och IPCC:s (Intergovernmental Panel of Climate Change) klimatmodeller (IPCC 2010).

Processen för att skapa och implementera en rumslig prognosmodell är komplicerad. Ofta används komplexa matematiska och statistiska ekvationer som ska motsvara observerade samband från årtal av forskning. Om sambanden är väldigt komplexa så genomförs i regel förenklingar i efterhand för att underlätta implementeringen av modellerna i datorprogrammet/en som ska köra modellen.

Prognosmodeller används i flera tillämpningsområden, allt ifrån trafikflygning till planering av skyddsområden för fridlysta arter. Väderleksmodeller används inom flygindustrin för att undvika turbulens och farliga vädermassor (Evans 2001). Vid planering av naturskyddsområden i områden med hög befolkningstäthet används befolkningsmodeller för att kartera vilka fridlysta arter som kan drabbas av en befolkningsexpansionen (Beissinger & Westphal 1998).

Ovanstående exempel är bara ett axplock för att understryka variationen av prognosmodeller och deras användningsområden. I denna rapport granskas en prognosmodell för skattning av fordonsolyckor på vägar och korsningar. Prognosmodeller för vägolyckor är vanliga (Hildebrand et al. 2008), både i Sverige och utomlands, och används vid flertalet tillämpningar, exempelvis vid planering av nya vägar. Ett annat användningsområde är planering och utvärdering av räddningsresurser.

2.2. Prognosmodell för planering av räddningstjänst

De vanligaste olyckstyperna som även krävt flest arbetstimmar under år 2008 var bränder och trafikolyckor. Förutom skogsbränder, så var olyckor orsakade av människor klart överrepresenterade i olycksstatistiken. Skogsbränder hade ökat mest under året och var även den kategori med flest insatser. Dock förekommer en betydligt större andel personskador vid trafikolyckor och bränder i byggnader än vid skogsbränder. Så mycket som 90 % fler personskador inträffar vid brand i byggnad än vid övrig brand. Den största andelen dödsfall sker vid trafikolyckor (Lundqvist et al. 2008). Trafikolyckor och bostadsbränder är alltså

kritiska olyckstyper ur hälsosynpunkt varvid en snabb utryckning vid dessa olyckskategorier kan vara livsavgörande.

Lagen om skydd mot olyckor (SFS 2003: 778) beskriver hur kommuner och länsstyrelser ska agera vid olyckor och räddningstjänst. MSB har som uppgift att hjälpa vid efterlevnaden av lagen och därmed bistå med verktyg och expertis rörande räddningstjänst.

Då räddningstjänst utförs så behövs räddningsresurser. Räddningsresurser är de enheter som används vid en utryckning t.e.x. släckbilar, brandutrustning och stegbilar. Det är av yttersta vikt att det finns en god tillgänglighet till räddningsresurser så att olycksplatsen kan nås så fort som möjligt. Av ekonomiska skäl finns det dock ett övre tak på hur mycket räddningsresurser ett administrativt område får tillgå. Därmed är det viktigt att planera sina räddningsresurser så att placeringen inom området blir optimal i förhållande till olycksrisken.

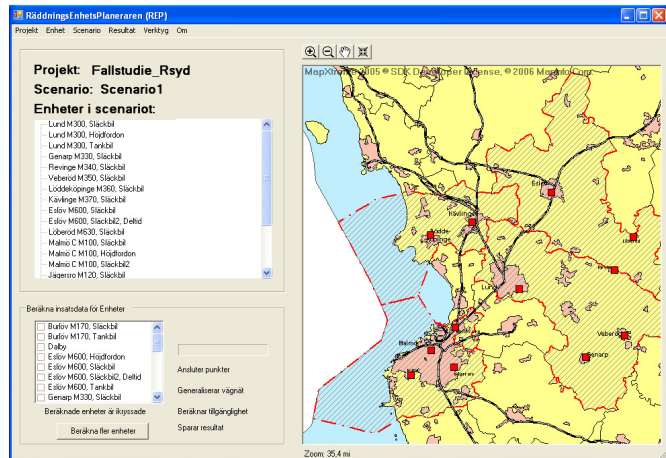
2.2.1. BeRädd och ROPA

BeRädd är ett projekt inom MSB som syftar till att ta fram ett planeringsverktyg för räddningstjänsterna. Projektet är ett samarbete mellan Linköpings universitet, Lunds universitet, statens Väg- och TransportforskningsInstitut (VTI), KnowIT, Tillgänglighetsanalys och Karlstads universitet. Ursprungligen var projektet uppdelat i två delar, båda finansierade av dåvarande Räddningsverket. I ett av projekten, *Geografisk tillgänglighetsanalys*, utvecklades ett tidigt verktyg *REP* (RäddningsEnhetsPlaneraren) som användes vid beräkning av tillgänglighet för Räddningstjänsten d.v.s. tillgänglighet till räddningsresurser utifrån olyckor .

År 2008 beslöts att ett nytt verktyg skulle tas i bruk, som en fortsättning av REP där lärdomar och expertis från båda projekten skulle tas till vara. Verktyget är ett datorprogram och går under namnet *ROPA* (Resurser, Optimering, Planering, Analys (eller Allokering)). Det ska hjälpa räddningstjänsterna att planera och utvärdera hur de använder sina resurser för bästa beredskap.

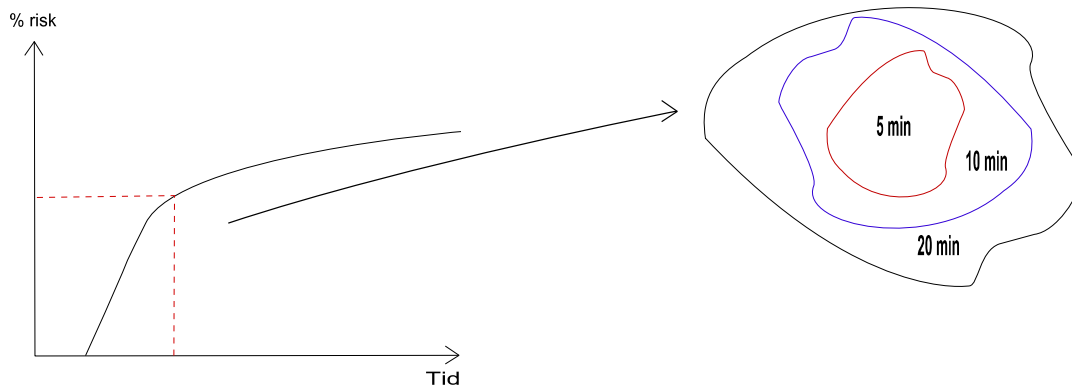
Programmet är konstruerat så att det ska kontrolleras med hjälp av ett enkelt gränssnitt liknande det som använts vid REP, se figur 7.

Figur 7: Figuren visar huvudmenyn i REP när ett scenario startats. Enheterna (räddningsresurser) som definierats i scenariot visas till höger i kartan som röda kvadrater. (Dahlgren et al 2007)



För varje scenario i ROPA definieras ett undersökningsområde, enheter som ska vara med i planeringen och risktypen som ska undersökas.

Programmet använder sig sedan av två analysmetoder för att planera räddningsresurserna: simulering och optimering. Simuleringen beräknar styrkeuppbyggnaden, täckningsgraden och medelinsatstiden för enheterna. Styrkeuppbyggnaden är en beräkning av mängden räddningspersonal som kan nå ett område på en viss tid. Täckningsgraden är den minsta körtiden som krävs för att få ut samtliga resurser som behövs för den avsedda olyckstypen. Täckningsgraden beräknas utifrån varje enhets medelinsatstid till varje riskruta där medelinsatstiden beräknas med nätverksalgoritmer baserat på tillgängliga vägar och positionen för enheten. Figur 8 illustrerar andel risk i ett område som kan nås under kortast tid med tillräckliga resurser (Dahlgren et al. 2007).

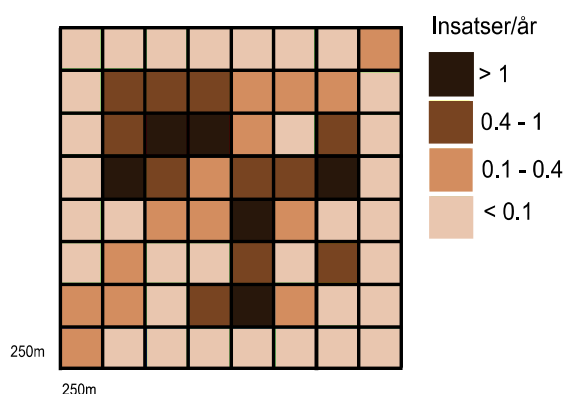


Figur 8: Grafen till vänster visar en hur en täckningsgradskurva kan se ut beroende på tid. Fun. Till höger illustreras det område som motsvarar det som är innanför det streckade området i grafen. Omarbetad från Dahlgren et al (2007).

Optimeringen innebär att programmet kan föreslå en alternativ placering utav räddningsresurserna för att optimera andel risk som teoretiskt kan nås utav de tillgängliga enheterna på så kort tid som möjligt. Programmet använder täckningsgraden och styrkeuppbyggnaden från simuleringen för att genomföra detta.

2.2.1.1. Prognoskikt

Risken uppskattas utifrån flera prognoskikt. Prognoskikt är rasterskikt med storleken 250 * 250 m eller 1000*1000 (i skrivande stund har upplösning på prognoskikten inte bestämts), där varje cells värde motsvarar en prognos för att en olycka ska inträffa av en viss olyckstyp, (se figur 10) (Dahlgren et al. 2007).

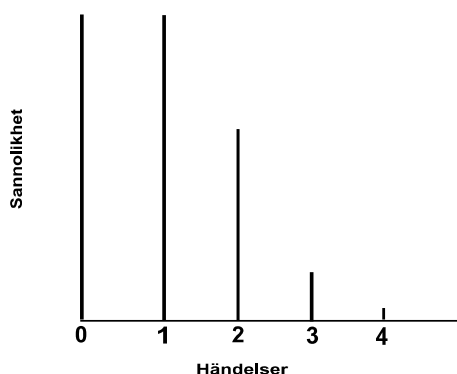


Figur 9: Exempel på ett prognoskikt.

Prognoskikten används för att predicera antal olyckor per dag i varje ruta. Sannolikheten för att en vägolycka ska inträffa inom ett område under en viss tid kan förklaras med Poissonfördelning (Blom et al 2005). Detta innebär att varje ruta har en Poissonfördelning av sannolika olycksvärden per dag, där det framräknade prognosvärdet är det mest sannolika och motsvarar medelvärdet μ i fördelningen $Po(\mu)$.

Poissonfördelningen är besläktad med binomialfördelningen. Binomialfördelningen uppstår då antalet händelser av samma händelsetyp inträffar vid oberoende upprepning av samma typ av händelseförsök. Binomialfördelning antar utseendet utav en Poissonfördelning då sannolikheten minskar att en händelse ska inträffa på ett ställe under en given tidsperiod. Typiska händelser som kan förklaras med en Poissonfördelning är antal olyckor på en väg under en dag, sönderfall av en radioaktiv partikel under en viss tid o.s.v., se figur 10 för en Poissonfördelning med $\mu=1$.

Figur 10: Poissonfördelning ($PO(\mu)$) då $\mu = 1$ (Omarbetad från Blom et al.2005)



ROPA använder sig av fyra prognosskikt för att beräkna tillgänglighet till räddningsresurser, två som behandlar bränder i bostäder och fritidshus samt två prognosskikt som beskriver fordonsolyckor på vägar.

Prognosskikt, brand i bostad

Jaldell (2009) beskriver hur förväntat antal bränder i bostäder kan modelleras. Modellen utgår från tre huvudkategorier av förklarande variabler: bostadsbyggnadsvariabler, befolknings- och familjedata samt arbetslöshets-, utbildnings-, inkomst- och ohälsodata. Tillsammans utgör de över 150 variabler som möjligt kan påverka antal bränder. Det visade sig att flertalet av variablerna är korrelerade med varandra. I och med det stora antalet har Jaldell (2009) använt sig av faktoranalys, som är en statistisk metod för att kombinera förklarande variabler, där korrelerade variabler togs bort. Därmed reducerades antalet variabler till 11 st. Dessa variabler är: antal lägenheter i flerfamiljshus, antal rad/par/kedjehus, antal friliggande villor, antal bostadsrätter i flerbostadshus, total befolkning, befolkning i åldern 0-19, antal födda utanför Norden, antal ohälsodagar totalt, antal eftergymnasialt utbildade, antal arbetslösa och antal grundskola- och gymnasieutbildade. Variablerna har sedan testats i en regressionsanalys mot förekomster av inträffade bränder för att se påverkan på antalet bränder. I analysen visade sig att samtliga variabler hade ett statistiskt signifikant samband med antalet inträffade bränder.

Utifrån modellen har testkörningar gjorts där prognosskikt med 1*1km stora rutor täckande hela Sverige producerats. Ursprungligen var det för avsikt att skikten skulle vara 250*250 m, men omprioriteringar gjordes pga avvikelser mellan fastighetspositionering och insatsdata.

Prognosskiktet har utvärderats mot insatsdata. Från testet framgick att förklaringsgraden blev 0,87 vilket är högt.(Jaldell 2009).

Prognosskikt för fordonsolyckor på motortrafikerade vägar och korsningar

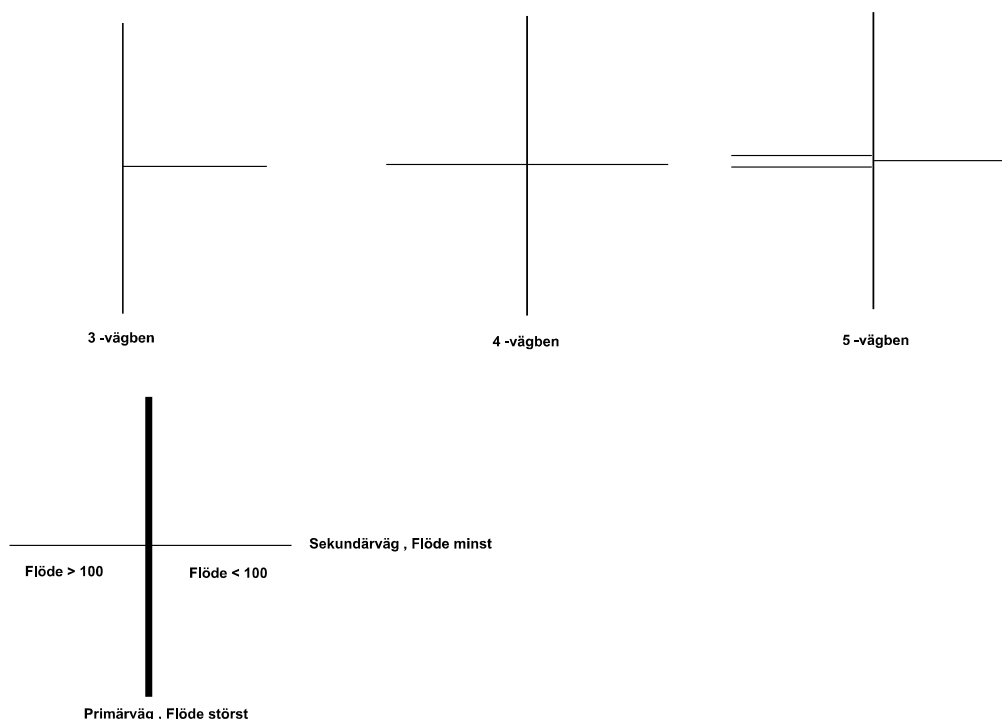
Prognosskikten för fordonsolyckor är uppbyggda av flera funktioner (dessa visas i sin helhet i fallstudien). Modellerna delar in olyckor på korsningsolyckor samt olyckor på väglänkar. Parametrarna som bestämmer olycksfrekvensen är ÅDT (ÅrsMedelTrafikflöde), längd på väglänken, normalvärde, relativ olycksrisk, olycksstatistik för kommunen och olika korsningsparametrar. Nedan beskrivs de olika parametrarna ÅDT, normalvärde och relativ olycksrisk:

ÅDT: En skattning om hur många fordon som kör på en väg varje dag. Principen för framtagandet är att vägnätet delas upp i ett stort antal sträckor som antas ha homogent flöde. Flödet på dessa sträckor mäts sedan flera gånger om året med elektronisk utrustning. Därefter modelleras resterande flöden utifrån dessa mätningar. Vägverket ser till att det finns uppgifter om årsmedelflöde på hela det statliga vägnätet (Vägverket *Allmänt om Vägverkets trafikmätningar* 2010).

Normalvärde: Normalvärdet är ett genomsnittligt olycksvärde som beräknas utifrån samband mellan faktiska olyckor och egenskaper för olika vägar och korsningar (Forsman & Vadeby 2009).

Relativ olycksrisk: Relativ olycksrisk är liksom normalvärdet ett genomsnittligt olycksvärde som är skattat utifrån olika samband. Skillnaden mot normalvärdet är att skattningen är mer osäker och i viss mån anses det vara ett schablonvärde. (Forsman & Vadeby 2009)

Olyckor på väglänkar är indelade in i tre kategorier, olyckor på statliga vägar, olyckor på kommunala vägar samt olyckor på enskilda vägar. Vid beräkning av olyckor på statliga vägar används normalvärdet, ÅDTa och längden på väglänken. Kommunala vägar har inga uppgifter om ÅDT, istället används relativ olycksrisk, väglängden samt olycksstatistik för kommunen som undersöks. Vid skattning av olyckor på enskilda vägar så utnyttjas bara väglänkens längd. Modellerna för de statliga korsningarna använder ÅDT på de inkommande vägarna samt typ av korsning. (d.v.s. antal vägben i korsningen och fördelning av trafiken i korsningen d.v.s. om korsningen är sned eller balanserad, se figur 11.)



Figur 11: De tre översta figurerna visar exempel på korsningar med olika antal vägben. Nedre figuren visar exempel på sned korsning. Om inkommande flöde på båda de sekundära vägbenen understiger eller överstiger 100 så anses den vara balanserad.

Kommunala korsningar har liksom det kommunala vägnätet ingen information om flöde till korsningen. Modellen för olyckor på kommunala korsningar använder istället relativ olycksrisk för korsningen samt statistik om olyckor i kommunen.

I kategorin korsningar görs dessutom indelningar beroende på om korsningen är en vanlig tre/fyrvägs korsning eller om korsningen är en cirkulationsplats eller *trafikplats*.

Utifrån de olika olyckstyperna så beräknas antal singel- och flerfordonsolyckor per dygn med olika typer av korrigeringar. Dessa olycksrisker utgör prognoskikten som sedan används i ROPA.

Indata till prognoskikt för fordonsolyckor

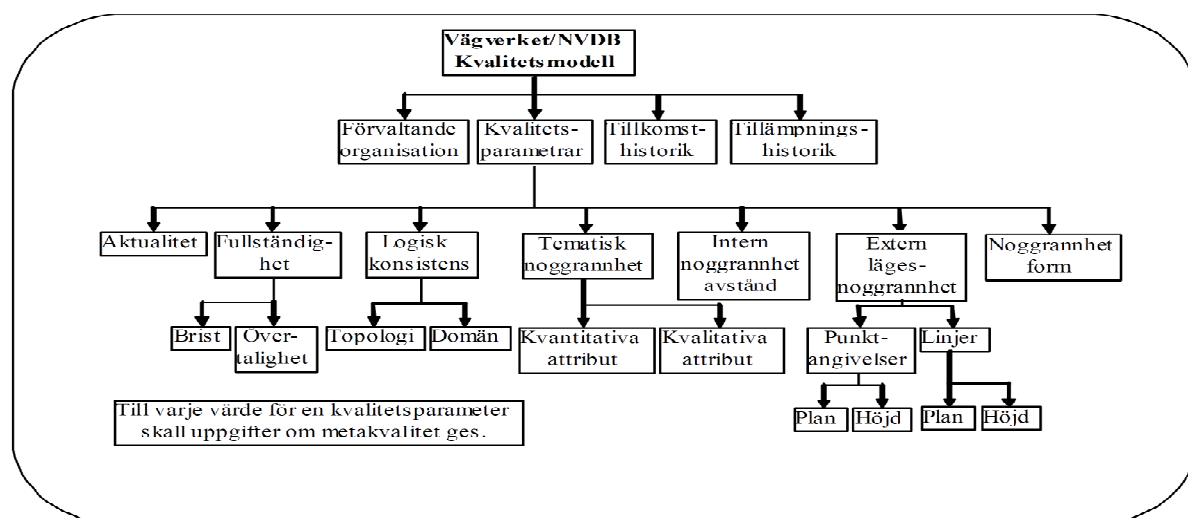
Indata till prognosmodellerna för fordonsolyckor är nationella vägdatan NVDB. NVDB är en geografisk databas vars syfte är att innehålla digital information om alla vägar i Sverige som beträds av motortrafik. NVDB är statligt finansierad och ansvarar för att rikets vägnät finns dokumenterat med god kvalitet. Databasen räknas som en av grunddatabaserna i samhället liksom t.e.x. statistiska centralbyråns SCB befolkningsdatabas och sjöfartsverkets djupdatabas. De som bidrar med insamlingen till NVDB är vägverket, skogsstyrelsen, lantmäteriet, kommunerna, försvaret och privata bidragsgivare, huvudman och samordnare av insamlingen är vägverket. Vägnätet är uppdelat i *enskilda vägar*, *kommunala vägar* samt *statliga vägar*. Tillsammans utgör de ca 560 000 km väg .

Geometrin i databasen består av linjer och noder som är positionerat i ett geodetiskt referenssystem, det kan levereras i de nationella referenssystemen SWEREF 99 och RT90. Knutet till geometrin finns attribut som t.ex. beskriver vägtyp, hastighetsbegränsning och breddklass.

NVDB följer ISO 19113 standarden för kvalitet av geografiska data. Figur 12 visar kvalitetsmodellen som NVDB följer. För att se om data uppfyller kraven används en avvikelsekvote, vilken helt enkelt är andelen data i datamängden som inte uppfyller ett gränsvärde delat med totala mängden data. Utvärderingen genomförs för varje kvalitetstema och parameter. För kvalitetsteman som fullständighet och tematisk noggrannhet beräknas avvikelseknoten relativt lätt eftersom kontrollerna enbart testar huruvida de är "rätt" eller "fel" angivna. Däremot vid utvärdering av kontinuerliga teman som lägesnoggrannhet är det svårare att ange ett sant värde, istället används konfidensintervall från maximal avvikelse. Avvikelsekvoten ger då ett procentuellt avvikelse från gränsvärdena.

Kvalitetskraven på vägarna är direkt knutet till sin inbördes rangordning av funktionalitet. NVDB har satt upp 4 olika kravklasser på kvalitet som styrs av vilken funktionell vägklass som vägen har. Vilket exempelvis innebär att det inte är samma krav på en enskild väg som en statlig.

Huvuddelen av kvalitetskontrollen utförs av personal på NVDB genom olika beräkningsmetoder och kontroller. En del kvalitetskontroller måste dock vara utförda av producenterna innan data levereras till NVDB (Mats-Ove Mattson 2009).



Figur 12: NVDB:s kvalitetsmodell (Vägverket 2009).

De attribut som används från NVDB i prognosmodellerna är: trafikflöde (ÅDT, finns i tre utformningar, ÅDT/fordon, ÅDT/lastbilar och ÅDT/axelpar), maximal hastighetsgräns, kommunkod (olika kommuner i Sverige har olika kommunkoder), tätortsid, väghållare (d.v.s. statlig, kommunal eller enskild), vägtyp (som är en kod mellan 1 och 6 som redogör för vad det är för väg), vägens bredd (indelat i olika breddklasser) samt funktionell vägklass (vilken rangordnar betydelsen av vägen).

Utöver ovanstående attribut behövs längden på väglänken i flera av beräkningarna. Vilken kan beräknas med Pythagoras sats utifrån positionen på noderna som bygger upp väglänken.

Hastighet, vägtyp, vägbredd och funktionell vägklass används för att erhålla normalvärden och relativa olycksrisker för länkarna och korsningarna. (Forsmann & Vadeby 2009)

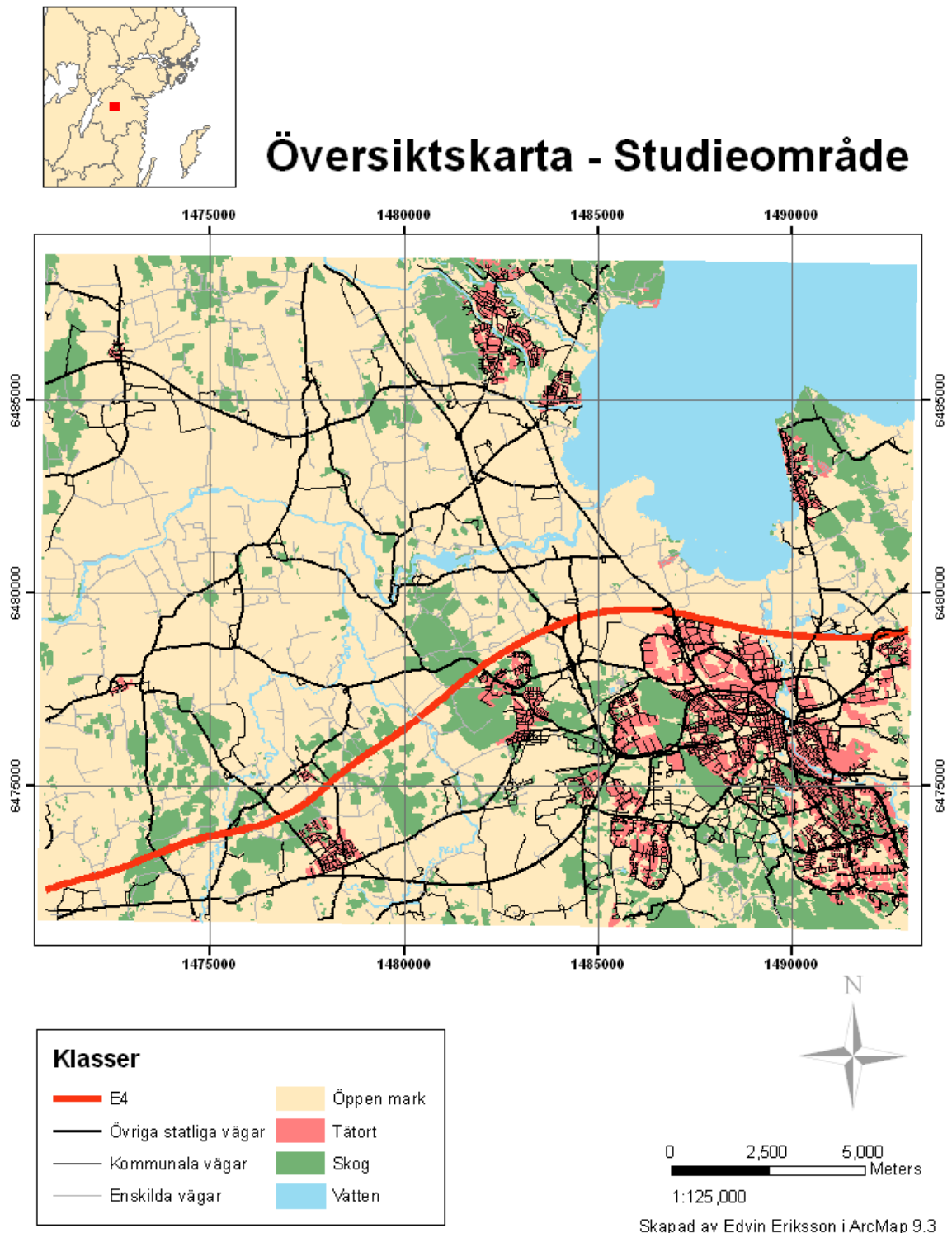
3. Fallstudie: Kvalitetsanalys av prognoskikt för vägolyckor

Fallstudien består av en kvalitetsanalys av prognoskikten för fordonsolyckor, med fokus på att skatta osäkerheten för indata samt hur den osäkerheten fortplantas i prognosmodellen i varje ruta. Även utvärdering mot olycksdata (polisrapporterade olyckor från STRADA) studeras. Både hur prognoskiktet svarar mot olycksstatistik inom olika typer av områden samt hur hela prognoskiktet utvärderar mot utfallsdata.

3.1. Områdesbeskrivning

Undersökningen har genomförts i ett rektangulärt testområde i Östergötland med merparten av testområdet i Linköpings kommun, se figur 13.

Området är ca 373 km^2 stort och inkluderar Linköping, Malmslätt, Ekängen, Slaka, Vikingstad, Berg slussar och Ljungsbro i Linköpings kommun samt Klockrike i Motala kommun. Genom området löper E4, med tre stora trafikplatser: Linköping- Östra, Västra och Norra. I området finns det totalt sett 47 statliga korsningar, 2561 kommunala korsningar samt 1445 statliga väglänkar, 5496 kommunala och 2321 enskilda väglänkar. Området har valts i samråd med VTI och bör därmed vara tillräckligt varierat och innehålla tillräckligt många element för att utvärdera prognosmodellen.

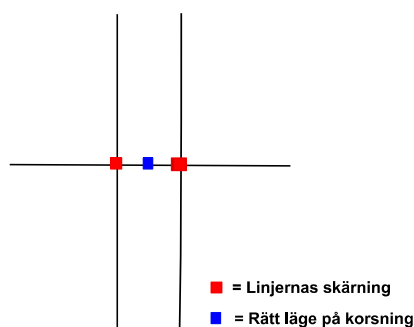


Figur 13: Karta över testområdet som använts i fallstudien. Tematiska data kommer från digitala kartbiblioteket (<https://butiken.metria.se-/digibib/index.php>)

3.2 Data

NVDB har använts som primär datakälla i studien (för förklaring av NVDB se sid 29). I NVDB finns all information som behövs för att beräkna antal olyckor på väglänkar enligt VTI:s modell.

För att beräkna olyckor på korsningar så behövs vektordata med korsningstyp. Att däremot identifiera dessa korsningar utifrån NVDB är svårt. Vid vissa förekommande fall så representeras en väg i NVDB av mer än en linje vilket beror på att vägen har 4 eller fler filer. En linje som korsar dessa två linjer kommer att generera två korsningar där det i verkligheten kanske bara är en (se figur 14 för illustration). Resultatet av det blir att olycksrisk för det området kan bli dubbelt så högt som det "borde" vara. För att undkomma dessa problem, och därmed få ett mer exakt korsningsskikt så digitaliserades samtliga korsningar istället för att använda automatiska lösningar.



Figur 14: Illustration över var linjerna skärs i NVDB i förhållande till rätt läge för korsningen.

De attribut som behövts från NVDB är funktionell vägklass, vägtyp, årsmedeltrafikflöde, hastighetsgräns, väghållare och attribut om vägen ligger i tätort eller ej.

Normalvärden, relativ olycksrisk och korsningsparametrar a , b och c har hämtats från tabell, se bilaga 1. Där värdena är tagna från vägverkets effektkatalog (Vägverket *effektsamband* 2008).

Dessutom så har data om kommunala gränser från SCB (Statistiska CentralByrån) används för att dela in området i olika kommuner.

Vid utvärderingen används polisrapporterade olyckor som erhållits av VTI. Olycksrapporteringen kommer från STRADA (Swedish Traffic Accident Data Acquisition). STRADA är en informationsdatabas som startade 2003 med syfte om att innehålla alla polisrapporterade trafikolyckor som sker i Sverige. Med informationen så lagras uppgifter om olyckstypen samt positionen för olyckan (Sjöo & Ungerbäck 2007). Tanken är att databasen ska hålla en hög lägesnoggrannhet men i nuläget finns det vissa kvalitetsbrister; bland annat är det svårt att knyta olyckorna till en viss position.

3.3 Metod

3.3.1 Metod för beräkning av prognosskiktet

Funktionerna som använts i beräkningen av prognosskiktet kommer från Forsman & Vadeby (2010).

Först så multipliceras trafikflödet per axelpar(ÅDTa) med 365 för att få trafikflödet per år(Q).

$$Q = 365 \times \text{ÅDTa} \quad (8)$$

Därefter beräknas antal olyckor på statliga vägar O_{LS} i varje ruta som följande ekvation där N är normalvärdet och L är längden i km på vägen i rutan.

$$O_{LS} = \frac{N \times Q \times L}{1000000} \quad (9)$$

Olyckor på statliga korsningar O_{KS} beräknades genom ekvation 10, där Q_t är totalt inkommande trafikflöde till korsningen och Q_s är totalt inkommande sekundärt flöde d.v.s. från det vägbenpar som har minst trafikflöde Q , a,b och c är korsningsparametrar.

$$O_{KS} = a \times Q_t^b \times \left(\frac{Q_s}{Q_t}\right)^c \quad (10)$$

Vid beräkning av olyckor på kommunala vägar så beräknades först den relativa olycksrisk på vägens längd(OR), där R är relativ olycksrisk och L är längden på vägen i km i rutan.

$$OR = R \times L \quad (11)$$

Därefter beräknades olyckor på de kommunala vägarna O_{LK} enligt ekvationen nedan, där summeringen av relativ olycksrisk*längd OR sker över alla kommunala länkar i testområdet. O_{obs} är ett snittvärde för observerat antal olyckor per kommun, för Linköping är det 173,15 och för Motala kommun så är det 41,5. Faktorn 0,5 är en justeringsfaktor.

$$O_{LK} = \frac{OR}{\sum OR} \times 0,5 \times O_{obs} \quad (12)$$

Olyckor på kommunala korsningar O_{KK} beräknades på liknande sätt som olyckor på kommunala väglänkar O_{LK} . Relativ olycksrisk R valdes från tabell beroende på vilken korsningstyp och hastighet på vägarna som går till korsningen. Observerat antal olyckor O_{obs} är densamma som vid beräkning av O_{LK} d.v.s 173,15 för Linköping kommun och 41,5 för Motala kommun.

$$O_{KK} = \frac{R}{\sum R} \times 0,5 \times O_{obs} \quad (13)$$

Den sista olyckstypen är olyckor på enskilda vägar O_E , vilken beräknades med ett enkelt samband beroende av vägens längd L i km.

$$O_E = 0,003 \times L \quad (14)$$

Efter att de olika olyckstyperna (O_{LS} , O_{KS} , O_{LK} , O_{KK} , O_E) beräknats så summerades de för varje ruta. Därefter beräknades antalet räddningsinsatser för varje olyckstyp $I_{LS,KS,LK,KK,E}$. Räddningsinsatserna är en korrigering av varje olyckstyp och korrigerar för bortfallet mellan polisrapporterade olyckor och insatser.

$$I_{LS} = 0,76 \times O_{LS}; \quad I_{KS} = 0,76 \times O_{KS} \quad (15)$$

$$I_{LK} = 0,76 \times O_{LK}; \quad I_{KK} = 0,76 \times O_{KK}$$

$$I_E = 0,76 \times O_E$$

Slutligen användes resultaten från insatskorrigeringen $I_{LS,KS,LK,KK,E}$ till att beräkna andelen singel $S_{LS,KS,LK,KK,E}$ och flerfordonsolyckor $F_{LS,KS,LK,KK,E}$ för de olika olyckstyperna

$$S_{LS} = 0,73 \times I_{LS} + 0,03(I_{KS} + I_{LK} + I_{KK} + I_E); \quad F_{LS} = 0,3 \times I_{LS} \quad (16)$$

$$S_{LK} = 0,63 \times I_{LK}; \quad F_{LK} = 0,37 \times I_{LK}$$

$$S_{KS} = 0,7 \times I_{KS}; \quad F_{KS} = 0,3 \times I_{KS}$$

$$S_{KK} = 0,63 \times I_{LK}; \quad F_{KK} = 0,37 \times I_{LK}$$

$$S_E = 0,7 \times I_{LK}; \quad F_E = 0,3 \times I_{LK}$$

Den annorlunda korrigeringen för singelolyckor på statliga vägar beror på att 3 % av alla olyckor är relaterade till kollisioner med vilt. Denna korrigeringen görs enbart för statliga vägar.

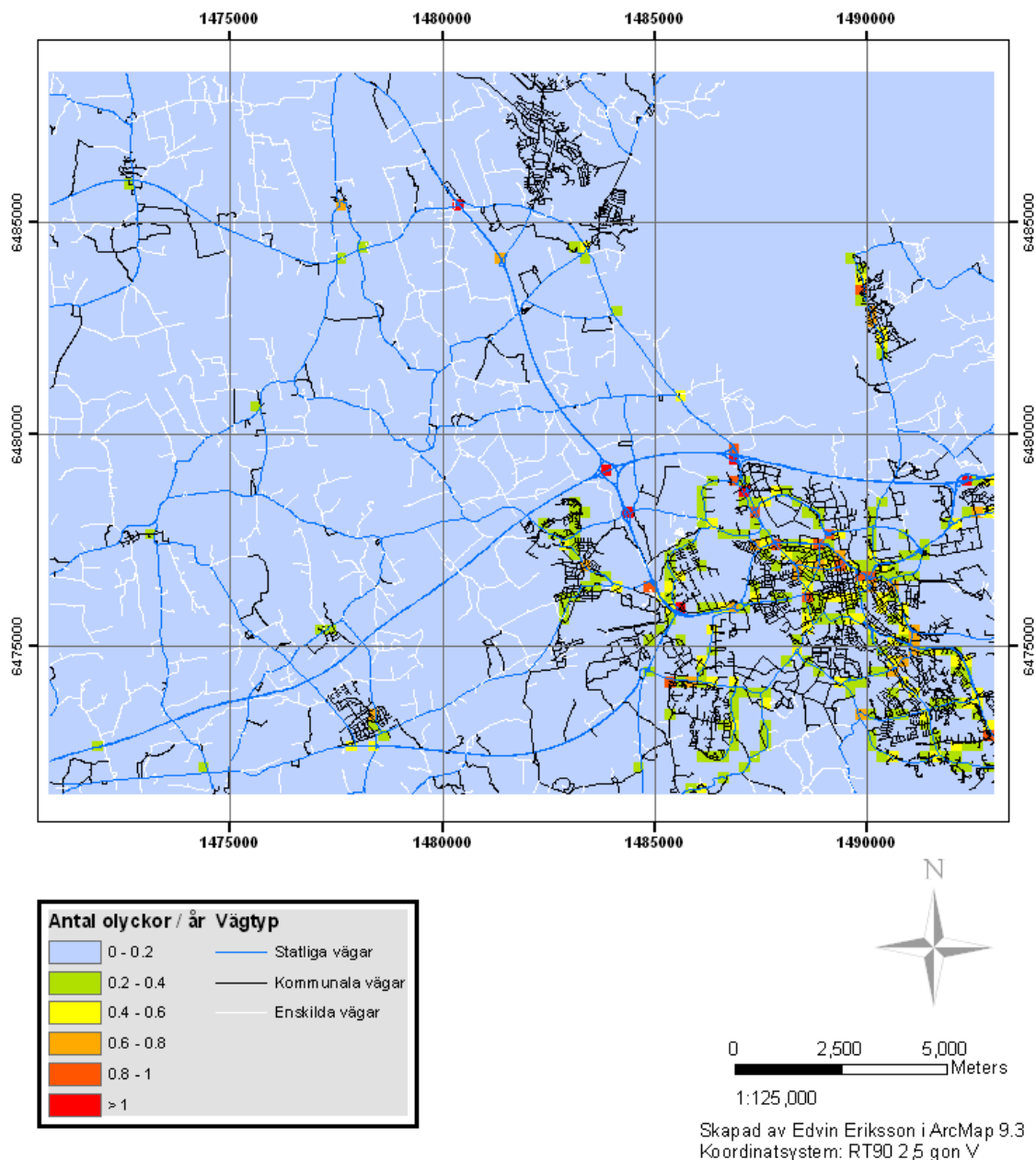
Slutligen summerades samtliga singel och flerfordonsolyckor i varje ruta för att erhålla ett prognoskikt av totalt antal olyckor per år

3.3.2 Implementering av prognoskiktet

ArcGIS 9.3 har använts som huvudsakligt program i implementeringen av prognoskiktet. Först skapades ett linjerutnät med 250*250m stora rutor som konverterades till polygonrutor. Därefter gjordes en uppdelning av vägarna utifrån olyckstyp, kommun och ruta. Uppdelningen utfördes dels med överlagringar mellan vägnät och korsningar med kommunsikt och rutsikt, samt dels med SQL (Structured Query Language) satser baserade

på attribut i NVDB d.v.s. väghållare, hastighetsgräns och tätortsid. Förutom ArcGIS så användes MatLab 2009b som stödprogram för att koppla väglänkarna till rätt normalvärde och relativ olycksrisk (från effektkatalogen, bilaga 1) baserat på attributen i väglänkarna (se bilaga 3a). ArcGIS användes även då samtliga korsningar digitaliserades, varvid all information som behövdes för beräkningen av korsningsolyckorna registrerades parallellt med digitaliseringen. Därefter beräknades varje väglänks längd i varje ruta. Vidare så beräknades de olika olyckstyperna för varje väg och summerades inom varje ruta på polygonrutnätet. Därefter konverterades polygonerna innehållande olika olyckstyper på vägar till raster bestående av de olika olyckstyperna. I samband konverterades punktskikten med olyckor i varje korsning till två rasterskikt. Dessa var olyckor på kommunala och staliga korsningar respektive. Slutligen beräknades prognosskiktet för totalt antal prognostiserade olyckor per år med olika rasterberäkningar baserade på ekvation 15 och 16.

Arbetsgången och verktygen som används i ArcGIS visas i bilaga 2a och 2b Resultatet av implementeringen kan ses i figur 15.



Figur 15: Prognostiserade olyckor från VTI:s modell. Summan av singel och flerfordonsolyckor per år.

3.3.3 Metod för felfortplantningen

3.3.3.1 Dataundersökning – Osäkerhet i data

Syftet med undersökningen av data var att hitta och uppskatta en medelosäkerhet i de olika variablerna som bygger upp flödesmodellen. De föreslagna osäkerheterna baseras på dokumentation i rapporter av NVDB och Vägverket samt muntlig kommunikation med Åsa Forsman på VTI.

Osäkerhet i vägarnas längd (L)

Längden på varje väglänk är direkt relaterad till positionen på varje nod som bygger upp väglänkarna. Längden beräknas med pythagoras sats och i princip så är det enda felet som kan uppstå från en sådan beräkning ett avrundningsfel. Dessa fel bedöms vara alldeles för små för att tas med i osäkerheten för längdangivelsen. Enligt Mattson (2009) finns det metoder inom NVDB:s arbetsgrupp för att studera osäkerheten i hur väl varje väg representerar den verkliga vägens utsträckning. Osäkerheten i denna representation bedöms vara $\pm 2\%$.

Osäkerhet i parametrar från vägverkets effektkatalog (N och R)

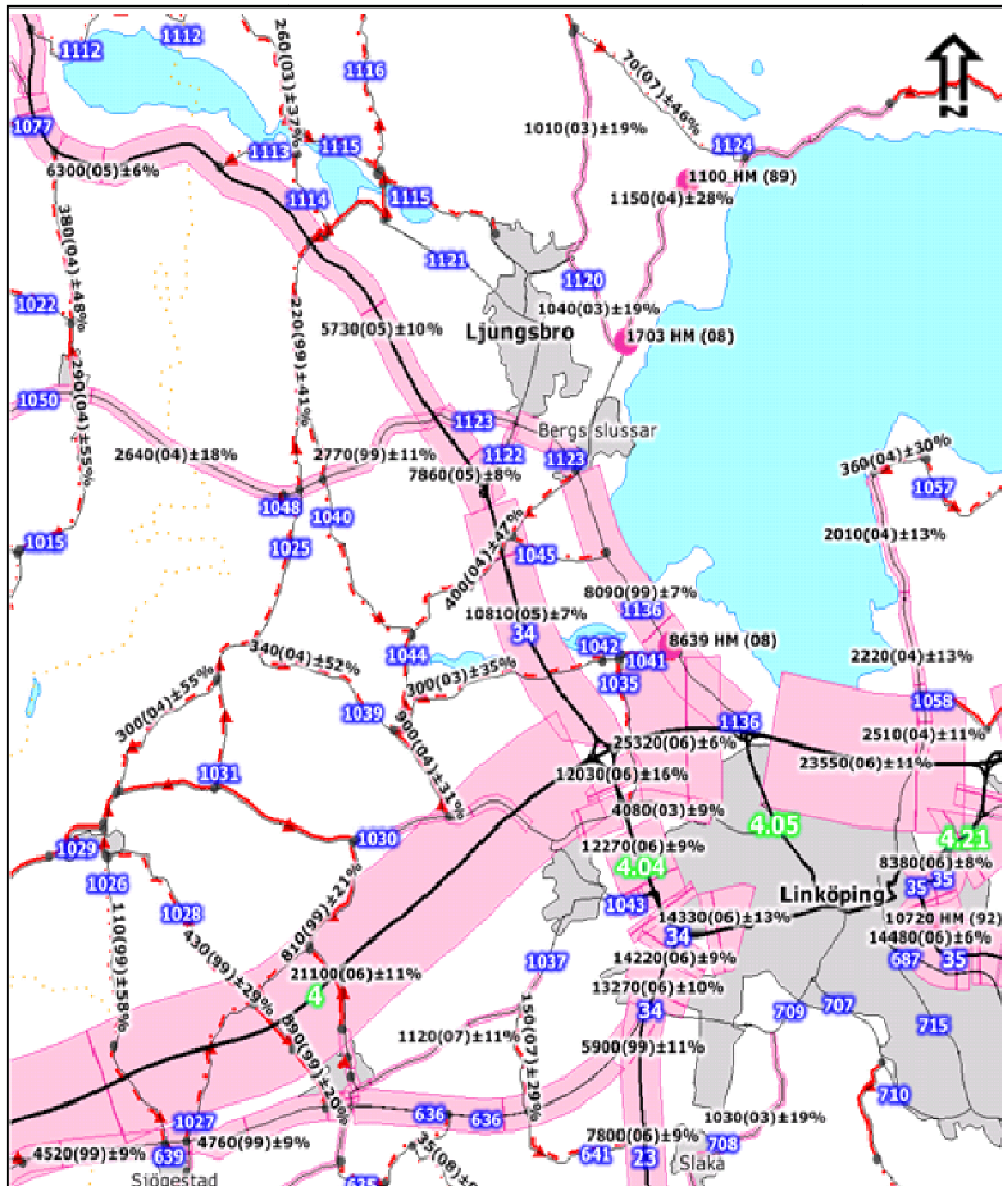
Osäkerheten i normalvärde är ca 10 %. Detta baseras på ett medel av standardavvikelsen utav modellvärdena som beräknar normalvärdet delat med varje normalvärde. Skattningen anses vara tillräckligt representativ för denna studie. Dock är osäkerheten förmodligen lite hög för många av vägarna (Forsman 2010).

Osäkerheten i relativ olycksrisk (*R*) rekommenderades till 15 %. Värdet baseras i stora drag på samma resonemang som vid osäkerheten för normalvärdet. (Forsman 2010)

Osäkerhet i trafikflödet

För osäkerheten i flödet finns det underlag för att göra en osäkerhetsskattning för varje väg. Vägverket beskriver att osäkerheten för flödet är lägre desto större flödet är vilket beror på att det sker fler mätningar på stora vägar med höga trafikflöden än små vägar med ett litet trafikflöde (Vägverket 2010).

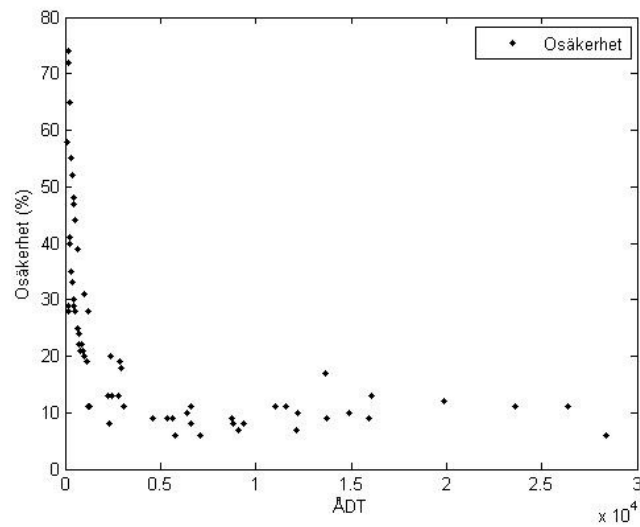
Uppgifter om trafikflöde och trafikmätningar kan nås via en webbtjänst som illustrerar information om trafikflöde på samtliga vägar som har uppgifter om flöde. Informationen inkluderar: storleken på flödet, årtalet för flödesmätningarna samt osäkerheten i flödesuppskattningen för vägen, se figur 16.



Figur 16: Figuren illustrerar testområdet med fokus på fordonsflöde. Vid varje vägsnitt finns ett värde som anger storleken på flödet/dygn, året då mätningarna genomförts (värdet i parantes) och osäkerheten i skattningen för flödet (i procent). Hämtad från vägverket konsult:s kartor över trafikflöden (<http://gis.vv.se/tfk2/tfk/indextfk.aspx?config=tfk>)

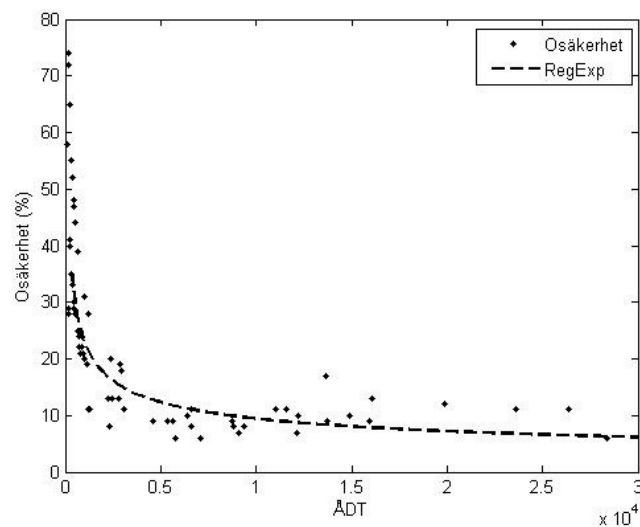
Däremot har det inte funnits tillgång till osäkerhetsvärdena i figur 16 i digital form. Istället har en empirisk modell utnyttjats utifrån sambandet mellan osäkerhet och trafikflöde.

För att få samband (mellan osäkerhet och flöde) så registrerades information om samtliga vägars flöde inom området samt deras respektive osäkerhet. Flödes uppgifterna multiplicerades med en faktor 1,12, enligt beräkningar från Forsman & Vadeby 2010 för att erhålla motsvarande flöde i axelpar som används i beräkning av statliga vägar och korsningar (antagande görs därmed om att osäkerheten i flöde axelpar är samma som osäkerheten för flöde fordon). Därefter plottades osäkerheten (%) mot flödet, se figur 17.



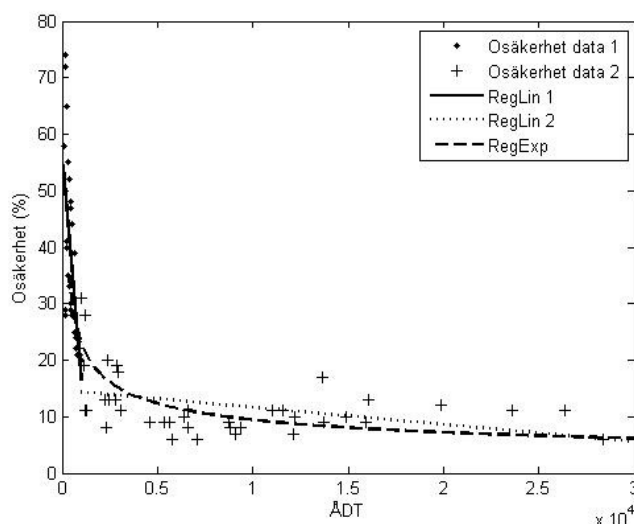
Figur 17: Spridningsdiagram utav osäkerhet mot trafikflöde axelpar (ÅDT).

Från figur 18 kan ett logaritmiskt samband anas mellan osäkerheten och flödet, därmed anpassades en logaritmisk regressionslinje till punktsvärmen, se figur 19



Figur 18: Spridningsdiagram utav osäkerhet mot trafikflöde axelpar (ÅDT) med en exponentiell regressionslinje.

Grafiskt så verkar den logaritmiska regressionslinjen kunna förklara sambandet ganska hyggligt. För enkelhetens skull delades datamängden upp i 2 mängder varvid två linjära regressionslinjer anpassades till datasvärmen, se fig 19.



Figur 19: Spridningsdiagram utav osäkerhet mot trafikflöde axelpar (ÅDT) med en exponentiell regressionslinje och två linjära regressionslinjer som bryter vid ett ÅDT på 1000.

Tröskelvärdet, d.v.s. värdet på flödet som valdes som delningspunkt, sattes till 1000. Värdet valdes utifrån observation av grafen och med hänseende på att hålla förklaringsgraden av de olika regressionslinjerna så hög som möjligt.

De linjära regressionsmodellerna nedan används i felfortplantningen som osäkerhetsuppskattningen för trafikflöde axelpar $\hat{A}DTa$.

$$Osäkerhet_{\hat{A}DT < 1000} = 56,1018 - 0,0396 \times \hat{A}DT \quad (17)$$

$$Osäkerhet_{\hat{A}DT > 1000} = 14,6471 - 3,0237 \times 10^{-4} \times \hat{A}DT$$

3.3.3.2 Analytisk felfortplantning

Metoden för att skatta hur osäkerheten i indata fortplantas i modellen är i denna studie analytisk felfortplantning med första ordningens Taylorutveckling, se ekvation 4. Detta avsnitt visar de beräkningar som används för att studera hur osäkerheten fortplantas i de olika stegen i prognosmodellen, där osäkerheten för indata baseras på bedömningar i föregående avsnitt 3.3.3.1.

Felfortplantningen för statliga vägar

Utifrån funktionen för olyckor på statliga vägar O_{LS} (ekvation 9), deriveras uttrycket med avseende på normalvärdet N , Längden L och Flödet Q .

$$\frac{dN}{dO_{LS}} = \frac{Q \times L}{1000000}; \quad \frac{dQ}{dO_{LS}} = \frac{N \times L}{1000000}; \quad \frac{dL}{dO_{LS}} = \frac{N \times Q}{1000000} \quad (18)$$

Därefter beräknades osäkerheten i staliga olyckor σ_{OLS} enligt följande ekvation där σ_N är osäkerheten i normalvärdet, σ_Q är osäkerheten i trafikflödet och σ_L är osäkerheten i väglänkens längd.

$$\sigma_{OLS} = \sqrt{\left(\frac{Q \times L}{1000000}\right)^2 \times \sigma_N^2 + \left(\frac{N \times L}{1000000}\right)^2 \times \sigma_Q^2 + \left(\frac{N \times Q}{1000000}\right)^2 \times \sigma_L^2} \quad (19)$$

Felfortplantningen för kommunala vägar

Först beräknas derivatan av funktionen för olyckor på kommunala väglänkar (ekvation 12) med avseende på OR (längden på länken L multiplicerat med relativ olycksrisk R för längden)

$$\frac{dOR}{dOLK} = \frac{0,5 \times O_{obs}}{\Sigma OR} \quad (20)$$

Osäkerheten för olyckorna på de kommunala väglänkarna σ_{OLK} baserat på osäkerheten i OR beräknas sedan genom följande ekvation:

$$\sigma_{OLK} = \sqrt{\left(\frac{0,5 \times O_{obs}}{\Sigma OR}\right)^2 \times \sigma_{OR}^2} \quad (21)$$

Där osäkerheten i OR σ_{OR} beräknas genom felfortplantning, σ_R är osäkerheten i R och σ_L är osäkerheten i L .

$$\sigma_{OR} = \sqrt{(L)^2 \times \sigma_R^2 + (R)^2 \times \sigma_L^2} \quad (22)$$

Felfortplantningen för enskilda vägar (Oe)

Första derivatan av funktionen för enskilda vägar (ekvation 14), med avseende på längden är 0,003 vilket ger följande uttryck för osäkerheten på de enskilda vägarna σ_{Oe} baserat på osäkerheten i längden.

$$\sigma_{Oe} = \sqrt{0,003^2 \times \sigma_L^2} \quad (23)$$

Felfortplantningen för kommunala korsningar

Derivering av kommunala korsningsfunktionen för olyckor (ekvation 14) med avseende på relativ olycksrisk R ger uttryck i ekvation 24.

$$\frac{dR}{dO_{KK}} = \frac{0,5 \times O_{obs}}{\Sigma R} \quad (24)$$

Osäkerheten för olyckor på kommunala korsningar baserat på osäkerhet i relativ olycksrisk beräknades genom följande uttryck.

$$\sigma_{O_{KK}} = \sqrt{\left(\frac{0,5 \times O_{obs}}{\Sigma R}\right)^2 \times \sigma_R^2} \quad (25)$$

Felfortplantningen för statliga korsningar (Oks)

Derivatan av funktionen för olyckor på statliga korsningar (ekvation 10) med avseende på totalt inkommande flöde Q_t ger uttryck i ekvation 26. Där a , b och c är korsningsparametrar och Q_s är inkommande flöde på det sekundära vägbenparet (se figur 12 för beskrivning).

$$\frac{dQ_t}{dO_{KS}} = Q_s^c \times a(b-c)Q_t^{b-c-1} \quad (26)$$

Då samma funktion (ekvation 10) deriveras med avseende på Q_s erhålles annat uttryck.

$$\frac{dQ_s}{dO_{KS}} = a \times c \times Q_t^{b-c} \times Q_s^{c-1} \quad (27)$$

Osäkerheten i de statliga korsningarna $\sigma_{O_{KS}}$ baserat på osäkerhet i Q_t σ_{Q_t} och osäkerheten i Q_s σ_{Q_s} beräknas sedan enligt:

$$\sigma_{O_{KS}} = \sqrt{(Q_s^c \times a(b-c)Q_t^{b-c-1})^2 \times \sigma_{Q_t}^2 + (a \times c \times Q_t^{b-c} \times Q_s^{c-1})^2 \times \sigma_{Q_s}^2} \quad (28)$$

Felfortplantning av justeringsekvationer och slutlig osäkerhet

Efter att osäkerheten för de olika olyckstyperna $\sigma_{O_{LS};O_{LK};O_{KS};O_{LS};O_{LS};O_{LS}}$ räknats ut så beräknas propageringen av indatas osäkerhet i insatskorrigering $\sigma_{I_{LS};I_{LK};I_{KS};I_{LS};I_{LS};I_{LS}}$ baserat på ekvation 15. Derivatan av insatskorrigering blir 0,76 för samtliga olyckstyper varav felfortplantningen med hänsyn till detta blir följande:

$$\begin{aligned} \sigma_{I_{LS}} &= \sqrt{0,76^2 \times \sigma_{O_{LS}}^2} & \sigma_{I_{LK}} &= \sqrt{0,76^2 \times \sigma_{O_{LK}}^2} & (29) \\ \sigma_{I_{KK}} &= \sqrt{0,76^2 \times \sigma_{O_{KK}}^2} & \sigma_{I_{KS}} &= \sqrt{0,76^2 \times \sigma_{O_{KS}}^2} & \sigma_{I_E} &= \sqrt{0,76^2 \times \sigma_{O_E}^2} \end{aligned}$$

Osäkerheten i singel σ_{SLS} och flerfordonsolyckor σ_{FLS} för de statliga vägarna beräknas genom ekvation 30 och 31 där konstanterna är derivatan av funktionen i ekvation 16 med avseende på insatserna $I_{LS}, I_{LK}, I_{KS}, I_{LS}, I_{LS}, I_{LS}$.

$$\sigma_{SLS} = \sqrt{0,73^2 \times \sigma_{I_{LS}}^2 + 0,03^2 \times \sigma_{I_{KS}}^2 + 0,03^2 \times \sigma_{I_{LK}}^2 + 0,03^2 \times \sigma_{I_{KK}}^2 + 0,03^2 \times \sigma_{I_E}^2}$$

$$\sigma_{FLS} = \sqrt{0,3^2 \times \sigma_{I_{LS}}^2} \quad (30)$$

Utifrån samma metod så beräknades osäkerheten i resterande singel och flerfordonsolyckor: för kommunala väglänkar $\sigma_{SLK}, \sigma_{FLK}$, enskilda väglänkar σ_{SE}, σ_{FE} , statliga korsningar $\sigma_{SKS}, \sigma_{FKS}$ samt för kommunala korsningar $\sigma_{SKK}, \sigma_{FKK}$.

$$\sigma_{SLK} = \sqrt{0,63^2 \times \sigma_{I_{LK}}^2} \quad \sigma_{FLK} = \sqrt{0,37^2 \times \sigma_{I_{LK}}^2} \quad (31)$$

$$\sigma_{SE} = \sqrt{0,7^2 \times \sigma_{I_E}^2} \quad \sigma_{FE} = \sqrt{0,3^2 \times \sigma_{I_E}^2}$$

$$\sigma_{SKS} = \sqrt{0,7^2 \times \sigma_{I_{KS}}^2} \quad \sigma_{FKS} = \sqrt{0,3^2 \times \sigma_{I_{KS}}^2}$$

$$\sigma_{SKK} = \sqrt{0,63^2 \times \sigma_{I_{KK}}^2} \quad \sigma_{FKK} = \sqrt{0,37^2 \times \sigma_{I_{KK}}^2}$$

Den sammanvägda totala osäkerheten σ_{TAO} i varje ruta, baserat på felpropageringen av osäkerheten i indata, beräknas slutligen genom följande uttryck

$$\sigma_{TAO} = \sqrt{\sigma_{SLS}^2 + \sigma_{FLS}^2 + \sigma_{SLK}^2 + \sigma_{FLK}^2 + \sigma_{SE}^2 + \sigma_{FE}^2 + \sigma_{SKS}^2 + \sigma_{FKS}^2 + \sigma_{SKK}^2 + \sigma_{FKK}^2} \quad (32)$$

3.3.4. Implementering av felfortplantningen

Felfortplantningen implementerades i MatLab, se bilaga 3 b. Resultatet i varje ruta adderades till ett vektor skikt (shape fil) bestående av 250*250m polygoner täckande området. Därefter konverterades skiktet (shape filen) till ett raster med samma utbredning i ArcGIS. Detta gjordes för att kunna dividera med prognosskiktet som beräknats tidigare (se bilaga 2 a "tao") och därmed få den relativa påverkan av osäkerheten i indata för varje ruta. Båda resultaten illustreras i figur 22 och figur 23.

3.3.5 Statistisk utvärdering mot utfall

3.3.5.1 Utvärdering områden

Av intresse är att se hur väl prognoskiktet svarar mot antal observerade olyckor inom större områden. Anledningen till en områdesvis jämförelse är att minimera slumpen utav olycksfrekvens av utvärderingsdata som kan ske p.g.a. bristande kvalitet. För att genomföra utvärderingen utnyttjas vissa statistiska resonemang.

Antal olyckor över tiden kan förklaras med Poissionfördelning (Blom et al. 2005) . Därmed kan det anses att prognoser för fordonsolyckor i varje ruta kan förklaras med intensiteten i en poissonfördelning över trafikolyckor. Fördelen med Poission fördelning är att den har vissa användbara egenskaper:

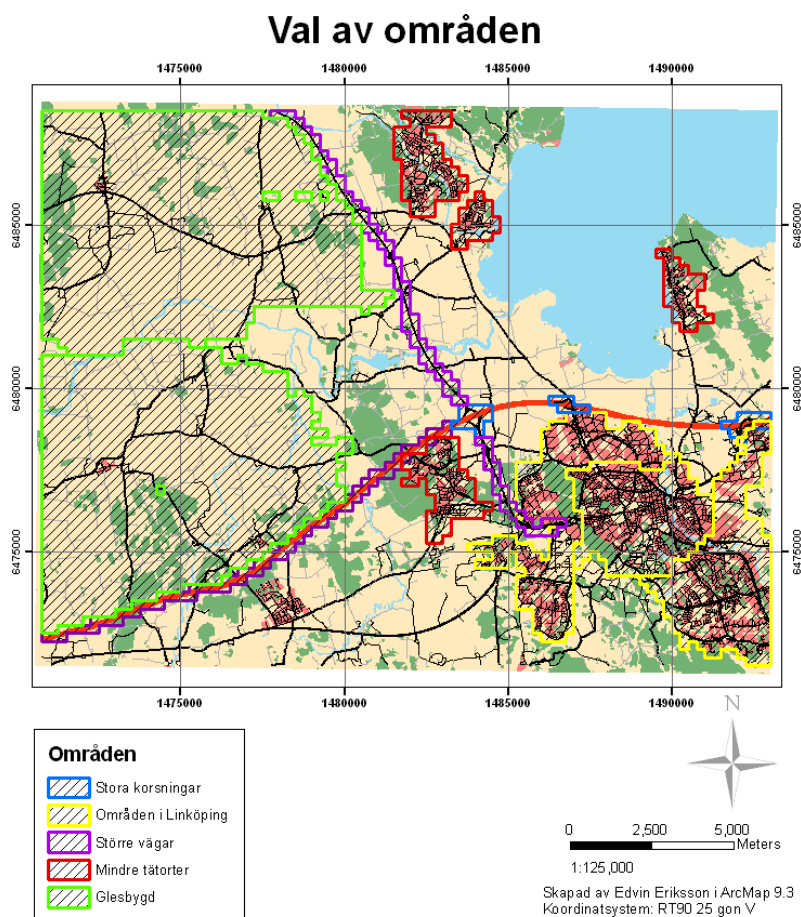
1. Sats 7.8 i Blom et al 2005 gäller vid oberoende av celler: Nämligen om X är Poissionfördelat med väntvärde $Po(\mu_1)$ och Y är Poissionfördelat med väntvärde $Po(\mu_2)$ då är summan $X + Y$ Poissionfördelat med väntvärde $Po(\mu_1 + \mu_2)$. Detta innebär att det går bra att aggregera olika μ , både i tiden och i rummet. I detta fallet innebär det aggregering av olika rutor i prognoskiktet.
2. Vidare gäller sats 7.7 att summan för väntvärdet μ är likställt med variansen vilket innebär att standardavvikelsen är roten av väntvärdet d.v.s. $\sqrt{\mu}$.
3. Poissionfördelningen har även en approximativ egenskap som innebär att vid stora värden på väntvärdet, ungefär $\mu > 15$, så kan normalapproximation göras d.v.s att X blir normalfördelat $X \in N(\mu, \sqrt{\mu})$. En fördel med att göra normalapproximation är att konfidensintervall kan skattas.

Ovantaende ansatser utnyttjas för att jämföra hur olika områdens prognostiserade olyckssvärden svarade mot faktiska olyckor inom tidsintervallet 2003-2008.

Val av områden

Intressanta områden att utvärdera hade varit statliga vägar, kommunala vägar, korsningar i tätorter, statliga korsningar och enskilda vägar utanför tätorter. Tyvärr kan inte en distinktion göras med urval på vägar och korsningar. Istället så utvärderades olika tätorter, områden i Linköping, områden täckande större statliga vägar, områden täckande de stora trafikplatserna samt glesbygdsområden, se figur 21 för illustrering av vilka områden som valts.

Gemensamt för områdena är att summan av den prognostiserade olycksfrekvensen under åren 2003-2008 överstiger 15 olyckor.



Figur 20: Karta över testområdet som visar val av intresseområden till utvärderingen. Gemensamt för testområdena är att $P_o(\mu)$ är större än 15 för tidsperioden 2003 -2008, d.v.s. att de aggregerade prognoserna för antal olyckor överstiger 15 under denna period.

Bilda konfidensintervall

Konfidensintervall är en skattning av osäkerheten av en företeelse och kan användas för undersöka om det "sanna" värdet hamnar inom det skattade värdets osäkerhetsintervall med en viss säkerhet(konfidens) (Olsson et al. 2005). Vanliga konfidensintervall som används är 95%, vilket då innebär att skattningen motsvarar det "sanna" värdet om värdet ligger inom 95% av skattningens fördelning av möjliga värden. Då det skattade värdet kan förklaras med en Poissonfördelning så går det att skapa ett approximativt 95% konfidensintervall enligt följande:(Blom et al.2005) :

$$I_{\mu} = \mu \pm 1,96 \times \sqrt{\mu} \quad (33)$$

Där I_{μ} är konfidensintervallet, μ är väntvärdet, $\sqrt{\mu}$ är standardavvikelsen och 1,96 är den standardiserade normalfördelningens värde vid gränsen för 95 % av fördelningens värde.

Det "sanna" värdet i denna tillämpning är de faktiska olyckorna som har skett i varje område. Det som ska studeras är om de olika områdenas 95 % konfidensintervall för prognostiserade olyckor täcker in STRADA olyckorna som har skett i området.

Praktisk utförande

Först konverterades prognoskiktet från raster till vektor format i ArcGIS. Därefter valdes områdena ut manuellt. Därefter extraherades olyckorna från STRADA inom varje område med överlagrings verktyg. Detta resulterade i tabeller innehållande prognosvärdena för tidsperioden samt STRADA olyckorna för tidsperioden. Dessa tabeller importerades därefter till MatLab med ett script som hanterar .dbf filer. I MatLab beräknades standardavvikelsen $\sqrt{\mu}$ för de olika områdenas prognosvärde och 95 % konfidensintervall. Slutligen plottades de olika områdenas prognostiserade olycksvärden med respektive konfidensintervall mot STRADA olyckorna, se figur 24.

3.3.5.2 Utvärdering hela prognoskiktet

När Jaldell (2009) utvärderade prognoskikten för bränder så undersökte han hur varje ruta i prognoskiktet svarade mot faktiska olyckor i samma ruta. Därmed är det även intressant med en utvärdering hur varje ruta i prognoskiktet svarar mot antal STRADA olyckor i varje ruta. Dels för att testa kvaliteten utav hela prognoskiktet och dels för att det inte bestämt vilken upplösning som prognoskikten ska ha i ROPA.

Valideringen utav prognoskiktet mot utfallsdata utfördes både med 250 och 1000 m stora rutor på prognoskiktet. Anledningen till att prognoskikt med upplösning på 1000 m också utvärderades var att Jaldell (2009) använt 1000 m stora rutor i prognoskiktet för bränder i bostäder.

Utvärderingsmått som använts är *MDA*, *RMSE*, och *r* se ekvation 7, 8 och 9.

Praktisk utförande

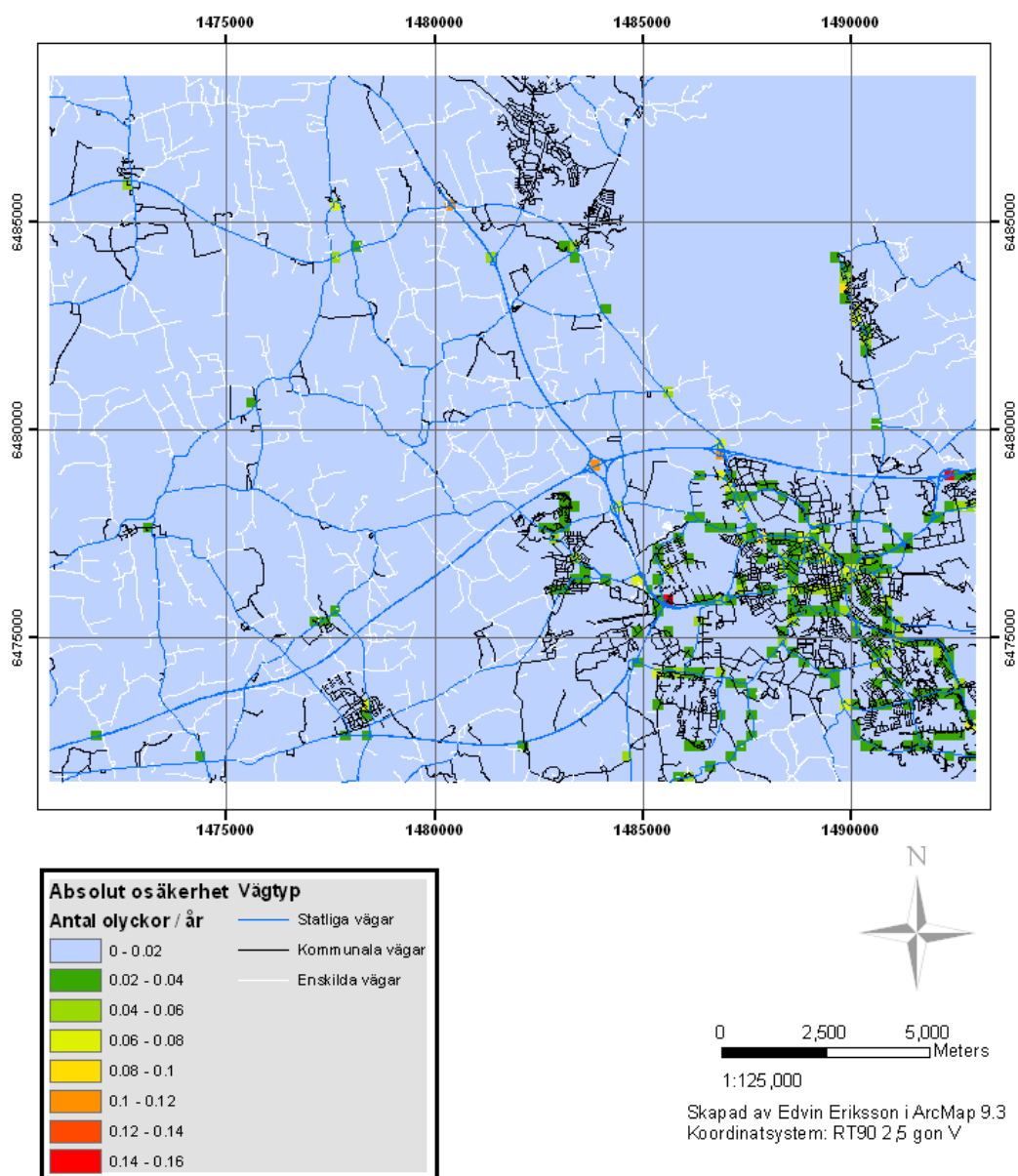
Först kopplades tabellen om polisrapporterade olyckor till ett rutnät innehållande polygoner med 250m rutor täckande området. Sedan konverterades polygonskiktet till raster med avseende på antal olyckor i varje polygon. Därefter ersattes "NonData" värden med 0 och värdet i varje multiplicerades 5 för att erhålla antal olyckor per år.

Prognoskiktet och STRADA skiktet aggregerades sedan med en faktor 4 för att få 2 skikt med 1000 m rutor. Därefter exporterades de 4 skikten, två prognoskikt och två utfallsskikt, till ASCII filer för enklare hantering i MatLab. I MatLab beräknades slutligen utvärderingsmått RMSE och MDA och *r*, figur 25 och 26 visar plottar mellan prognosticerade olycksvärden mot STRADA olyckor. Tabell 2 visar resultatet från utvärderingen.

3.4.Resultat

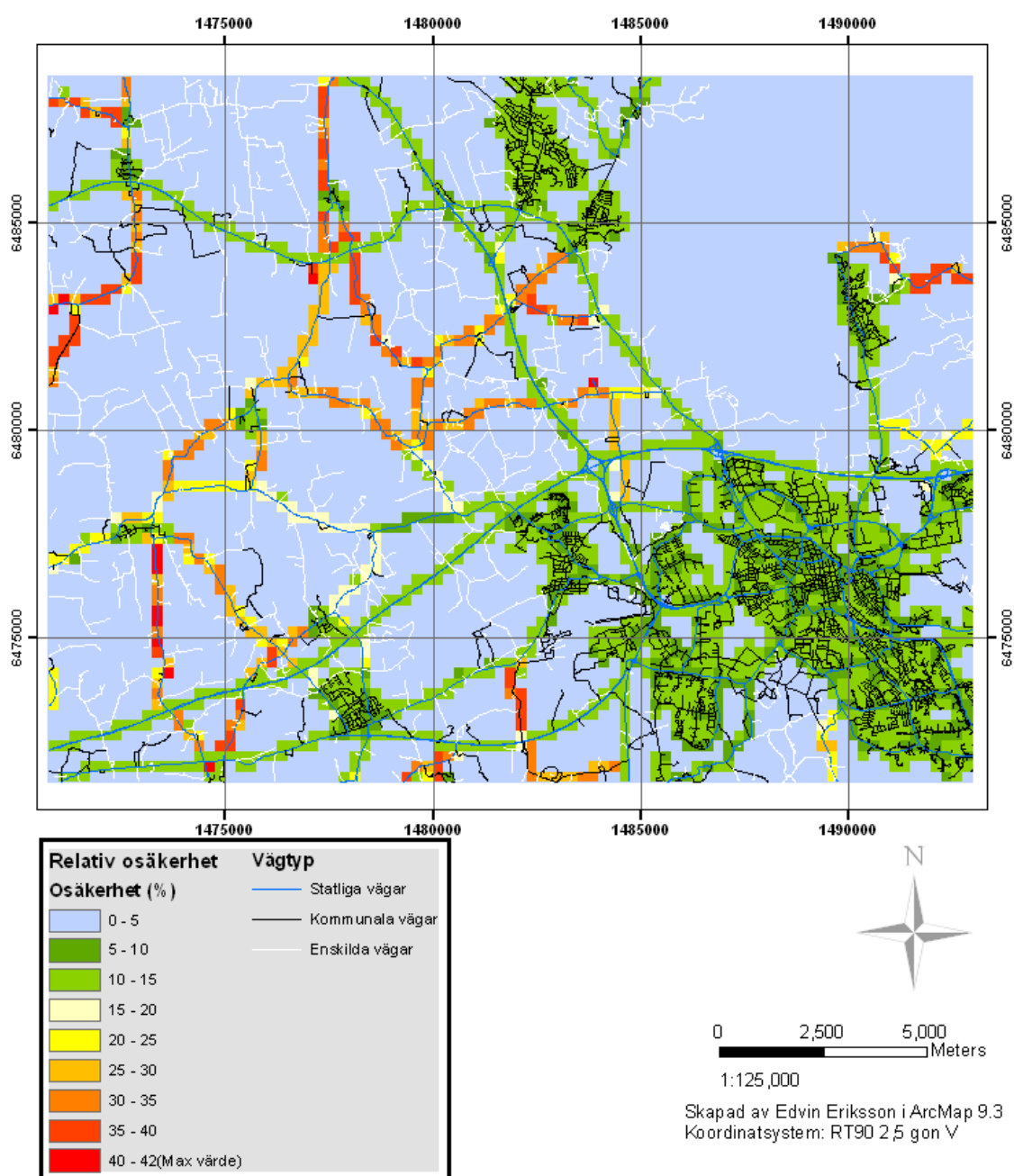
3.4.1 Resultat från felfortplantningsanalysen

De två följande figurerna illustrerar resultatet från felfortplantningen, i absoluta och relativa mått. Figur 21 visar felfortplantningens totala storlek för varje ruta i området (antal olyckor per år). Det som kan utläsas från figuren är att osäkerheten i indata har störst påverkan på rutorna i Linköpings tätort samt i stora korsningar i området. Osäkerhetens fortplantning på det enskilda vägnätet är litet i sammanhanget och är även väntat då den enda osäkerhetsfaktorn i indata på det enskilda vägnätet är 2 % av vägens längd i rutan.



Figur 21: Resultatet från analytisk felfortplantning med första ordningens Taylor utveckling.

Då resultatet från felfortplantningen dividerats med olycksvärdet från prognosskiktet, i varje ruta, så framträder en helt annan fördelning än figur 21. Figur 22 illustrerar den relativa påverkan av osäkerheten i indata för varje ruta d.v.s. hur stor felmarginal som kan väntas jämfört med prognosvärdet baserat på osäkerhet i indata. Det som utmärker är den höga relativa påverkan på mindre statliga vägar. Samtliga tätorter i området påverkas liknande av osäkerheten i indata. Likt figur 21 så har det enskilda vägnätet en låg osäkerhet även i relativa mått.

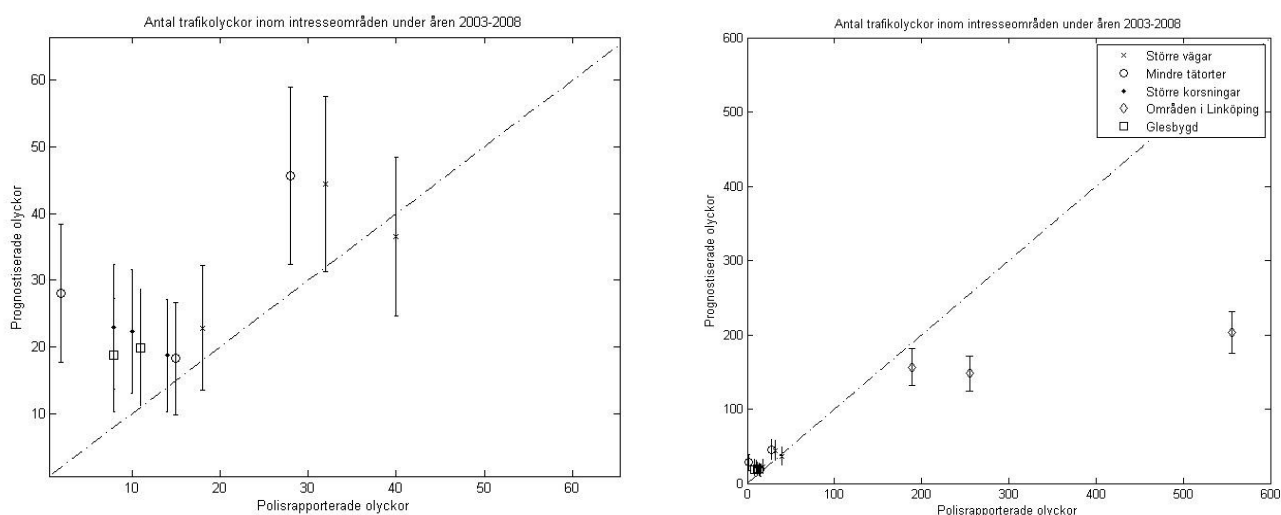


Figur 22: Relativ osäkerhet i området, baserat på analytisk felfortplantning med första ordningens Taylor utveckling.

3.4.2 Resultat från utvärdering mot utfallsdata

3.4.2.1 Utvärdering områden

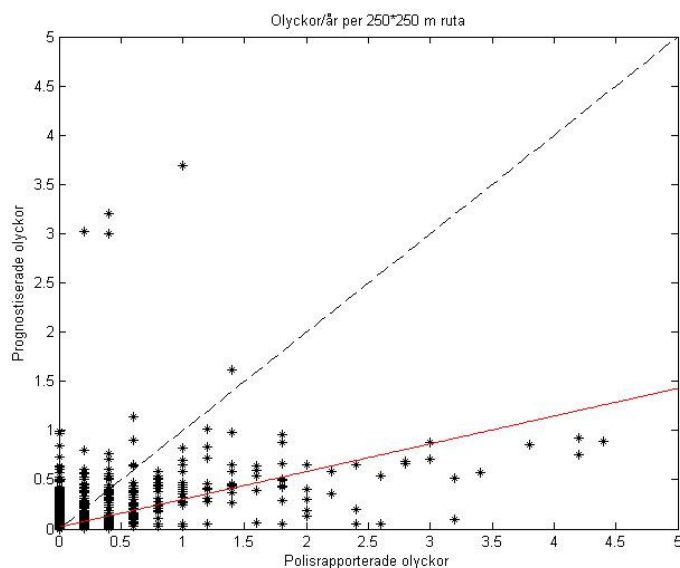
Figur 23 illustrerar resultatet från utvärderingen i särskilda intresseområden mot STRADA olyckorna under tidsperioden 2003-2008. På y axeln är det prognostiserade olycksvärdet och på x-axeln STRADA olyckorna. Områdena betecknas med olika symboler och linjen som går ut ifrån varje jämförelse är ett 95 % konfidensintervall baserat på ekvation 33. Nedre grafen är en förstoring av intervallet 0 – 60 olyckor av den övre grafen. Den streckade linjen motsvarar en perfekt överensstämmelse mellan modellen och utfallsdata. Det som kan utläsas är att samtliga konfidensintervallen för de större vägarna överlappar linjen. Ingen av konfidensintervallen områdena i Linköping överlappar. För de andra tre områdesklasserna så överlappar en av områdenas konfidensintervall den streckade linjen.



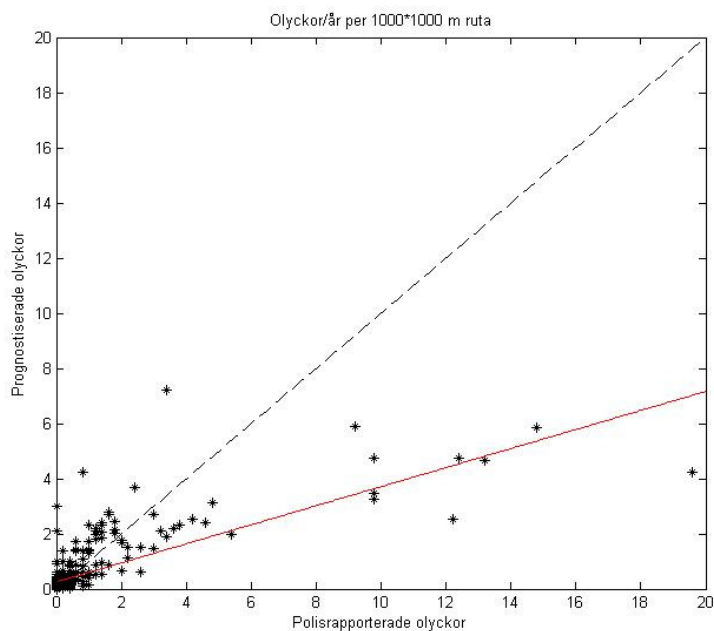
Figur 23: Grafen till höger visar de prognostiserade olycksvärdena under tidsperioden 2003-2008 inom olika områden. Figuren till vänster är en förstoring av intervallet 0-60 antal olyckor. Linjerna betecknar 95 % konfidensintervall.

3.4.2.2 Utvärdering hela prognoskiktet

Följande två figurer 24 och 25 visar resultatet från utvärderingen utav hela prognoskiktet mot STRADA för varje ruta, i upplösningen 250 m och 1000 m respektive. Generellt sett så är antal faktiska olyckor fler än de prognostiserade olyckorna för varje ruta. Då rutorna aggregerats till 1000 m rutor (fig 22) så finns det rutor med betydligt fler polisrapporterade olyckor än de prognostiserade.



Figur 24: Grafen visar olyckor från prognosmodellen med 250 m upplösning, plottade mot STRADA olyckorna i samma ruta. Den röda linjen är regressionslinjen.



Figur 25: Grafen visar olyckor från prognosmodellen med 1000 m upplösning, plottade mot STRADA olyckorna i samma ruta. Den röda linjen är regressionslinjen.

Resultatet från utvärderingsmetoderna visas i tabell 2. Där kan det utläsas att 250 m prognosskiktet, i jämförelse med utfallsdata, har ett lägre värde än 1000 m prognosskiktet för RMSE och MDA och r . (se ekvation 5,6 och 7 för förklaring av utvärderingsmått.)

Tabell 2. Resultatet från utvärderingen av hela prognosskiktet mot olycksstatistik från STRADA, både med 250 och 1000 m upplösning på rutor.

Upplösning	r	MDA	RMSE
250 m	0,415	0,0427	0,2015
1000 m	0,5009	0,4698	1,6621

4 Diskussion

Diskussionsdelen inleds med diskussion om resultaten av de olika momenten i fallstudien för att avslutas med förslag på fortsatta studier samt egna reflektioner. Fokus i rapporten har varit felfortplantningsanalysen och därmed behandlas den också först och utförligast.

4.1 Diskussion om felfortplantningen

4.1.1 Resultaten

Som den första kartan visar, figur 21, så är den absoluta osäkerheten störst vid de statliga korsningarna och tätorterna. För statliga korsningar så beror det delvis på att vid de statliga korsningarna så finns det både korsningar och vägar inom samma ruta vilket bidrar till att öka osäkerheten. Dessutom har de stora statliga korsningarna höga värden på trafikflöden till korsningarna vilket gör att osäkerheten med avseende på flödet blir högt (modellen för olyckor på de statliga korsningarna bygger i princip enbart på inkommande flöde). Anledningen till de höga absoluta osäkerhetsvärdena inom Linköping är att vägarna och korsningarna ligger tätt vilket ökar chansen att fler vägar och korsningar hamnar inom samma ruta och därmed fler bidragande faktorer till osäkerhet.

Den andra kartan som visar den relativa osäkerheten, figur 22 är förmodligen mer intressant än den absoluta. Här får man en uppfattning om den procentuella osäkerheten av varje ruta i förhållande till sitt medelvärde. Det spelar inte så stor roll vid de stora statliga korsningarna att den absoluta osäkerheten är hög eftersom de redan har ett högt olycksvärde. Mindre statliga vägar har den största relativa osäkerheten, vilket beror på att deras totala olycksrisk inte är så hög. Osäkerheten i dessa uppgår till 42 %. Detta värdet är förmodligen ganska högt och bör användas med försiktighet.

De flesta funktioner som bygger upp modellen är linjära varvid en första gradens Taylor ger ett trovärdigt resultat för osäkerheten i prognoskiktet baserat på osäkerheten i indata. Givetvis med utgångspunkt på att osäkerheten i de ingående faktorerna är rätt valda. Linjäriteten i funktionerna återspeglas också på resultatet där den relativa osäkerheten hamnar i liknande storleksordning som osäkerheten i data. Detta indikerar också att prognosmodellen är känslig för fel i indata.

4.1.2 Övrig diskussion

4.1.2.1 Ytterligare osäkerhet

Vid statliga korsningar föreligger förmodligen en större osäkerhet än vad kartan visar eftersom ingen hänsyn tagits till osäkerheten i korsningsparametrarna a , b och c . Dessutom är funktionerna för olyckor på statliga korsningar konkavt avtagande baserat på inkommande trafikflöde (både totalt och sekundärt), vilket innebär att osäkerheten framräknad med första

ordningens Taylorutvecklingen kan skilja sig något från den "faktiska" osäkerheten, se figur 6 för illustration.

4.1.2.2 Osäkerhet i indata kontra osäkerhet i modell

Denna studie tar bara hänsyn till hur osäkerheten i indata fortplantas i modellerna och inte hur osäker modellen är i sig. I båda kartorna kan det skådas att osäkerheten är lägst för det enskilda vägnätet O_E . Den enda orsaken till att osäkerheten är låg för O_E är att modellen är väldigt förenklad. I realiteten är förmodligen olycksfrekvensen på enskilda vägar beroende av så många andra faktorer än vägens längd. Därmed kanske osäkerheten egentligen är störst för det enskilda vägnätet eftersom modellen inte tar hänsyn till någon av dessa potentiella faktorer. Så även om de mindre statliga vägarna har större osäkerhet som ett resultat av osäker indata så har förmodligen det kommunala och enskilda vägnätet en större osäkerhet i olycksrisk som resultat av en osäker modell.

Vid skattning av total osäkerhet är det viktigt att ta hänsyn till fel i modellen. Heuvelink (1998) anser att om hänsyn inte tas till hur väl modellerna representerar det geografiska fenomenet så blir osäkerhetsbedömningen undervärderad. Problemet i detta fallet är att det är svårt att studera hur väl varje olycksfunktion representerar verkligheten eftersom det finns kvalitetsbrister i lägesnoggrannheten i STRADA (Sjöö & Ungerbäck 2007). Detta gör att det är svårt att knyta olyckorna till respektive väg eller korsning. Om tillgång funnits till ett större och mer detaljerat dataset över olyckor så hade detaljstudier om varje modells osäkerhet också varit möjlig att genomföra.

4.1.2.3 Eventuella problem med metoden

Att beräkna första gradens Taylor på diskreta värden bör undvikas (Heuvelink 1998) men eftersom R och N inte kan betraktas som diskreta variabler utan istället som medelvärden utav flera beräkningar är det inget problem. Praktiskt sett så anpassas ett linjärt beroende mellan O_{KK}/O_{KL} och R samt mellan O_{LS} och N .

Det kan tyckas att det finns ett problem vid felfortplantningen av de kommunala länkarna och korsningarna. Summan av R ($\sum R$) och OR ($\sum OR$) (d.v.s. den relativa olycksrisken och längden på länken * relativ olycksrisk) har använts som konstanter i beräkningen. Värdet på R och OR som undersöks är ju en del av summan vilket innebär att summan också rimligtvis bör ha någon form av osäkerhet. För större områden minskar detta problem vilket gör att $\sum R$ och $\sum OR$ kan behandlas som konstanter.

4.1.2.4 Effekter av data behandling

Konvertering mellan olika format och data behandling brukar generellt bidra till osäkerheten i resultatet. I analysen har flera konverteringar mellan vektor och raster samt vice versa utförts, (se bilaga 2 a). Den potentiella osäkerheten som bidras från dessa konverteringar anses vara så små och därmed har ingen hänsyn tagits till detta. Förklaringen är eftersom

konverteringarna hela tiden utförts mellan raster och polygonrutor med samma upplösning och samma geometriska form.

4.1.3 Diskussion om osäkerhetsval hos parametrar

4.1.3.1 Normalvärde

Osäkerheten för normalvärdet skattades som ett enkelt samband d.v.s. standardavvikelsen / normalvärdet. Forsman (2010) tror att detta värde är högt för flera av vägarna. Detta är säkert fallet men det påverkar förmodligen inte den höga osäkerheten på de mindre statliga vägarna eftersom det verkar otroligt att de skulle ha en högre precision på normalvärdet än de stora statliga vägarna.

4.1.3.2 Relativ olycksrisk

Osäkerheten för relativ olycksrisk valdes med utgångspunkt på att det med största sannolikhet är större än osäkerheten i normalvärdet (Forsman 2010). Här hade om möjligt ett ännu större osäkerhetsvärde kunnat användas. Dessutom hade möjligen olika osäkerhetsvärden kunnat användas för olika tätorter.

4.1.3.3 Flöde

I uppskattningen av osäkerheten i trafikflödet så genomfördes ett empiriskt samband. Det kan diskuteras hur bra det var att anpassa två linjära regressionslinjer som ett samband mellan osäkerhet och flöde för statliga vägar och korsningar. Valet gjordes av pragmatiska skäl och med tanke på att Vägverket själva anser att osäkerheten minskar med ökat flöde. Om mer tid funnits till förfogande så hade saken kunnat undersökas vidare. Ännu bättre hade varit att få tillgång till vägverkets data som de publicerat på karttjänsten, fig 14. Då hade man även kunnat undersöka mer i detalj hur mätningarna genomförts och hur osäkerheten skattats för varje väg. Fördelen med ett matematiskt uttryck är att det lättare skulle kunna anpassas till ett större område utan att behöva samla in så mycket data om osäkerheten.

4.2 Diskussion om utvärdering mot polisrapporterade olyckor från STRADA

4.2.1 Utvärdering inom intresseområden

Områdesmässigt så svarar prognoskikten väl mot polisrapporterade olyckor. Generellt sett så har prognoskikten en högre olycksfrekvens än de faktiska utfallen med undantag för samtliga områden inom Linköping tätort och en av de statliga vägarna. Ingen av konfidensintervallen för områdena i Linköping överlappar heller ej överensstämmeelinjen. En tolkning av detta är att modellen parerar bättre för olyckor inom de större vägarna än

stora tätorter. För de övriga områdesklasserna överlappar 1 område vilket påvisar att de faktiska olyckorna inte skiljer sig helt och hållet från den teoretiska fördelningen av möjliga värden på prognoskiktet inom dessa områden.

Resultaten ska tolkas med försiktighet, ett större urval av områden hade behövts undersökas för att kunna uttala sig mer rättvist om resultaten. En kommentar är emellertid att prognosmodellerna om möjligt skulle behöva justeras för större olycksrisk i stora tätorter.

En kritiker ställer nog sig frågan vid det här laget om det verkligen går att utvärdera mot polisrapporterade olyckor när dessa används för beräkningarna av normalvärdena. Det är klart att det måste tas i beaktande men faktum kvarstår att olycksstatistik för hela Sverige använts vid beräkning av medelvärde för normalvärdena vilket gör att det problemet förmodligen är ganska litet.

4.2.2 Utvärdering hela prognoskiktet

Vid närmare studie av graferna från utvärderingen (figur 24, 25) kan det observeras att regressionslinjen ligger lägre än överensstämmelse linjen. Detta indikerar på att olycksrisken är undervärderad i prognoskikten, jämfört med utvärderingsdata från STRADA under åren 2003-2008, oavsett vilken upplösning på prognoskikt som används.

Det är mindre medelavvikelse mellan prognoskiktet och utfallsdata från STRADA på 250 m nivå än 1000 m eftersom NRMSE och MDA värdet är lägre. Däremot så är korrelationskoefficienten högre på 1000m nivå varav den linjära korrelationen är större. Det finns alltså ett något högre samband på 1000 m nivå. Dock så finns det väldigt höga olyckssiffror från STRADA jämfört med prognoskiktet i denna upplösning (se figur 26), vilket även bekräftas med det markant högre RMSE värdet för 1000 m rutor, 1,66 jämfört med 0,201. Därav är det nog lämpligare att använda 250 m upplösning på prognoskikt i ROPA för att undvika att prognosvärden som betydligt undervärderar olycksrisken.

Vid användande av prognoskikt med 250 m upplösning på rutor så kan det förväntas en medelavvikelse (baserat på MDA) på ca 0,043 olyckor per år jämfört med STRADA statistik. Detta måttet ska inte övertolkas då det är ett väldigt förenklat utvärderingsmått som inte tar hänsyn till variationen inom skiktet. Däremot så har MDA används vid liknande utvärderingsstudier av prognosmodeller för fordonsolyckor (Hildebrand et al. 2008). Jämfört med Hildebrand et al. (2008) så ligger MDA värdet i denna studie lägre än de prognosmodeller som de utvärderar.

Det är viktigt att komma ihåg att positioneringen för STRADA olyckorna är inte helt precis. Det finns ett kvalitetsproblem relaterat till lägesnoggrannheten i STRADA vilket gör att en olycka sannolikt kan ligga i en annan ruta. Detta bör dock inte påverka medelavvikelsen särskilt mycket.

4.3 Diskussion om fortsatta studier och egna reflektioner

Under studiens förlopp gjordes en hel del generaliseringar, förenklingar och olika avvägningar för att kunna genomföra studien. Det har uppkommit en hel del obesvarade frågor och önskemål från min sida om hur en osäkerhetsstudie skulle kunna göras bättre och därmed också öka användbarheten.

4.3.1 Fortsatta studier

Den viktigaste fortsättningen av arbetet är att skatta osäkerheten i de olika modellerna d.v.s. olyckor på statliga-, kommunala vägar och korsningar samt olyckor på enskilda vägar. Som diskuterats tidigare så hade ett större och mer precist statistiskt underlag behövts för att genomföra denna uppskattning. Visserligen har en utvärdering genomförts mot utfallsdata men denna ska tolkas mer som en utvärdering av hela prognoskiktet och inte av de olika modellerna. Därmed går det heller inte att använda detta resultat för att uttala sig om osäkerheten i de olika modellerna. Om det däremot hade varit möjligt att skatta en osäkerhet för de olika modellerna, så finns det metoder för att ta med dessa osäkerheter i felfortplantningen och därmed få ett mer komplett osäkerhetsresultat. (Heuvelink 1998).

Vidare så hade det varit intressant att göra en osäkerhetsberäkning för hela Sverige. För att bättre bedöma hur noggrann prognoskiktet är över stora områden. Då hade givetvis en metodik behövts tas fram för att extrahera rätt korsningar från NVDB, då det är ett alltför tidsödande jobb att digitalisera alla korsningar för hand. SWECO Position i Karlstad har utvecklat en metodik för att extrahera korsningar från NVDB för att kunna använda i implementeringen av prognoskiktet. Under studiens gång så har jag utvärderat min implementering mot deras och funnit att det på vissa ställen finns stora skillnader i olycksvärdena mellan våra implementeringar. Detta tyder på att det inte är helt lätt att automatiskt extrahera de korsningar som behövs för uppgiften. Optimalt hade varit om NVDB hade tillhandahållit information om korsnings id och typ av korsning för korsande vägar.

I samband med att göra en osäkerhetsbedömning av prognoskiktet över hela Sverige så hade en djupare studie om parametrarnas osäkerhet eftersträvas. Med fokus då på att studera modellerna och sambanden som skattar relativ olycksrisk och normalvärdena i samråd med VTI för att erhålla en mer trovärdig osäkerhetsbedömning för dessa parametrar. En utökad studie av potentiella osäkerhetskällor hade även inkluderat undersökning av parametrarna a,b och c för de statliga korsningarna. Detta hade i och för sig blivit en lite svårare uppgift då dessa parametrar togs fram på 70-talet av tidigare medarbetare på VTI (Forsman 2010). Rönnbäck (2004) har utvecklat en enkät som kan användas i osäkerhetskattningen av olika variabler. Enkäten är utformad så att användaren inte behöver besitta några nämnvärda statistiska kunskaper, vilket gör att fler personer kan bidra till uppskattningen. I en utökad studie så hade enkäten kunnat användas och skickats till fler

personer som har kunskap om de ingående variablerna och därmed fått en ännu bättre osäkerhetsbedömning.

Som komplement hade en Monte Carlo-simulering kunnat användas för att jämföra osäkerhetsbedömningarna. Monte Carlo-simuleringen har fördelen att den kan användas oavsett funktionernas utformning. Av intresse hade i så fall varit att studera hur osäkerheten i de statliga korsningarna skiljer sig åt då dessa funktioner är de enda som inte är linjära. Heuvelink (1998) har implementerat ett felfortplantningsverktyg ADAM i en forskningsprogramvara som heter PCRaster. Möjligen hade denna programvaran varit användbar vid fortsatta felfortplantningsanalyser.

Bättre utvärderingsstudier mot olycksdata hade kunnat göras. Exempelvis att med hjälp av olika statistiska metoder kombinera utvärderingsmomenten med osäkerhetsuppgifterna från felfortplantningsanalysen. Då hade man kunnat studera om de faktiska olyckorna ligger inom en rimlig fördelning av sannolika värden i prognoskikten baserat på osäkerheten. Samma ansats hade förmodligen varit möjlig att genomföra på områdesnivå.

4.3.2 Egen reflektion

Syftet med rapporten var att uppskatta kvaliteten hos prognoskikten genom att studera hur osäkerheten i indata fortplantas i modellen samt att utvärdera prognoskiket mot utfallsdata. En intressant diskussion är om prognoskiktets kvalitet svarar till sitt tänkta syfte nämligen om de är tillräckligt noggranna för att använda vid planering av räddningstjänst.

Utifrån felfortplantningsanalysen så kan man se att osäkerheten i prognoskikten baserat på osäkerhet i indata är högst varierande med tämligt höga osäkerhetsvärden på sina ställen och dessutom är modellen relativt känslig för fel i indata. Däremot så svarar modellen hyggligt mot faktiska olyckor, i alla fall då större områden jämförs (undantaget är områdena i Linköping). Det som egentligen hade varit avgörande för att bedöma användbarheten av prognosmodellen, vilket även Axelsson (2010) anser, är att undersöka hur olika osäkerheter i prognoskiket påverkar beräkningarna i ROPA. Utan en sådan undersökning är det väldigt svårt att säga att kvaliteten är god nog eller vice versa. Om läsaren erinrar definitionen av kvalitet som beskrivs tidigare i rapporten: *"Kvalitet är egenskaper hos produkt eller tjänst som är av betydelse för dess förmåga att tillgodose kända krav"*. Enligt denna definition så beskrivs kvalitet utifrån sitt användningsområde och ROPA är prognoskiktets användningsområde. Därmed är nog det lämpligaste sättet att utvärdera användbarheten att ta hänsyn till osäkerhetsanalysen och se hur väl prognoskikten fungerar i det faktiska arbetet vid planering av räddningstjänst.

En del viktiga lärdomar har inhämtats gällande osäkerhet och kvalitet i geografiska analyser. Nämligen att även väldigt genomtänkta och genomarbetade rumsliga modeller kan generera relativt stora osäkerheter baserat enbart på osäkerhet i indata. Där noggrannheten av osäkerhetsanalysen står i proportion till hur pass bra osäkerhetsbedömningen av indata.

Däremot så ger felfortplantning av osäkerhet i indata ingen information om hur bra en modell svarar mot "verkligheten".

Felfortplantningsanalyser viktiga verktyg för att uppskatta kvaliteten utav en geografisk analys och att negligera en sådan kan få oanade konsekvenser.

5. Slutsats

5.1 Specifika slutsater

- Baserat på en uppskattad osäkerhet i indata, som valts i samråd med VTI och från Vägverkets och NVDB:s osäkerhetsbedömningar, så kan man vänta sig en varierande kvalitet/osäkerhet i prognoskiktet för fordonsolyckor. Storleken på osäkerheten av de prognostiserade olyckorna uppgår till ungefär 10-15 % i tätorter, 10-15 % på större statliga vägar och varierande osäkerhet upp mot ca 40 %, på mindre statliga vägar. På enskilda vägar utanför tätorter kan man vänta sig en osäkerhet i prognostiserade olyckor mellan 2-3 % baserat på osäkerhet i indata.
- Resultaten från felfortplantningsanalysen i denna studie ger enbart ett mått på kvaliteten baserat på osäkerhet i indata. För en total osäkerhetsbedömning krävs det att modellernas representation av det geografiska fenomenet tas i beaktande. Vid en sådan analys kan det väntas en större osäkerhet överlag men särskilt på det enskilda vägnätet då denna modell är väldigt förenklad.
- Prognoskiken från VTI indikerar på god kvantitativ överensstämmelse i Linköpings omnejd, mot faktiska olyckor från STRADA under åren 2003-2008, på områden med större statliga vägar. Sämre överensstämmelse indikeras i Linköping tätort.
- För att bedöma användbarheten utav prognoskikten baserat på kvaliteten krävs vidare studier i hur osäkerheten i prognoskikten påverkar slutresultaten i ROPA.

5.1 Allmänna slutsater

- Att använda en geografisk modell okritiskt utan kvalitetsgranskning är ej att rekommendera då det kan förekomma flera potentiellt stora osäkerhetskällor som kan påverka resultatet.
- För att kunna uppskatta alla möjliga osäkerheter som finns i indata i en geografisk modell så behövs detaljstudier rörande varje osäkerhetskälla.
- Analytisk felfortplantning med första ordningen Taylor fortplantning är ett effektivt verktyg för att studera hur olika osäkerhetskällor påverkar kvaliteten av geografiska analyser, förutsatt att modellerna som bygger upp analysen är approximativt linjära i förhållande till indata och att osäkerhetsuppgifterna för indata är rätt valda.

Referenser

Artiklar och Böcker

Beissinger, S, R & Westphal, M, I (1998). *"On the Use of Demographic Models of Population Viability in Endangered Species Management"*. Journal of wildlife management. Vol 62. s 821-841. Allen_Press

Burrough, P, A & McDonnen, R, A (1998). *"Principles of Geographical Information systems"*. Oxford University Press. Oxford

Brage, T (2004) *"Kvalitetsgranskning av servicedatabasen Sisyla"* Lund universitet, Institutionen för naturgeografi och ekosystemanalys, Seminarieuppsats 105, Tillgänglig via <http://www.geobib.lu.se>

Crosetto, M & Tarantola, S (2000) *"Uncertainty and sensitivity analysis: tools for GIS-based model implementation"* International Journal of Geographical Information Science" (2001), Vol 15, No 5, s 415-437

Devillers, R. Bédard, Y. Jeansoulin, R. (2005) *"Multi dimensional management of geospatial data quality information for its dynamic use within GIS"* Phogrammetric Engineering & Remote Sensing. Vol 71. No 2. February 2005. s 205-215

Eastman, J, R & Hong, J (2000). *"Application of fuzzy measures in multi-criteria evaluation in GIS"*, International Journal of Geographical Information Science. Vol. 14. Nr 2. s. 173-184

Evans, E, J (2001) *"Tactical Weather Decision Support To Complement "Strategic" Traffic Flow Management for Convective Weather"* 4th USA/Europe Air Traffic Management R&D Seminar Santa Fe, New Mexico 3-7 December

Forsling, G & Neymark, M (2004), *"Matematisk analys en variabel"* LIBER, ISBN : 9789147051885

Frank, A, U, (2008). *"Analysis of dependence of decision quality on data quality"*. Journal of Geographical Systems Vol 10, 2008, s 8-59

Gunnarsson, J. Ollvik, L(2002). *"Geodetisk mätningsteknik"*. kompendium på Lunds Tekniska Högskola. Institutionen för teknik och samhälle. KFS AB.

Heuvelink, G (1998). *"Error Propagation in Environmental Modelling with GIS"* London. Taylor & Francis Ltd. ISBN: 0-7484-0743-X HB

Hildebrand, E. Robichaud, K. Ye, H (2008) "*Evaluation of accident prediction for rural highways*". Canadian Journal of Civil Engineering. Vol 35. s 647-651

Hwang, D. Karimi. A, H. Byun, D, W (1997) "*Uncertainty analysis of environmental models within GIS environments*" Computers & Geosciences. Volume 24, Issue 2. 20 March 1998. S, 119-130

Jacobsson, A (2008) *On the future of topographic base information management in Finland and Europe*. Helsinki. National Land Survey of Finland. 101. ISBN: 951-48-0192-X

Jacobsson, A & Marttinen J (2003). "*Data quality management of reference dataset. present practice in european national mapping agencies and a proposal for a new approach*". Proceedings of the 21st International Cartographic Conference (ICC). *Cartographic Renaissance*. ISBN: 0-958-46093-0

Leung, Y. Ma, L. Goodchild, M, F (2004). "*A general framework for error analysis in measurement-based GIS*". Journal of Geographical Systems (2004). Vol 6. s 325-354.

McBratney A. B. & Odeh. I O.A. (1997). "*Application of fuzzy sets in soil science: fuzzy logic measurements and fuzzy decisions*". Geoderma. Vol 77. s. 85-113

Meyer,S.L. (1975) "*Data Analysis for Scientists and Engineers*". Wiley ISBN 0-471-59995-6

Oksanen, J. Sarakoski T (2005)."*Error propagation analysis of DEM-based drainage basin delineation*". International Symposium on Spatial Data Quality (2005). Vol. 26. s 3085-3102

Olsson, U. Englund, J, Engstrand, U (2005)."*Biometri – Grundläggande biologisk statistik*". Studentlitteratur. Lund. Sweden. ISBN: 978-91-44-04577-1

Rönnbäck, Britt-Inger, (2004) "*Are uncertain uncertainties useful?: towards improved quality assessment of spatial data*" Doctoral Thesis: 2004: 13, Civil and Environmental Engineering / Geographical Information Technology, Luleå University of Technology. ISSN: 1402-1544

Stehman, S, V & Czaplewski, R, L(1998)"*Design and analysis for thematic map accuracy assessment: Fundamental principles*". Remote sens. Environ(1998).Vol 64. s 331 -344, Elsevier Science Inc. 1998. 655 Avenue of the Americas. New York, NY 10010

Tastan, H. Altan, M.O (2002). "*Spatial data quality – Concepts and ISO Standards*". OEEPE/ISPRS Joint Workshop on Spatial Data Quality Management. 21-22 mars 2002. Istanbul. Turkiet.

Thompson, M,K(2000) "*Quality evaluation procedure for the quantitative horizontal positional accuracy assessment of orthophotos of the Swiss National Park*"

Verigin, H & Lanter, D (1995) "Data-quality enhancement techniques in layer-based geographic information systems". Comput. Environ. And Urban Systems (1995). Vol 19.s 23-36

Wayne, L (2001) "Institutionalize metadata before it Institutionalizes you". URISA (2001) Federal geographic data committee (2006)

Wasström, C. Lönnberg, G. Harrie, L (2008) "Kvalitetsaspekter" Geografisk Informationsbehandling metoder och tillämpningar. Red: Harrie, Lars. s 237-255. Forskningsrådet Formas. ISBN: 978-540-6015-3

Worboys, M (2004) "GIS a computing perspective" CRC Press LLC, 2000 N.W. Corporat Blvd. Boca Raton. Florida 33431 ISBN: 0-415-28375-2

Zhang, J. Goodchild, F, M (2002) "Uncertainty in Geographical Information". Taylor&Francis. New Fetter Lane, London. ISBN: 0-415-27723-X

Internet

2010-03-21 Hitta.se
<http://www.hitta.se>

2010-03-25 Lantmäteriet. *Nya HMK-Kvalitet*
http://www.lantmateriet.se/templates/LMV_Page.aspx?id=15918

2010-02-04. SMHI . *Hydrologiska sannolikhetsprognoser*.
<http://www.smhi.se/kunskapsbanken/hydrologi/hydrologiska-sannolikhetsprognoser-1.5284?print=yes>

2010-02-09 Skanetraffiken
<http://www.skanetraffiken.se>

2009-09-18. Nationalencyklopedin. *Kvalitet*. Lång.
<http://www.ne.se/ludwig.lub.lu.se/lang/kvalitet/234210>

2009-09-18. Nationalencyklopedin *Fel*. Lång
<http://www.ne.se/ludwig.lub.lu.se/lang/fel/168082>

2009-10-01. Nationalencyklopedin *Vaghet*. Lång
<http://www.ne.se/vaghet>

2009-10-28. Nationalencyklopedin *Kvantitativ metod*. Lång
<http://www.ne.se/lang/kvantitativ-metod>

2009-11-04. Nationalencyklopedin *Regressionsanalys*. Lång
<http://www.ne.se/regressionsanalys>

2009-10-28. Nationalencyklopedin. *Statistik*. Lång.
<http://www.ne.se/lang/statistik>

2009-10-29. Vägverket. *Defintioner om vägar*.
http://www22.vv.se/templates/page3____19807.aspx

2010-03-16 Vägverket *Allmänt om vägverkets Trafikmätningar*
<http://www.vv.se/Startsida-foretag/vagar/Vagnatet/Trafikfloden-och-medelhastigheter/Allmant-om-Vagverkets-trafikmatningar/>

2010-03-16 Vägverket *Effektsamband*
http://publikationswebbutik.vv.se/Shopping/Default____3366.aspx

Personlig kommunikation

Axelsson Anders, Projektledare BeRädd, MSB, maj 2010

Forsman Åsa, Forskare, VTI, oktober 2009, mars 2010

Harrie Lars, Lektor, Lunds universitet, mars 2010

Övrigt

Axelsson, A (2008). "*Berädd – ett verktyg för planering av räddningstjänst*". Statusrapport Räddningsverket (MSB) 2008-12-18

Goodchild, M (2000). "*Communicating the results of accuracy assessment: Metadata, digital libraries, and fitness for use*". National center of Geographic Information and Analysis, Univerity of California. Hämtad från: <http://www.geog.ucsb.edu/~good/papers/325.pdf> 10-03-05

Jaldell, Henrik (2009) "*Förväntat antal bränder – underlag för BeRädd*". Myndigheten för samhällsskydd och beredskap (2009-08-26)

Lundqvist, M. McIntyre C, Hedman U (2008) "*Statistik och analys – Räddningstjänst i siffror*" Myndigheten för samhällsskydd och beredskap. Publikationsnummer: MSB 0107-09. ISBN: 978-91-7383-046-1

Matsson, M-O (2009) *"NVDB-Specifikation av innehåll, Vägnät version 5.4"* Vägverket , enhet Väg – och trafikdata. 2009:86. ISSN 1401-9612

Sjöö, Bengt. Ungerbäck, Ann-Christin (2007) *"Nytt nationellt informationssystem för skador och olyckor inom hela vägtransportssystemet"* Vägverket, ISSN: 1401-9612

SIS, (Swedish Standards Institute). *"Samverkande GIS med ISO 19100 – En handbok om teknisk ramverk för geografisk information"* Stanli - SIS projektområde för geografisk information. Utgåva 1 September 2004

SFS 2003:778 *"Lag om skydd vid olyckor"* Förvarsdepartementet (2003-11-20). Källa: Regeringskansliets rättsdatabaser

Bilaga 1- Tabeller från vägverkets effektkatalog

Figur 1: Figuren är ett utklipp från vägverkets effektkatalog och visar vilket normalvärde som ska väljas på statliga vägar beroende på vad det är för vägtyp, breddklass och hastighetsgräns.

	Vägtyp	Hastighetsgräns	Breddklass	Normalvärde
Vanlig väg, ej motoriserad trafikbredd (ML)	2kf	70	<8m	0,153
	2kf	70	>=8m (inkl. ML)	0,107
	2kf	80	<8m	0,107
	2kf	80	>=8m (inkl. ML)	0,077
	2kf	90	<8m	0,113
	2kf	90	>=8m (inkl. ML)	0,082
	2kf	100	<8m	0,092
	2kf	100	>=8m (inkl. ML)	0,072
	2kf	110	<8m	0,098
Vanlig väg motoriserad trafikbredd (MML)	MLV	80		0,063
	MLV	90		0,067
	MLV	100		0,084
	MLV	110		0,089
	MML	80		0,073
	MML	90		0,077
Motorväg, alla utformningar	MV	70		0,174
	MV	80		0,082
	MV	90		0,088
	MV	100		0,062
	MV	110		0,067
4kf, alla utformningar	4kf	80		0,099
	4kf	90		0,105
	4kf	100		0,083
	4kf	110		0,088
Vanlig väg	2kf	50		0,181
	2kf	60		0,118
	>2kf	50		0,290
	>2kf	60		0,157

Figur 2: Figuren är ett utklipp från vägverkets effektkatalog och visar vilken relativ olycksrisk som ska väljas på kommunala korsningar beroende på funktionell vägklass, korsningstyp och hastighetsgräns.

Funktionell vägklass	Hastighet på vägen med högst flöde	3-vägs korsning			4-vägs korsning		5 eller fler ben
		Typ 1	Typ 2 och 3	Typ 1	Typ 2 och 3	Typ 2 och 3	
alla ben klass 0-6	30 - 50 km/h	2,250	1,500	2,813	1,875	1,875	
	60 - 70 km/h	2,813	1,875	3,516	2,344	2,344	
ett ben klass 7-9, 2 eller 3 ben klass 0-6	30 - 50 km/h	1,500	1,000	2,250	1,500	1,500	
	60 - 70 km/h	1,875	1,250	2,813	1,875	1,875	
minst två ben klass 7-9	30 - 50 km/h	1,500	1,000	1,875	1,250	1,250	
	60 - 70 km/h	1,875	1,250	2,344	1,563	1,563	
alla ben klass 7-9	30 - 50 km/h	0,000	0,000	0,000	0,000	0,000	
	60 - 70 km/h	0,000	0,000	0,000	0,000	0,000	

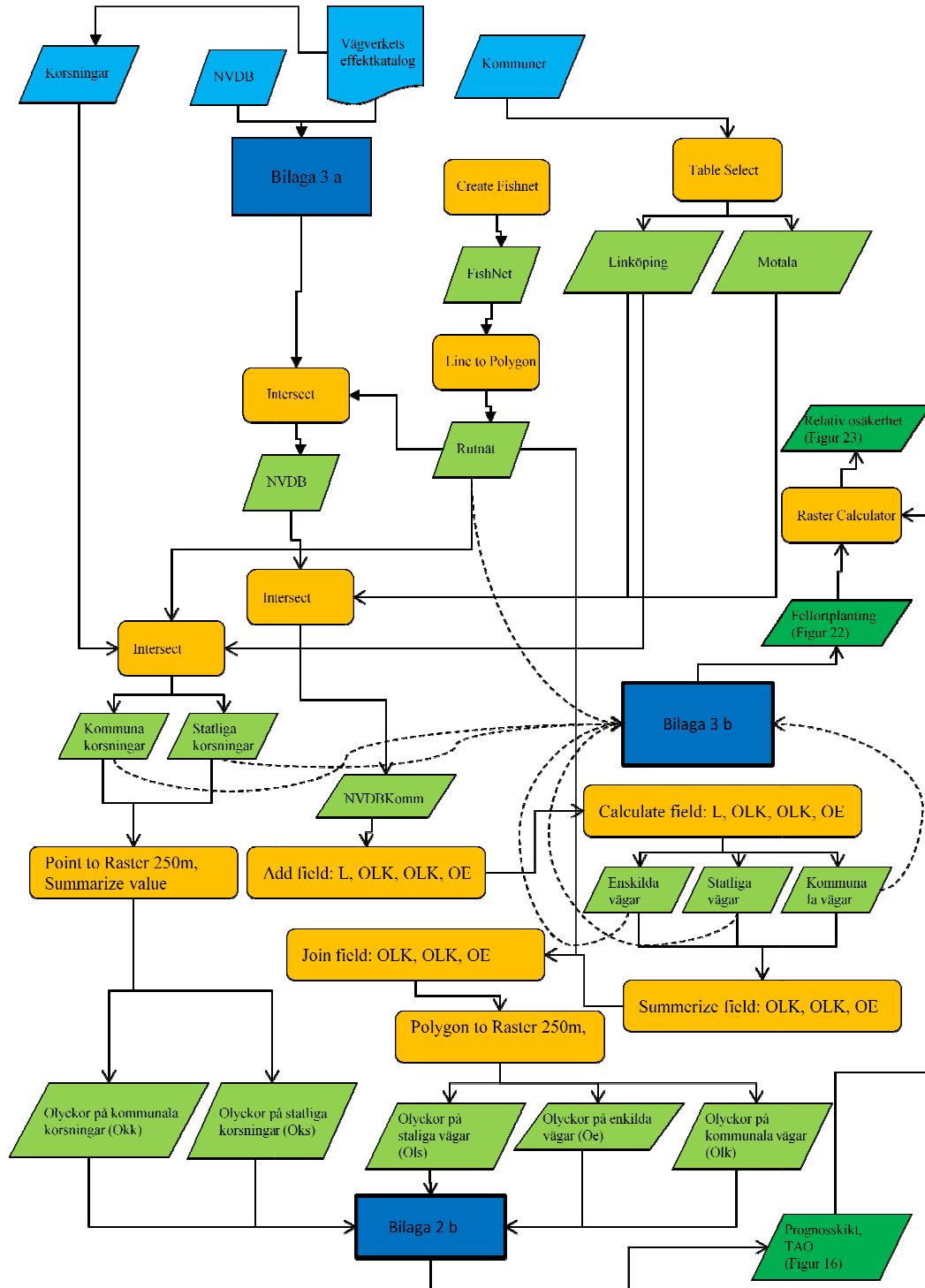
Figur 3: Figuren är ett utklipp från vägverkets effektkatalog och visar vilken relativ olycksrisk som ska väljas på kommunala väglänkar beroende på funktionell vägklass och hastighetsgräns.

Funktionell vägklass	Hastighetsgräns	Relativ olycksrisk
3	30-50 km/h	2,9
4	30-50 km/h	2,6
5	30-50 km/h	1,9
6	30-50 km/h	0,8
7	30-50 km/h	0,2
8	30-50 km/h	0,2
9	30-50 km/h	0,2
3	60-70 km/h	2,1
4	60-70 km/h	1,9
5	60-70 km/h	1,5

Figur 4: Figuren är ett utklipp från vägverkets effektkatalog och visar vilka konstanter a,b och c som ska väljas på statliga korsningar beroende på korsningstyp och hastighetsgräns.

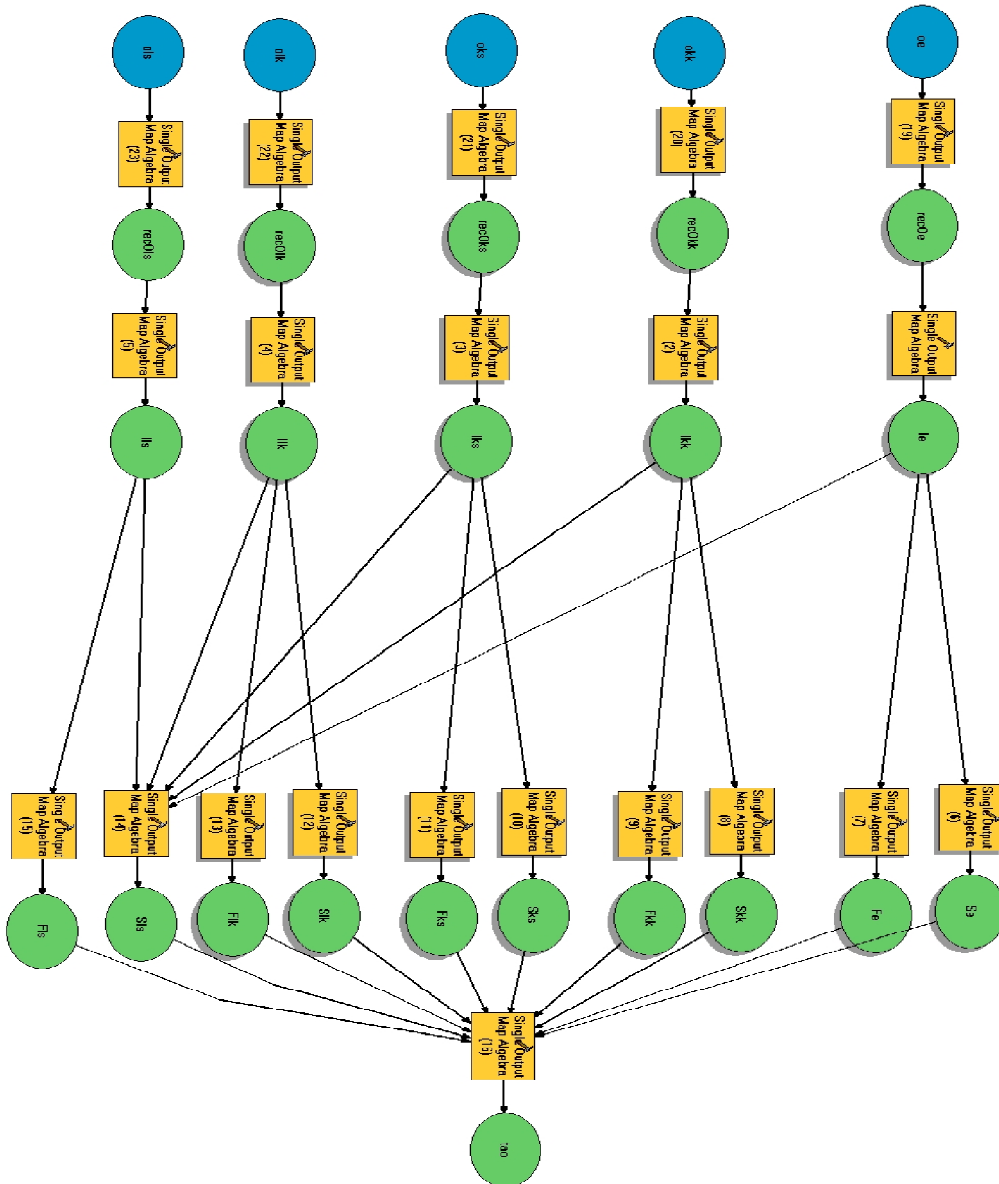
Hastighet på vägen med högst flöde	Korsningstyp	3-vägs korsning			4-vägs korsning			4-vägs snedfördelad		
		a	b	c	a	b	c	a	b	c
40 eller 50 km/h	1	1.877 x 10 ⁻⁶	1,45	0,50	3.905 x 10 ⁻⁶	1,45	0,60	2.065 x 10 ⁻⁶	1,45	0,50
	2	6.893 x 10 ⁻⁶	1,20	0,00	9.148 x 10 ⁻⁶	1,20	0,00			
	3	1.502 x 10 ⁻⁶	1,45	0,50	0.939 x 10 ⁻⁶	1,45	0,50			
60 eller 70 km/h	1	22.264 x 10 ⁻⁶	1,25	0,45	41.448 x 10 ⁻⁶	1,25	0,55	24.490 x 10 ⁻⁶	1,25	0,45
	2	7.220 x 10 ⁻⁶	1,20	0,00	9.572 x 10 ⁻⁶	1,20	0,00			
	3	16.030 x 10 ⁻⁶	1,25	0,45	10.019 x 10 ⁻⁶	1,25	0,45			
80 eller 90 km/h	1	25.080 x 10 ⁻⁶	1,25	0,45	41.448 x 10 ⁻⁶	1,25	0,55	27.588 x 10 ⁻⁶	1,25	0,45
	2									
	3	18.058 x 10 ⁻⁶	1,25	0,45	11.286 x 10 ⁻⁶	1,25	0,45			
100 eller 110 km/h	1	25.080 x 10 ⁻⁶	1,25	0,45	41.448 x 10 ⁻⁶	1,25	0,55	27.588 x 10 ⁻⁶	1,25	0,45
	2									
	3	18.058 x 10 ⁻⁶	1,25	0,45	11.286 x 10 ⁻⁶	1,25	0,45			

Bilaga 2 a Implementering av prognosskikten och osäkerhetsanalysen



Figur Schemat visar den huvudsakliga arbetsgången i studien. De ljusblå objekten illustrerar indata bestående av NVDB, Kommuner, egen digitaliserade korsningar och vägverkets effekttabell från Bilaga 1. De gula rutorna visar de verktyg som används i ArcGIS vilka resulterar i delresultat illustrerade som ljusgröna romber. Mörkblå rutor hänvisar till andra bilagor. Slutresultaten illustreras som mörkgröna romber, där figurnumret i parantes hänvisar till figur i rapporten.

Bilaga 2 b Utsnitt från ArcGIS modelbuilder



Figur Flödesschema från ArcGIS modelbuilder som visar det sista steget i att skapa prognoskiktet (*tao*). De gula rutorna illustrerar verktyg där de sista beräkningarna i prognosmodellen genomförts d.v.s. justering för insatser och uppdelning i singel och flerfordonsolyckor. De blåa ellipserna är indata bestående av raster innehållande de olika olyckstyper. De gröna ellipserna är resultat från beräkningarna.

Bilaga 3 a: Matlab Script: Tilldelning av relativ olycksrisk och normalvärden för vägar i NVDB

```
% Program som lägger till värden för N och R till de statliga och kommunala
% länkarna respektive.
%
%Författare : Edvin Eriksson
%
%Laddar in shapefilen med NVDB data samt två textfiler med normalvärden och
%relativa olycksrisker.

%Textfiler med normalvärden och relativolycksrisk som hittas i figur 1 och
%3 i bilaga 1.
%
sLNorm = load...
    ('SLNorm.txt');
kLRel =load...
    ('KLR.txt');

%Shapefil med vägar från NVDB
nvdb = shaperead...
    ('NVDB');

% Kör igenom alla vägar och om kriterierna stämmer överens så får väglänken
% ny attributinformation
for i = 1: length(nvdb(:,1))
    nvdb(i,1).N=0;
    nvdb(i,1).R=0;
    nvdb(i,1).BK=0;
    if (nvdb(i,1).VAGHA_6 == 1)
        nvdb(i,1).R = 0;
        if(nvdb(i,1).BREDD_156 < 8 && nvdb(i,1).BREDD_156 > 0)
            nvdb(i,1).BK = 1;
        elseif(nvdb(i,1).BREDD_156 >= 8.0)
            nvdb(i,1).BK=2;
        else
            nvdb(i,1).BK=0;
        end
    end
    for j = 1: length(sLNorm(:,1))
        if((nvdb(i,1).VAGTY_41 == 2 || nvdb(i,1).VAGTY_41 == 5)...
            && nvdb(i,1).F_HOGST_22 >=70)
            if(nvdb(i,1).VAGTY_41 == sLNorm(j,1) && ...
                nvdb(i,1).F_HOGST_22 == sLNorm(j,2) && ...
                nvdb(i,1).BK == sLNorm(j,3))
                nvdb(i,1).N = sLNorm(j,4);
            end
        else
            if(nvdb(i,1).VAGTY_41 == sLNorm(j,1) && ...
                nvdb(i,1).F_HOGST_22 == sLNorm(j,2))
                nvdb(i,1).N = sLNorm(j,4);
            end
        end
    end
    elseif(nvdb(i,1).VAGHA_6 == 2 || (nvdb(i,1).VAGHA_6 == 3 ...
        && nvdb(i,1).F_HOGST_22 <=50))
        for j = 1: length(kLRel(:,1))
            if(nvdb(i,1).KLASS_181 == kLRel(j,1) && ...
                nvdb(i,1).F_HOGST_22 == kLRel(j,2))
                nvdb(i,1).R = kLRel(j,3);
            end
        end
    end
end

%skriver ut shapefilen med ny attributinformation
shapewrite(nvdb,'nvdbNew.shp');
```

Bilaga 3 b: Implementering av felfortplantningen

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Huvudprogram som beräknar det totala felet och adderar det till en %
% .shp fil i form av ett polygonrutnät med 250*250 m stora celler. Input%
% i programmet är utsnitt i NVDB som processerats i ArcMap för att %
% Programmet använder MathWorks inbyggda Mapping toolbox som konvererar %
% .shp filer till MathWorks geostructure filer. Medelfelet för varje %
% ruta beräknas med första gradens Taylor utveckling. De övriga stegen %
% finns i funktionerna errorOLS, errorOLK,errorOKK, errorOE och %
% errorOKS. %
%
%
% Författare: Edvin Eriksson %
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Adressen till shapefilerna som innehåller vägarna,korsningarna samt
%skiktet som ska uppdateras med värden för osäkerhet

pathOKK = ('C:\TempData\EdvinTAINTEBORT\ErrorPropagation\okk');
pathOLK = ('C:\TempData\EdvinTAINTEBORT\ErrorPropagation\olk');
pathOLS = ('C:\TempData\EdvinTAINTEBORT\ErrorPropagation\ols');
pathOE = ('C:\TempData\EdvinTAINTEBORT\ErrorPropagation\oe');
pathOKS = ('C:\TempData\EdvinTAINTEBORT\ErrorPropagation\OKS');
pathError = ('C:\TempData\EdvinTAINTEBORT\ErrorPropagation\errorPoly');

%Läser in .shp filerna med en inbyggd MathWorksfunktion: shaperead.
unsortedOKK = shaperead(pathOKK);
unsortedOLK = shaperead(pathOLK);
unsortedOLS = shaperead(pathOLS);
unsortedOE = shaperead(pathOE);
unsortedOKS = shaperead(pathOKS);
error = shaperead(pathError);

%Sorterar vägarna med avseende på rutid för att påskynda algoritmen,
%FID_fishPo är rutid för vägarna

[unused, order] = sort([unsortedOLK(:).FID_fishPo]);
olk = unsortedOLK(order);

[unused2, order2] = sort([unsortedOLS(:).FID_fishPo]);
ols = unsortedOLS(order2);

[unused3, order3] = sort([unsortedOE(:).FID_fishPo]);
oe = unsortedOE(order3);

[unused4, order4] = sort([unsortedOKS(:).FID_fishPo]);
oks = unsortedOKS(order4);

[unused5, order5] = sort([unsortedOKK(:).FID_fishPo]);
okk = unsortedOKK(order5);

%Längden på strukturvektorererna definerar slutet på loopen
%Indexen är start index för varje loop vilka ändras inuti loopen så
%att looparna hoppar över celler som redan besökts.

lengthOLS = length(ols(:,1));
lengthOLK = length(olk(:,1));
lengthOE = length(oe(:,1));
lengthOKS = length(oks(:,1));
lengthOKK = length(okk(:,1));

%start index för vägarna och korsningarna
indexOLS = 1;
indexOLK = 1;
indexOE = 1;
```

```
indexOKS = 1;
indexOKK = 1;

for i = 1 : 6052

    %De summerade felen i varje ruta i det första steget
    sumOLS = 0;
    sumOLK = 0;
    sumOE = 0;
    sumOKS = 0;
    sumOKK = 0;

    %Loopar igenom alla vägar och korsningar i varje ruta, om det inte
    %finns några vägar eller korsningar inom rutan så avbryts loopen,
    for j= indexOLS: lengthOLS
        if(i == (ols(j,1).FID_fishPo + 1))
            errOLS = errorOls(ols(j,1));
            sumOLS = errOLS +sumOLS;
            if(j ~= lengthOLS)
                if((ols(j,1).FID_fishPo)~= (ols(j+1,1).FID_fishPo))
                    indexOLS = j+1;
                    break;
                end
            end
        end
    end

    for j= indexOLK: lengthOLK
        if(i == (olk(j,1).FID_fishPo + 1))
            errOLK = errorOlk(olk(j,1));
            sumOLK = errOLK +sumOLK;
            if(j ~= lengthOLK)
                if((olk(j,1).FID_fishPo)~= (olk(j+1,1).FID_fishPo))
                    indexOLK = j+1;
                    break;
                end
            end
        end
    end

    for j= indexOE: lengthOE
        if(i == (oe(j,1).FID_fishPo + 1))
            erroE = errorOe(oe(j,1));
            sumOE = erroE +sumOE;
            if(j ~= lengthOE)
                if((oe(j,1).FID_fishPo)~= (oe(j+1,1).FID_fishPo))
                    indexOE = j+1;
                    break;
                end
            end
        end
    end

    for j= indexOKS: lengthOKS
        if(i == (oks(j,1).FID_fishPo + 1))
            errOKS = errorOKS(oks(j,1));
            sumOKS = errOKS +sumOKS;
            if(j ~= lengthOKS)
                if((oks(j,1).FID_fishPo)~= (oks(j+1,1).FID_fishPo));
                    indexOKS = j+1;
                    break;
                end
            end
        end
    end

    for j= indexOKK: lengthOKK
        if(i == (okk(j,1).FID_fishPo + 1))
            errOKK = errorOKK(okk(j,1));
            sumOKK = errOKK +sumOKK;
            if(j ~= lengthOKK)
                if((okk(j,1).FID_fishPo)~= (okk(j+1,1).FID_fishPo));
```

```

                indexOKK = j+1;
                break;
            end
        end
    end
end
disp(i)

%Nästa steg är att beräkna påverkan på insatserna ILS, ILK, IKS, IE
errILS = sqrt(0.76^2*sumOLS^2);
errILK = sqrt(0.76^2*sumOLK^2);
errIKS = sqrt(0.76^2*sumOKS^2);
errIE = sqrt(0.76^2*sumOE^2);
errIKK = sqrt(0.76^2*sumOKK^2);

%Sista steget utifrån VTI modellen beräknar felet i FE, SE, FKS, SKS,
%SLS, FLS, SLK, FLK
errFE = sqrt(0.3^2*errIE^2);
errSE = sqrt(0.7^2*errIE^2);
errFKS = sqrt(0.3^2*errIE^2);
errSKS = sqrt(0.7^2*errIKS^2);
errSLS = sqrt((0.73^2*errILS^2)+(0.03^2*errILK^2)+...
(0.03^2*errIE^2)+(0.03^2*errIKS^2)+(0.03^2*errIKK^2));
errFLS = sqrt(0.3^2*errILS^2);
errSLK = sqrt(0.63^2*errILK^2);
errFLK = sqrt(0.37^2*errILK^2);
errSKK = sqrt(0.63^2*errIKK^2);
errFKK = sqrt(0.37^2*errIKK^2);

%Medelfelet för varje ruta beräknas slutligen som summan av felen i
%ovanstående beräkning.

%Osäkerheten i varje ruta med avseende på OLS, OLK, OE och OKS
error(i,1).errorOLS =sumOLS;
error(i,1).errorOLK =sumOLK;
error(i,1).errorOE =sumOE;
error(i,1).errorOKS = sumOKS;
error(i,1).errorOKK = sumOKK;

%Total osäkerhet
error(i,1).errorAll = sqrt(errFE^2 + errSE^2 + errFKS^2 ...
+ errSKS^2 + errSLS^2 + errFLS^2 + errSLK^2 + errFLK^2 ...
+ errSKK^2 + errFKK^2);

%Osäkerhet i singelolyckor
error(i,1).errorS = sqrt(errSE^2 + errSKS^2 + errSLS^2 + ...
errSLK^2 + errSKK^2);

%Osäkerhet i multiolyckor
error(i,1).errorF = sqrt(errFE^2 + errFKS^2 + errFLS^2 + ...
errFLK^2 + errFKK^2);
end

% Skriver ut en .shp fil med osäkerhetsinformation
outPath = 'C:\TempData\EdvinTAINTEBORT>ErrorPropagation\errorMat';
shapewrite(error,outPath);

-----

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Funktion som beräknar osäkerheten i OLK(j,1)
%
% Input = väglänk x
% Output = osäkerhet c
%
% Författare: Edvin Eriksson
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

function [c] = errorOlk(x)

%Parametrar
uncertL= 0.02; %Osäkerhet i längden
uncertR = 0.15; %Osäkerhet i R
L = x.L/1000.0; %Längden
R = x.R; %Relativ olycksrisk

%Relativ osäkerhet i Olk(j,1).L och Olk(j,1).R
errorL = L * uncertL;
errorR = R *uncertR;

%Felfortplantning av osäkerhet i L och R i OR, L är derivatan med avseende
%på R och tvärtom R
errorOR = sqrt((L^2*errorR^2)+(R^2*errorL^2));

%Parametrar om summan av relativ olycksrisk * längd, OR samt Oobs som
%varierar beronde på kommun
sumOR = 276.662023;
Oobs = 0;

if(strcmp(x.KNKOD,'0580')==1);
    Oobs = 173.5;
elseif(strcmp(x.KNKOD,'0583')==1);
    Oobs = 41.5;
end

%derivatan av OLK(j,1) med avseende på OR
term = (1/sumOR)*0.5*Oobs;

%Den totala osäkerheten(c) i OLK(j,1) beräknas genom felfortplantning, se ekvation 23
c = sqrt(term^2*errorOR^2);

-----

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Funktion som beräknar osäkerheten i OLS(j,1)
%
% Input = väglänk x
% Output = osäkerhet c
%
% Författare: Edvin Eriksson
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [c] = errorOls(x)

%parametrar
treasV = 1500; %Tröskelvärde för osäkerhetsregressionerna
uncertN = 0.1; %Osäkerheten i Normalvärdet är 10 %
uncertL= 0.02; %Osäkerheten i längden är 2 %
L = x.L/1000; %Längden
ADT = x.ADT_A_113; %Flöde axelpar
N = x.N; %Normalvärde
Q = 365*ADT; %Flöde per år

%Relativ osäkerhet i OLS(j,1) beroende på ADT
if(ADT < treasV)
    uncertADT = (56.1018 + (-0.0396*ADT))/100;
else
    uncertADT = (14.6471 + (((-3.0237)*(10)^(-4))*ADT))/100;
end

%Relativ osäkerhet i OLS(j,1).ADT
errorADT = ADT*uncertADT;

%Osäkerhet beroende av Q, beräknas genom felfortplantning, 365 är
%derivatan av Q med avseende på ADT
errorQ = sqrt(365^2*errorADT^2);

```

```
errorN = N * uncertN; %Relativ osäkerhet i OLS(j,1) beroende på N
errorL = L * uncertL; %Relativ osäkerhet i OLS(j,1) beroende på L

%Derivering av OLS(j,1) med avseende på L, Q och N
termOne = N*Q/1000000;
termTwo = L*N/1000000;
termThree = L*Q/1000000;

%Den totala osäkerheten(c) i OLS(j,1) beräknas genom felfortplantning
c = sqrt((termOne^2*errorL^2) + (termTwo^2*errorQ^2) + ...
        (termThree^2*errorN^2));
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Funktion som beräknar osäkerheten i Oe(j,1)
%
% Input = väglänk x
% Output = osäkerhet c
%
% Författare: Edvin Eriksson
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
function [c] = errorOe(x)

%parametrar
uncertL= 0.02; %Osäkerheten i längden 2%
L = x.L /1000; %Längden på länken i rutan

%Relativ osäkerhet för längden Oe(j.1).L
errorL = L * uncertL;

%Den totala osäkerheten(c) i Oe(j,1) beräknas genom felfortplantning
%0.03 är derivatan av Oe med avseende på längden
c = sqrt(0.003^2*errorL^2);
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Funktion som beräknar osäkerheten i OKK(j,1)
%
% Input = korsning x
% Output = osäkerhet c
%
% Författare: Edvin Eriksson
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
function [c] = errorOKK(x)

%parametrar
uncertR = 0.15; %Osäkerheten i R är 15 %
Oobs = x.Oobs; %Observerat antal olyckor
sumR = x.SumR; %Summan av R i alla korsningar
R = x.R; %Relativ olycksrisk för korsningen

%Relativ osäkerhet för OKK(j,1).R
errorR = R*uncertR;

%derivatan av OKK med avseende på R
term = (0.5*Oobs*(1/sumR));

%Den totala osäkerheten(c) i OKK(j,1) beräknas genom felfortplantning
c = sqrt(term^2*errorR^2);
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Funktion som beräknar osäkerheten i OKS(j,1)
%
% Input = korsning x
% Output = osäkerhet c
%
% Författare: Edvin Eriksson
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [z] = errorOKS(x)

%Parametrar

a = x.a;%Konstanter a,b,c som används i modellen
b = x.b;
c = x.c;
treasV = 1500;%tröskelvärde i flöde-osäkerhets regressionen
Qs = x.Qs; %Inkommande flöde sekundärvägar och totalt
Qt = x.Qt;

% Villkor för procentuell osäkerhet, om flödet är mindre än 1500 använd
% regressions modell 1 annars modell 2

if(Qs <treasV)
    uncertQs= (56.1018 + (-0.0396*Qs))/100;%Regressionsfunktion 1
    if(Qt<treasV)
        uncertQt= (56.1018 + (-0.0396*Qs))/100;
    else
        uncertQt= (14.6471 + (((-3.0237)*(10)^(-4))*Qt))/100;%Regressionfunktion 2
    end
else
    uncertQs= (14.6471 + (((-3.0237)*(10)^(-4))*Qs))/100;
    uncertQt= (14.6471 + (((-3.0237)*(10)^(-4))*Qt))/100;
end

%Relativ osäkerhet för Qt och Qs
errorQt = Qt*uncertQt;
errorQs = Qs*uncertQs;

% Derivering av OKS med hänseende på Qt(termOne) och Qs(termTwo)
termOne = a*(b-c)*Qs.^c*Qt^(b-c-1);
termTwo = a*c*Qt.^(b-c)*Qs.^(c-1);

%Den totala osäkerheten(c) i OKS(j,1) beräknas genom felfortplantning
z = sqrt((termOne^2*errorQt^2)+(termTwo^2*errorQs^2));

-----
```

Lunds Universitets Naturgeografiska institution. Seminarieuppsatser. Uppsatserna finns tillgängliga på Naturgeografiska institutionens bibliotek, Sölvegatan 12, 223 62 LUND. Serien startade 1985. Hela listan och själva uppsatserna är även tillgängliga på <http://www.geobib.lu.se/>

The reports are available at the Geo-Library, Department of Physical Geography, University of Lund, Sölvegatan 12, S-223 62 Lund, Sweden.

Report series started 1985. The whole complete list and electronic versions are available at <http://www.geobib.lu.se/>

- 156 Cederlund, Emma (2009): Metodgranskning av Klimatkommunernas lathund för inventering av växthusgasutsläpp från en kommun
- 157 Öberg, Hanna (2009): GIS-användning i katastrofdrabbade utvecklingsländer
- 158 Marion Früchtl & Miriam Hurkuck (2009): Reproduction of methane emissions from terrestrial plants under aerobic conditions
- 159 Florian Sallaba (2009): Potential of a Post-Classification Change Detection Analysis to Identify Land Use and Land Cover Changes. A Case Study in Northern Greece
- 160 Sara Odelius (2009): Analys av stadsluftens kvalitet med hjälp av geografiska informationssystem.
- 161 Carl Bergman (2009): En undersökning av samband mellan förändringar i fenologi och temperatur 1982-2005 med hjälp av GIMMS datasetet och klimatdata från SMHI.
- 162 Per Ola Olsson (2009): Digitala höjddata och höjdsystem. Insamling av höjddata med fokus på flygburen laserskanning.
- 163 Johanna Engström (2009): Landskapets påverkan på vinden -sett ur ett vindkraftperspektiv.
- 164 Andrea Johansson (2009): Olika våtmarkstypers påverkan på CH₄, N₂O och CO₂ utsläpp, och upptag av N₂.
- 165 Linn Elmlund (2009): The Threat of Climate Change to Coral Reefs
- 166 Hanna Forssman (2009): Avsmältningen av isen på Arktis - mätmetoder, orsaker och effekter.
- 167 Julia Olsson (2009): Alpina trädgränsens förändring i Jämtlands- och Dalarnas län över 100 år.
- 168 Helen Thorstensson (2009): Relating soil properties to biomass consumption and land management in semiarid Sudan – A Minor Field Study in North Kordofan
- 169 Nina Cerić och Sanna Elgh Dalgren (2009): Kustöversvämningar och GIS - en studie om Skånska kustnära kommuners arbete samt interpolationsmetodens betydelse av höjddata vid översvämningssimulering.
- 170 Mats Carlsson (2009): Aerosolers påverkan på klimatet.
- 171 Elise Palm (2009): Övervakning av gåsbete av vass – en metodutveckling
- 172 Sophie Rychlik (2009): Relating interannual variability of atmospheric CH₄ growth rate to large-scale CH₄ emissions from northern wetlands
- 173 Per-Olof Seiron and Hanna Friman (2009): The Effects of Climate Induced Sea Level Rise on the Coastal Areas in the Hambantota District, Sri Lanka - A geographical study of Hambantota and an identification of vulnerable

- ecosystems and land use along the coast.
- 174 Norbert Pirk (2009): Methane Emission Peaks from Permafrost Environments: Using Ultra-Wideband Spectroscopy, Sub-Surface Pressure Sensing and Finite Element Solving as Means of their Exploration
- 175 Hongxiao Jin (2010): Drivers of Global Wildfires — Statistical analyses
- 176 Emma Cederlund (2010): Dalby Söderskog – Den historiska utvecklingen
- 177 Lina Glad (2010): En förändringsstudie av Ivösjöns strandlinje
- 178 Erika Filppa (2010): Utsläpp till luft från ballastproduktionen år 2008
- 179 Karolina Jacobsson (2010): Havsisens avsmältning i Arktis och dess effekter
- 180 Mattias Spångmyr (2010): Global of effects of albedo change due to urbanization
- 181 Emmelie Johansson & Towe Andersson (2010): Ekologiskt jordbruk - ett sätt att minska övergödningen och bevara den biologiska mångfalden?
- 182 Åsa Cornander (2010): Stigande havsnivåer och dess effect på känsligt belägna bosättningar
- 183 Linda Adamsson (2010): Landskapsekologisk undersökning av ädellövskogen i Östra Vätterbranterna
- 184 Ylva Persson (2010): Markfuktighetens påverkan på granens tillväxt i Guvarp
- 185 Boel Hedgren (2010): Den arktiska permafrostens degradering och metangasutsläpp
- 186 Joakim Lindblad & Johan Lindenbaum (2010): GIS-baserad kartläggning av sambandet mellan pesticidförekomster i grundvatten och markegenskaper
- 187 Oscar Dagerskog (2010): Baösbergsgrottan – Historiska tillbakablickar och en lokalklimatologisk undersökning
- 188 Mikael Månsson (2010): Webbaserad GIS-klient för hantering av geologisk information
- 189 Lina Eklund (2010): Accessibility to health services in the West Bank, occupied Palestinian Territory.