

Uheldsmodeller for veje i åbent land

Forprojekt



Søren Underlien Jensen

8. april 2011

Indhold

Sammenfatning.....	5
1. Indledning.....	11
1.1 Formål med forprojektet	13
1.2 Rapportens opbygning	13
2. Anvendelse af uheldsmodeller.....	15
2.1 Udpegning af sorte pletter.....	17
2.2 Valg af vej- og krydstype ved anlæg af nye veje	25
2.3 Vurdering af forventet uheldsbesparselse	31
2.4 Før-efter uheldsevalueringer	33
2.5 Opsummering	34
3. Uheldsmodeller, et indblik	37
3.1 Nuværende danske uheldsmodeller	37
3.2 State-of-the-art uheldsmodeller	42
3.2.1 Variable	43
3.2.2 Funktionsudtryk	46
3.2.3 Residualer og empirical Bayes	50
3.2.4 Estimering, optimering og teststatistik.....	56
3.2.5 Alvorlighed af uheld og personskader	59
3.2.6 Forudsætninger	66
3.3 Opsummering	71
4. Datagrundlag.....	75
5. Forslag til nye uheldsmodeller	81
5.1 Ønskværdige indledende analyser	81
5.2 Initierende vej- og krydstypeinddeling	82
5.3 Procedure for modeludarbejdelse	85
6. Analyser af designelementer	89
6.1 Relevante analyser	91
6.2 Potentielle analysemetoder	93
6.2.1 Tværsnitsstudier.....	93
6.2.2 Case-control studier	97
6.2.3 Bi- og multivariate modeller	99
6.3 Opsummering	100
Litteraturliste	101

Sammenfatning

Vejdirektoratet har bedt Trafitec om at belyse, hvordan fremtidige uheldsmodeller kan opstilles for kryds og strækninger i det åbne land. Modellerne skal anvendes til at opgøre et forventet antal uheld, personskader samt uheldsomkostninger for vej- og krydstyper, der er nævnt i *Vejregler for åbent land* i det omfang, som er muligt. Med et **forventet antal uheld** menes det antal af uheld, der i gennemsnit kan forventes at forekomme på en vej eller i et kryds af en given type, ved givne trafikmængder og indenfor en given tidsperiode.

I en række situationer er det også relevant med en opgørelse af estimerede uheld, personskader og uheldsomkostninger for eksisterende veje og kryds. Dette kan eksempelvis anvendes til udpegning af sorte pletter og vurderinger af forventede uheldsbesparelser af trafiksikkerhedstiltag. Med et **estimeret antal uheld** menes det antal uheld, der ville forekomme et givet sted og i en given periode, når man tager højde for den tilfældige variation i de rapporterede uheld.

Der findes mange forskellige typer af uheldsmodeller, men nærværende rapport fokuserer på modeller, der kan beregne et **forventet** og evt. et **estimeret** antal uheld og personskader, som angivet ovenfor. Sådanne uheldsmodeller kan bl.a. anvendes i arbejdet med:

- Udpegning af sorte pletter
- Valg af vej- og krydstype ved anlæg af nye veje
- Vurdering af forventet uheldsbesparelse
- Før-efter uheldsevalueringer

To sæt af uheldsmodeller må opstilles, hvis uheldsmodeller skal anvendes rimeligt korrekt indenfor de fire nævnte områder. Det ene sæt af modeller er i rapporten kaldt for **sortpletmodeller**. Disse anvendes til at opgøre forventede og estimerede uheld og personskader for et bredt spektrum af steder, f.eks. i forbindelse med sortpletudpegning. De kan også anvendes til opgørelse af estimerede uheld og personskader i forbindelse med vurdering af forventet uheldsbesparelse. Tillige anvendes de til opgørelse af estimerede uheld og personskader samt angivelse af generel udvikling i trafiksikkerhed i forbindelse med før-efter uheldsevalueringer. Sortpletmodeller er oftest basismodeller eller faktormodeller, der typisk ser ud på følgende måde:

$$\begin{array}{ll} \text{Basismodeller:} & \text{Strækning: } UHT = a \times N^P \\ & \text{Kryds: } UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2} \end{array}$$

$$\begin{array}{ll} \text{Faktormodeller:} & \text{Strækning: } UHT = a \times N^P \times e^{\sum_{i=1}^n b_i x_i} \\ & \text{Kryds: } UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2} \times e^{\sum_{i=1}^n b_i x_i} , \end{array}$$

hvor UHT er uheldstætheden, N er ÅDT på strækningen, N_{pri} er ÅDT ind i krydset fra primærvejen, N_{sek} er ÅDT ind i krydset på sekundærvejen, variablene x_i beskriver stedets udformning, samt a , P , P_1 , P_2 og b_i er konstanter.

Sortpletmodeller er derimod ikke egnede til at blive anvendt til beregning af forventede uheld og personskader i forbindelse med valg af vej- og krydstype ved anlæg af nye veje samt ved omfattende ombygninger af eksisterende veje og kryds. Det skyldes to forhold. For det første er sortpletmodeller sjældent præcise nok med hensyn variable for trafik og fysisk udformning til at kunne beskrive den aktuelle variant af vej- og krydstype med hensyn til antal uheld og personskader. For det andet er variable i (og udeladte variable for) sortpletmodeller som regel behæftet med så stor samvariation og endogenitet, at de estimerede konstanter i sortpletmodeller ikke er udtryk for kausale (årsag-virkning) sammenhænge. Sortpletmodeller kan derfor ikke anvendes til "alt-andet-lige" betragtninger, da alt andet ikke er lige i disse modeller. Hvis man ved anlæg af et nyt kryds stod og skulle vælge, om krydset skulle udformes med venstresvingsspor eller ej, så ville sortpletmodeller ikke kunne anvendes til at informere, om de sikkerhedsmæssige konsekvenser ved at vælge det ene alternativ frem for det andet.

Det andet sæt modeller er benævnt **grundmodeller**. Disse modeller kan anvendes sammen med tilhørende sikkerhedsfaktorer (**SF**'er) til opgørelse af forventede uheld og personskader i forbindelse med valg af vej- og krydstype ved anlæg af nye veje samt ved omfattende ombygninger af eksisterende veje og kryds. En grundmodel er en model, der angiver den eller de "rene" (kausale) sammenhænge mellem uheldstæthed og trafikmængde. Grundmodellen gælder for en specificeret variant af en vej- eller krydstype. Modellen er alene baseret på observationer, som er steder af den pågældende variant. En eller flere SF'er bruges til at opgøre forventede uheld og personskader for varianter, der adskiller sig lidt eller noget fra den variant, som grundmodellen er baseret på. En SF baseres på pålidelige studier, f.eks. gode før-efter uheldsevalueringer eller veldokumenterede med-uden studier. Kombinationen af grundmodel og SF'er vil typisk se således ud:

$$\begin{aligned} \text{Grundmodel + SF'er: Strækning: } UHT &= a \times N^P \times SF_1 \times SF_2 \times \dots \times SF_i \\ \text{Kryds: } UHT &= a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2} \times SF_1 \times SF_2 \times \dots \times SF_i \end{aligned}$$

Hvis en SF er 0,8, så svarer det til, at sikkerheden forbedres med 20 procent. Det kunne eksempelvis være en SF for tilstedeværelsen af vejbelysning, hvis grundmodellen er baseret på en variant uden vejbelysning.

Det er i dag uklart, hvor meget et alternativ må adskille sig fra strækninger eller kryds, der indgik i grundmodellen, før en forventet uheldstæthed ikke længere er troværdig.

Kombinationen af grundmodeller og SF'er kan derimod ikke anbefales at blive anvendt til sortpletudpegning. Dels er det i praksis uhensigtsmæssigt, da det vil kræve mange relativt detaljerede oplysninger om eksisterende steder. Dels vil

grundmodeller og SF'er ikke pålideligt kunne anvendes til at beregne et estimeret antal af uheld og personskader. Man er således nødt til at opstille to sæt uheldsmodeller, hvis modeller skal bruges på de fire førnævnte anvendelsesområder.

Nuværende danske uheldsmodeller

Vejdirektoratets seneste uheldsmodeller er sortpletmodeller udformet som basismodeller. De er af ældre dato og trænger til at blive opdateret både med hensyn til modelspecifikation og data. Disse uheldsmodeller kan ikke anvendes til valg af vej- og krydstype, og kan ikke beskrive antallet og alvorligheden af personskader. De kan heller ikke anvendes til at opgøre det estimerede antal uheld, da den valgte fordeling af residualer (Poisson) ikke muliggør brug af empirical Bayes, som anvendes til at beregne det estimerede antal uheld.

State-of-the-art uheldsmodeller

Gennemgangen af state-of-the-art uheldsmodeller tyder på følgende:

- Det er hensigtsmæssigt at undersøge fordelingen af uheld eller personskader før en model opstilles, hvilket i hovedtræk vil sige middelværdi og varians.
- Modeller for tætheden af uheld og personskader kan i næsten alle tilfælde opstilles med baggrund i en Negativ binomial (Poisson-gamma) fordeling. Det specielle ved en Negativ binomial model (NB model) er, at modellens middelværdi (det forventede antal uheld) er lavere end variansen. Det bør dog altid undersøges, hvilken fordeling der er rimelig at antage. Det kan eksempelvis gøres for en foreløbig basismodel og en senere færdig model.
- NB modeller kan opstilles med en fast eller varierende spredningsparameter. Spredningsparameteren beregnes ud fra modellens middelværdi og varians. Med varierende spredningsparameter menes, at den afhænger af en eller flere af modellens uafhængige variable. En varierende spredningsparameter kan være en stor fordel. Det er derfor vigtigt at undersøge, hvordan spredningsparameteren bedst kan beskrives.
- Estimeringen af relativt simple NB modeller kan fint gøres med teknikken med maximum-likelihood.
- Udvalget af potentielle uafhængige variable er oftest den forhåndenværende datatilgængelighed. Det er meget sjældent, at særlige registreringer udføres for at opstille uheldsmodeller. Blandt potentielle uafhængige variable udvælges modellers uafhængige variable oftest med baggrund i statistiske test, dog er det ikke ualmindeligt, at trafikmængder indgår selvom disse ikke opfylder kravene i statistiske test. Det anbefales at benytte signifikansniveauet for variablens sammenhæng med uheldstætheden samt AIC (ved modeller uden årsfaktorer eller HGLM modeller med årsfaktorer) eller QIC (ved GEE modeller med årsfaktorer) for modellen til at vælge, om den uafhængige variabel skal indgå i modellen. Det anbefales at optage variable trinvis i modellen efter signifikansniveau. (AIC er Akaike's Information Criterion, HGLM er Hierarchi-

cal Generalized Linear Models, QIC er Quasi-likelihood Information Criterion, og GEE er Generalized Estimating Equation)

- Brug af årsfaktorer kan være en fordel, hvis der er tilstrækkeligt med data for hvert år (eller anden periodelængde f.eks. to eller tre år). Områdefaktorer kan også være en fordel, dog er brug af områdefaktorer ikke velundersøgt.
- Variable for trafik (eksponering) kan med fordel opstilles så detaljerede som muligt f.eks. opdelt på trafikstrømme og trafikarter, i stedet for overordnede tal for årsdøgntrafik.
- Det anbefales at undersøge flere funktionsudtryk, og valget af funktionsudtryk bør være understøttet af statistiske tests. Simple udtryk, f.eks. uden synergi- og additive led, bør foretrækkes frem for komplekse udtryk, hvis forskelle i AIC eller QIC er beskedne.
- Modeller for personskader opstilles bedst for hver skadesgrad (NB modeller), hvis der er tilstrækkeligt med data til at gøre det. Alternativt må hhv. en model for personskadetæthed (NB model) samt en model for fordeling af skader efter skadesgrad (ordinal probit eller logit) opstilles.
- Samvariation mellem modellens uafhængige variable bør beskrives, og der bør tages stilling til, om det er nødvendigt at håndtere stærk samvariation.
- Det anbefales at gøre brug af empirical Bayes til opgørelse af estimerede uheld og personskader.
- Det anbefales, at modellens forklaringskraft udtrykkes ved Elvik's indeks eller Freeman-Tukey R^2 .
- Konfidensinterval og standardafvigelse for estimerede konstanter i modellen bør angives.
- Modellens gyldighedsområde bør beskrives.
- Det forekommer ikke at være velundersøgt, hvordan kryds' udbredelse bedst defineres. Det forekommer samtidigt, at den oftest benyttede basismodel for kryds ($UHT = a \times N_{pri}^{P_1} \times N_{sek}^{P_2}$) er diskutabel.

Fremtidige uheldsmodeller for veje og kryds i det åbne land

Det tilgængelige datagrundlag for fremtidige uheldsmodeller er undersøgt. Dog er ikke alle relevante variable for trafik og fysisk udformning studeret. Før arbejdet med at udvikle uheldsmodeller iværksættes er det ønskværdigt at gennemføre et antal analyser. De analyser skal lede frem til klare definitioner af kryds, ramper og flettestrækninger i relation til uheldsmodellering. Analyser for minimumskrav til trafikmængder for steder, der skal indgå i uheldsmodeller, er også ønskelige.

Ud fra oplysningerne om tilgængelige data er der givet et bud på en initierende vej- og krydstypeinddeling for fremtidige uheldsmodeller. Dette bud er opdelt i hhv. sortpletmodeller og grundmodeller. For sortpletmodellerne gælder desuden, at den foreslåede inddeling for modeller for personskader ikke er samme som for modeller for uheld.

Forslag til initierende vej- og krydstypeinddeling for sortpletmodeller til beregning af forventede og estimerede **uheld**stætheder i det åbne land:

Strækninger:	Kryds:
Motorveje	Kryds med 3 ben/grene
Flettestrækninger-ramper ved motorveje	Kryds med 4 ben/grene
2-sporede veje med/uden midterrabat	Prioriterede T-kryds
3-4-sporede veje med/uden midterrabat	Forsatte prioriterede T-kryds
	Prioriterede F-kryds
	Rundkørsler
	Signalregulerede T-kryds
	Signalregulerede F-kryds

Der foreslås i alt fire uheldsmodeller for strækninger og otte modeller for kryds. For 2-sporede veje kan det dog vise sig, at flere modeller kan udarbejdes. Det kan vise sig, at spredningsparameteren i modeller for kryds med hhv. 3 og 4 ben/grene er for stor til, at modellerne kan siges at være hensigtsmæssige. Det kan evt. være en fordel at sammenlægge rundkørsler i åbent land med rundkørsler i byområder. En sådan løsning kan evt. også være som en fordel med signalregulerede T-kryds.

Forslag til initierende vej- og krydstypeinddeling for sortpletmodeller til beregning af forventede og estimerede **personskade**tætheder i det åbne land:

Strækninger:	Kryds:
Dræbte – alle strækninger	Dræbte – alle kryds
Alvorlige skader – motorveje	Alvorlige skader – alle kryds
Lette skader – motorveje	Lette skader – alle kryds
Alvorlige skader – andre vejstrækninger	
Lette skader – andre vejstrækninger	

Der foreslås fem personskademodeller for strækninger og tre modeller for kryds. Det må komme an på en prøve, hvorvidt det er hensigtsmæssigt at lade ramper og flettestrækninger indgå i kategorien motorveje. Med andre vejstrækninger menes øvrige veje i det åbne land og motortrafikveje.

Der foreslås tre sæt grundmodeller for strækninger i det åbne land. Med ét sæt grundmodeller menes én model for uheldstæthed, én model for personskadetæthed samt én model for fordeling af skader efter skadesgrad. De tre sæt er:

- *Motorvejsstrækninger* med 4 kørespor og nødspor, 110/130 km/t hastighedsbegrænsning, ingen vejbelysning, ingen afstandsmærker, kørespor ca. 3,5 meter brede, ...
- *Højklasset 2-sporede veje* uden midterrabat, med midtlinje og kantpæle, med minimum x,x meter brede kantbaner, uden ind- og udkørsler, horisontalkurver

skal minimum have radius af xxxx meter, 80/90 km/t hastighedsbegrænsning, ingen vejbelysning, ingen overhalingsforbud, kørespor 3,2-4,0 meter brede, ...

- *Lavklasset 2-sporede veje* uden midterrabat, uden kantbaner, mellem 0 og x ind- og udkørsler pr. km, ingen vejbelysning, horisontalkurver med minimum radius af xxx meter, 60/70/80 km/t hastighedsbegrænsning, ingen vejbelysning, kørespor 2,5-3,5 meter brede, ingen overhalingsforbud, ...

Disse grundmodeller har trafikmængde (ÅDT) som uafhængig variabel. Det kan dog være relevant også at benytte f.eks. hastighedsbegrænsning, bredde af kørespor og kantbaner, forekomst af midtlinje samt antal ind-/udkørsler som andre uafhængige variable – for at opnå et tilstrækkeligt antal observationer.

Det er uvist, om det er muligt at opstille grundmodeller for kryds. Problemet er, at antallet af uheld og personskader for velspecificerede varianter af kryds er ukendt. Det kan dog være relevant at analysere følgende krydstyper, hvoriblandt mere velspecificerede varianter kan identificeres:

- Prioriteret T-kryds uden kanalisering
- Prioriteret T-kryds med kanalisering på både primær- og sekundærvej
- Prioriteret F-kryds med kanalisering på både primær- og sekundærvej
- Signalreguleret T-kryds
- Signalreguleret F-kryds
- Rundkørsel med 1 cirkulationsspor og 4 vejgrene

Egentlige krydstællinger af trafikstrømme i kryds anses for nødvendige til opstilling af relevante grundmodeller for F-kryds og rundkørsler. I T-kryds kan trafikstrømme beregnes ud fra tællinger på de tre strækninger før krydset. For alle fem nævnte krydstyper kan det være relevant at sammenlægge kryds i åbent land med kryds i byområder, og benytte en variabel for zone.

Der er i afsnit 5.3 udarbejdet en overordnet procedure for, hvordan disse sortpletmodeller og grundmodeller kan udvikles.

Analyser af designelementer

Der er fortsat behov for at udføre med-uden studier (analyser af designelementer) for at opgøre sikkerhedseffekter (SF'er) for visse vej- og trafiktekniske tiltag. En liste med designelementer i kryds og på strækninger, der kunne være relevante at undersøge i med-uden studier, er opstillet. Tre metodisk forskellige måder at udføre med-uden studier er gennemgået. Det er hhv. tværsnitstudier, case-control studier samt bi- og multivariate modeller. Set i lyset af listen af designelementer, der kunne være relevante at undersøge, forekommer det, at tværsnitstudier og case-control metoden er relevante at anvende. Det anbefales at gøre brug af en datafangst med mange krav til variable for trafik og fysisk udformning, frem for en metodik med komplekse korrektioner for skævheder, herunder samvariation og endogenitet blandt uafhængige variable.

1. Indledning

Uheldsmodeller har været anvendt af danske vejbestyrelser til sortpletudpegning og andre formål i over 20 år. De danske uheldsmodeller angiver sammenhænge mellem uheldstæthed (**UHT**) pr. år pr. km vej eller pr. kryds og årsdøgntrafik (**ÅDT**) på vejstrækningen eller ind i krydset. Uheldsmodellernes konstanter benævnes ofte a og p. I Danmark anvendes uheldsmodeller, der internationalt kaldes for basismodeller, og er potensmodeller. Disse modeller har følgende former:

$$\text{Strækning: } UHT = a \times N^P \quad (1)$$

$$\text{Kryds: } UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2}, \quad (2)$$

hvor UHT er uheldstætheden, N er ÅDT på vejen i begge retninger, N_{pri} er ÅDT ind i krydset fra primærvejen, N_{sek} er ÅDT ind i krydset på sekundærvejen, og a, P, P_1 og P_2 er konstanter. Primærvejen i krydset er de to vejben uden vigepligt i vigepligtsregulerede kryds, eller de to vejben i lyskryds, der har den største trafik.

Med kommunalreformen blev mulighederne for at opstille uheldsmodeller ændret, da vej- og trafikdata for de fleste amtsveje, der overgik til kommunerne, udgik af Vejman.dk. Der er ikke opstillet uheldsmodeller og udgivet a/p-værdier siden da. De seneste modeller bygger på uheldsdata fra 2001-2005, hvilket er ved at blive et problem, idet modellerne ændrer sig med tiden. Der er brug for nye a/p-værdier, hvis f.eks. bekæmpelse af sorte pletter fortsat skal være en aktivitet i det danske trafiksikkerhedsarbejde.

En anden problemstilling er vej- og krydstypeinddelinger (a/p-typer). Her er der to ønsker. I relation til sortpletudpegning ønskes, at a/p-typerne matcher eksisterende uheldsbelastede steder bedst muligt. Derfor er det ret uinteressant med a/p-værdier for f.eks. 2+1 veje, da der kun findes få veje af denne slags. Og derfor er det også mindre vigtigt med a/p-værdier for svagt trafikerede veje, da der kun sker få uheld på de veje. I relation til valg af vej- og krydstyper i trafikplanlægning af nye veje og større ombygninger af veje og vejnet har man derimod brug for a/p-typer, der kan beskrive uheldstætheden for de fremtidige veje. Og så er det vigtigt med mere pålidelige vurderinger af forventede uheldsbespareser af trafiksikkerhedstiltag og andre vej- og signaltekniske projekter.

Vejdirektoratet har bedt Trafitec om at belyse, hvordan fremtidige uheldsmodeller kan opstilles for kryds og strækninger i det åbne land. Modellerne skal anvendes til at opgøre et forventet antal uheld, personskader samt uheldsomkostninger for vej- og krydstyper, der er nævnt i *Vejregler for åbent land* i det omfang, som er muligt. Med et **forventet antal uheld** menes det antal af uheld, der i gennemsnit kan forventes at forekomme på en vej eller i et kryds af en given type, ved givne trafikmængder og indenfor en given tidsperiode. Det er også relevant at udarbejde modeller for en række eksisterende vej- og krydstyper i det åbne land, der ikke er

nævnt i vejreglerne. I Vejregler for åbent land findes 9 overordnede vejtyper og 9 overordnede krydstyper, se tabel 1.

Vejtyper	Krydstyper
Motorvej 4-sporet	Prioriterede (stop- eller vigepligtsreguleret) T-kryds
Motorvej 6-sporet	Forsatte prioriterede T-kryds
1-sporet vej	1-sporet rundkørsel
2-sporet vej	2-sporet rundkørsel
2-sporet vej med midterrabat/-areal	Signalreguleret T-kryds
2+1 vej med midterrabat/-areal	Signalreguleret F-kryds
2+1 vej	2-planskryds, forbindelsesanlæg
4-sporet vej med midterrabat	2-planskryds, tilslutningsanlæg
6-sporet vej med midterrabat	2-planskryds, hanske

Tabel 1. Overordnede vej- og krydstyper for nyanlæg i Vejregler for åbent land.

Af tabel 1 kan erfares i alt 18 vej- og krydstyper. I Vejreglerne er samtidig nævnt en række varianter af disse vej- og krydstyper. Det kan være relevant, at fremtidige uheldsmodeller kan opgøre et forventet antal uheld for varianterne. Af eksisterende vej- og krydstyper, der ikke er angivet i tabel 1, kan nævnes motorvej med 8-10 spor, 4-sporet vej uden midterrabat og prioriterede F-kryds.

For tiden anlægges især nye 2-sporede veje og 4-sporede motorveje i det åbne land. Ved nyanlæg udføres især prioriterede T-kryds, 2-planskryds og rundkørsler i det åbne land. Af omkostningstunge ombygninger udføres tillige en del f.eks. bundne venstresving i signalregulerede kryds, kanaliseringer af kryds, shunts i rundkørsler, midterrabatter, vejbelysning, cykelstier og vejudvidelser. Af billige ombygninger kan nævnes afmærkninger såsom midt- og kantlinjer, rumleriller, afstandsmærker og forvarsling af kryds.

I den ideelle verden skulle et sæt af uheldsmodeller gerne muliggøre:

- Opgørelse af forventede uheld, personskader og uheldsomkostninger for alle relevante typer og varianter af vejstrækninger og kryds til brug for f.eks. valg af vej- og krydstype i forbindelse med nyanlæg og udpegning af sorte pletter.
- Opgørelse af estimerede uheld, personskader og uheldsomkostninger for alle eksisterende vejstrækninger og kryds. Dette kan eksempelvis anvendes til udpegning af sorte pletter, vurderinger af forventede uheldsbesparelser, og før-efter uheldsevalueringer. Med et **estimeret antal uheld** menes det antal uheld, der ville forekomme et givet sted og i en given periode, når man tager højde for den tilfældige variation i de rapporterede uheld.

1.1 Formål med forprojektet

Forprojektets hovedformål er at opstille et forslag til en ramme for at udarbejde uheldsmodeller for veje og kryds i åbent land. Med ramme menes, at der opstilles forslag til vej- og krydstypeinddelinger samt metodiske procedurer for at udvikle uheldsmodeller for disse vej- og krydstyper.

Udover en ramme for udarbejdelse af uheldsmodeller skal forprojektet beskrive, hvordan trafiksikkerhedsmæssige analyser af designelementer kan gennemføres. Med analyser af designelementer menes, at der ud fra sammenligninger af forskellige veje eller kryds kan opgøres, hvordan de forskellige designelementer påvirker trafiksikkerheden, f.eks. sammenhænge mellem trafiksikkerhed og midterøens diameter i rundkørsler eller sammenhænge mellem køresporets bredde og trafiksikkerhed. Resultater fra analyserne kan anvendes til at forfine uheldsmodeller og til beregninger af forventede uheldsbesparselser af vejtekniske ændringer.

I forprojektet belyses tillige forskellige metodiske muligheder for at udarbejde uheldsmodeller og analyser af designelementer. Derfor har forprojektet også til formål at belyse forskellige opbygninger af uheldsmodeller, herunder formeludtryk og modelkomplekser, indgangsværdier og forudsætninger, vej- og krydstypeinddelinger, måder at inddrage uheldsalvorlighed samt statistiske vurderinger af modellers pålidelighed, gyldighedsområde og usikkerhed. Det er også væsentligt at belyse, hvordan analyser af designelementer kan give troværdige resultater.

Forslaget til en sådan ramme skal også ses i lyset af, hvad modeller og resultater kan anvendes til samt i hvilke sammenhænge de anvendes. Endelig er det vigtigt, at denne ramme tager højde for, hvilke data der findes eller kan forventes at blive tilgængelige på længere sigt.

1.2 Rapportens opbygning

Rapporten er bygget op omkring seks kapitler. Fokus i alle kapitler er på uheldsmodeller, der kan angive forventede og estimerede antal uheld og personskader i en given periode på en vej eller i et kryds af en given udformning og med en given trafikmængde.

Kapitel 2 *Anvendelse af uheldsmodeller* giver et indblik i hvad uheldsmodeller kan bruges til. Der ses også på hovedtrækkene i den hidtidige anvendelse af uheldsmodeller i Danmark, og de problemstillinger, der er forbundet hermed.

Kapitel 3 *Uheldsmodeller – et indblik* beskriver detaljeret tidligere og nuværende uheldsmodeller i Danmark. Derudover findes en beskrivelse af, hvordan uheldsmodeller på det højeste internationale niveau (state-of-the-art) kan udarbejdes. Fokus her er alene på uheldsmodeller, der tager udgangspunkt i rapporterede uheld og personskader indtruffet på eksisterende veje og kryds. Kapitellet afsluttes

med beskrivelser af opbygningen af uheldsmodeller, der kan anvendes til valg af vej- og krydstype.

Kapitel 4 *Datagrundlag* angiver, hvad der findes af tilgængelige data om veje og kryds, trafik samt uheld og personskader på nuværende tidspunkt. Der angives tillige sandsynlige udviklinger i tilgængelige data i årene fremover.

Kapitel 5 *Forslag til nye uheldsmodeller* angiver et forslag til proces for udarbejdelse af uheldsmodeller med udgangspunkt i rapporterede uheld og personskader indtruffet på eksisterende veje og kryds. I forslaget indgår tillige en initierende vej- og krydstypeinddeling.

Kapitel 6 *Analyser af designelementer* beskriver, hvordan den sikkerhedsmæssige betydning af vej- og krydsdesign kan opgøres ud fra faktormodeller, case-control studier samt bi- og multivariate uheldsmodeller. Vanskeligheder og usikkerheder, der er forbundet med sådanne studier og modeller, angives også. Der ses ikke på før-efter uheldsevalueringer. Kapitlet indeholder også forslag til en konkret serie af analyser af designelementer, samt hvordan analyserne i praksis kunne udføres.

2. Anvendelse af uheldsmodeller

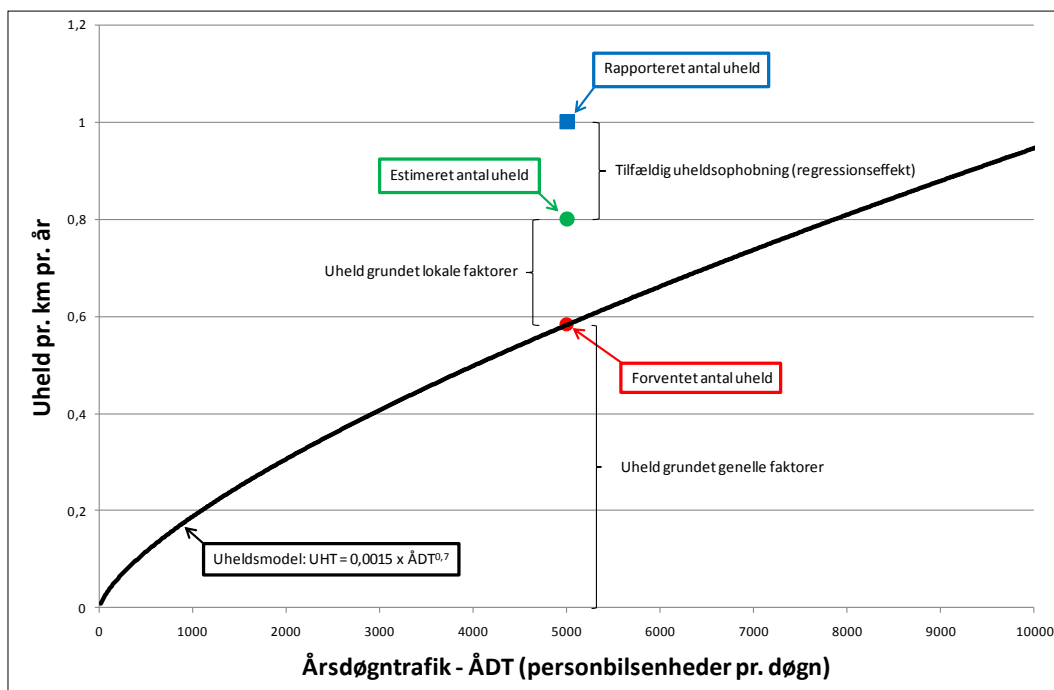
Kapitlet omhandler uheldsmodeller, som anvendes til at beregne hhv. et **forventet** og et **estimeret** antal uheld og personskader. Det er særdeles vigtigt at forstå disse to begreber og forskellen på dem:

- Med et **forventet** antal uheld menes det antal uheld, der normalt og i gennemsnit blev rapporteret for strækninger eller kryds af en given udformning med en given trafikmængde og i en given periode. Et stort antal af strækninger eller kryds anvendes til at angive, hvad der er normalt og gennemsnitligt.
- Med et **estimeret** antal uheld menes det antal uheld, der for et givet sted – et helt konkret sted – skete i en given periode, når der er taget højde for den tilfældige variation (regressionseffekt) i de rapporterede uheld. Det vil sige det antal uheld, der ville være sket, hvis man betragtede en særdeles lang periode, hvor betydningen af regressionseffekten er minimal. Det estimerede antal uheld opgøres ud fra det rapporterede antal uheld for det konkrete sted og det forventede antal uheld for et sådant sted, samt en opdeling i hhv. systematisk og tilfældig variation i uheld for de steder, der indgik i uheldsmodellen, som blev anvendt til at beregne det forventede antal uheld. Det er den såkaldte Empirical Bayes (EB) metode, der benyttes. Man ser altså ikke på en særdeles lang periode for det konkrete sted, men opgør hvordan den systematiske og den tilfældige variation er på lignende steder. Årsagen er, at hovedparten af de steder, hvor man vil beregne et estimeret antal uheld, jævnligt ombygges – og man kan derfor ikke betragte en særdeles lang periode af det konkrete sted med én given udformning.

Regressionseffekten er forskellen mellem et rapporteret antal uheld i en relativ kort periode og det ”reelle” uheldsniveau, hvis man kunne opgøre dette niveau ud fra en uendelig lang periode. Det er tilfældigheden i uheldsforekomsten, der er baggrunden for, at regressionseffekten overhovedet forekommer. Man kan sige, at regressionseffekten er udslaget af en skæv stikprøve. Regressionseffekten kan være positiv eller negativ, altså kan der være rapporteret flere eller færre uheld end forventet. Regressionseffekten er detaljeret beskrevet af Hauer (1997).

Hauer (1997) præsenterede EB metoden, der kan angive et estimat for regressionseffekt for et enkelt kryds eller en strækning ud fra det rapporterede antal uheld samt en uheldsmodel, der passer med den aktuelle type og variant af krydset eller strækningen. Efterfølgende er EB blevet afprøvet, testet, osv. I forskningskredse opfattes EB som state-of-the-art eller den bedste metode til at tage højde for regressionseffekt på et givet sted (se f.eks. *Vistisen, 2002; Madsen, 2005; Elvik, 2007; Persaud og Lyon, 2007*).

For at få en bedre forståelse af de to begreber, forventet og estimeret antal uheld, er de illustreret grafisk i figur 1. Det forventede antal uheld er således de uheld, der normalt ville ske og skyldes generelle faktorer. Det forventede antal uheld er i figur 1 angivet med en tyk sort linje, der er tegnet på baggrund af uheldsmodellen med udtrykket: $UHT = 0,0015 \times \text{ÅDT}^{0,7}$. For et konkret sted med 5.000 biler pr. døgn er det forventede antal uheld således 0,58 uheld pr. km pr. år og angivet med en rød prik.



Figur 1. Uheld for en strækning opdelt i uheld grundet hhv. lokale og generelle faktorer samt tilfældig uheldsophobning (regressionseffekt).

I figur 1 er med blå plet angivet antallet af rapporterede uheld (1 uheld pr. km pr. år). I Danmark har det været almindeligt at udpege sorte pletter ud fra forskellen mellem antallet af forventede og rapporterede uheld samt at vurdere forventede uheldsbesparelser af trafikssikkerhedstiltag alene ud fra rapporterede uheld. Det må siges at være uhensigtsmæssigt, da man ikke tager højde for regressionseffekten. Når der ikke tages højde for regressionseffekten, udpeges der flere falske sorte pletter, og uheldsbesparelser fejlvurderes ofte – typisk overvurderes de.

EB metoden gør det muligt at estimere det mest sandsynlige antal uheld, der vil ske det konkrete sted på lang sigt. Det estimerede antal uheld ligger altid mellem det forventede og det rapporterede antal uheld. Derved opdeles forskellen mellem forventet og rapporteret antal uheld i to klumper – dels uheld grundet lokale faktorer og dels en tilfældig variation i uheldsforekomsten (regressionseffekt). Når EB benyttes til udpegning af sorte pletter, så siges, at en sort plet er, hvor antallet af uheld grundet lokale faktorer er så stort, at der er en stor chance for at identificere et tiltag, der kan give en god uheldsforebyggende effekt. Lokale faktorer er egent-

ligt systematisk variation i uheldsforekomsten, som den opstillede uheldsmodel ikke beskriver.

Det er vigtigt at nævne, at det rapporterede antal uheld kan være lavere end det forventede antal uheld. I så fald bliver regressionseffekten og uheld grundet lokale faktorer til negative talstørrelser. Der sker altså færre uheld end forventet, fordi lokale faktorer gør stedet sikrere. Og regressionseffekten er negativ, da der tilfældigvis er sket færre uheld end hvad der ellers ville have gjort på lang sigt.

Opgørelsen af forventede og estimerede antal uheld kan anvendes til en række forskellige arbejdsprocesser indenfor trafiksikkerhedsområdet, f.eks.:

- Udpegning af sorte pletter
- Valg af vej- og krydstype ved anlæg af nye veje
- Vurdering af forventet uheldsbesparelse
- Før-efter uheldsevalueringer

2.1 Udpegning af sorte pletter

Et væsentligt anvendelsesområde for uheldsmodeller er udpegning af sorte pletter. Sorte pletter er typisk korte vejstrækninger eller kryds, hvor der er sket flere uheld end man kunne forvente. I nogle lande arbejdes også med lange strækninger (grå strækninger), hvor der er sket flere uheld end forventet (*Elvik, 2007*). I få lande bruges modeller, der kan opgøre, om der er sket flere personskader end forventet.

Sorte pletter er vigtige at udpege, fordi man finder frem til steder, hvor der ofte er rig mulighed for at forbedre trafiksikkerheden. En uheldsanalyse af en sort plet peger typisk på et eller flere specifikke uheldsproblemer, og efterfølgende peges på et eller flere tiltag, der kan afhjælpe disse problemer. Det er ikke altid, at man kan identificere et uheldsmønster med tilhørende uheldsproblemer og tiltag. Det betyder ikke nødvendigvis, at der er tale om en falsk sort plet, men at vej- eller krydsudformningen og trafikantadfærden fører til et diffust uheldsmønster.

Evalueringer tyder på, at den sikkerhedsmæssige effekt ikke er procentuelt større ved udbedringer af sorte pletter i forhold til udbedringer af andre uheldsbelastede lokaliteter, når udbedringen er ens – altså samme tiltag (*Elvik, 2007*). Forestil dig to prioriterede F-kryds, hvor kun det ene kryds er en sort plet. I begge kryds anlægges venstresvingsbaner. Man har her fundet, at sikkerhedseffekten opgjort i procent er nogenlunde ens i sådanne to kryds, når der tages højde for regressions-effekten. Men når effekten opgøres i det absolutte antal forebyggede uheld og personskader, så er det antal størst ved udbedring af sorte pletter, fordi det er dér, der sker flest uheld. Noget andet er, at mange uheldsanalyser i forbindelse med sortpletarbejdet leder frem til særlige lokale tiltag såsom fjernelse af sigthindrende genstande, uhensigtsmæssig afmærkning, for høj vejkant, faste genstande, osv.

Det er forskelligt, hvordan sorte pletter udpeges rundt om i verden. Nedenfor er gengivet et udsnit af de anvendte metoder:

- Uheldstæthed (uheld pr. år)
- Uheldsfrekvens (uheld pr. motorkøretøj)
- Uheldsomkostning el. lign. (uheldsomkostning pr. år eller motorkøretøj)
- Kritisk uheldsfrekvens (uheld pr. motorkøretøj er over "kritisk" niveau)
- Kritisk niveau af visse uheldssituationer (uheld pr. år eller motorkøretøj)
- Modelbaseret udpegning, rapporteret / forventet, ej EB baseret
- Modelbaseret udpegning, estimeret / forventet, EB baseret

Mange vejbestyrelser udpeger sorte pletter ud fra uheldstætheden – nogle gange opdelt på typer af strækninger og kryds. Et problem med den udpegningsmetode er, at man typisk udpeger de steder, hvor der er mest trafik, og de byder langt fra altid på gode muligheder for at forbedre trafiksikkerheden. Derved bruges mange ressourcer på uheldsanalyser, der ikke leder frem til udbedringer af de udpegede sorte pletter.

Andre vejbestyrelser har udpeget sorte pletter ud fra uheldsfrekvensen, igen oftest opdelt på typer af strækninger og kryds. Den metode er bedre end uheldstæthedsmetoden, men indebærer alligevel problemer. Her udpeges ofte steder med relativ lidt trafik og høj positiv regressionseffekt. Det skyldes, at uheldsfrekvensen ofte er stærkt afhængig af trafikmængden – jo mere trafik desto lavere uheldsfrekvens. Samtidig tages der ikke højde for regressionseffekten.

Ved at udpege efter uheldsomkostning eller en anden vægtningsmetode, der sætter de forskellige alvorlighedsgrader af uheld og personskader i forhold til hinanden, fås en vigtig dimension ind i udpegningen. Et særdeles stort problem med udpegninger, der tager hensyn til alvorlighed, har været, at dræbte eller dødsulykker vægtes meget højt, og derfor udpeges ofte blot steder med dræbte. Brugen af EB har delvist løst problemet, da personskadens alvorlighed er forbundet med en stor grad af tilfældighed. Men der er to uløste metodiske problemer, dels afhænger antallet af personskader pr. uheld af uheldets alvorlighed, og dels er alvorligheden af personskader i et uheld indbyrdes afhængige.

Ved udpegning efter kritisk uheldsfrekvens inddeles vejnettet i typer af strækninger og kryds, og hver type underinddeles i hensigtsmæssige intervaller for trafikmængde. For hvert interval opstilles så en kritisk uheldsfrekvens, som findes ud fra gennemsnittet og variansen i uheldsfrekvens for de kryds eller strækninger, der indgår i intervallet samt et velvalgt signifikansniveau. Denne metode har egentlig vist sig at fungere ganske godt, hvor man ad åre finder ud af, hvilke vej- og krydstyper og intervaller der leder frem til de mest effektive sortpletudbedringer. Et problem er, at der ikke tages højde for regressionseffekt, så der er en tendens til overvurdere effekten af udbedringerne.

Kritisk niveau af visse uheldssituationer har typisk været brugt til at udpege kryds eller strækninger med bestemte uheldsproblemer eller steder, hvor bestemte tiltag kunne være relevante at implementere. Afhængig af formålet med udpegningen har man typisk brugt antallet af uheld af visse uheldssituationer eller andelen, som visse uheldssituationer udgør af alle uheld, til at fastsætte det kritiske niveau. En fordel er, at udgifterne til uheldsanalyser efterfølgende reduceres, mens bagdelen er, at andre potentielle uheldsproblemer og tiltag overses. Sådanne udpegninger er udført med og uden tal for trafikmængder, og med og uden brug af EB.

I Danmark har Vejdirektoratet og de tidligere amter i mange år anvendt en modelbaseret udpegning, der ikke gør brug af EB og derfor udpeges sorte pletter ud fra forholdet mellem forventede og rapporterede uheld. Vejnettet inddeles i typer af strækninger og kryds, og sættet af modeller kan bruges til at beregne forventede antal af uheld for alle typer og ved alle trafikmængder. Et problem er, at der ikke tages højde for regressionseffekt, hvilket øger antallet af falske sorte pletter, og fører til overvurderede uheldsbesparelser.

Nogle få vejbestyrelser i udlandet udfører i dag modelbaseret udpegning af sorte pletter med brug af EB. Denne metode udpeger flest sande sorte pletter og peger således mest præcist på de steder, hvor der oftest er gode muligheder for at forbedre trafiksikkerheden. En fordel ved EB er, at regressionseffekten bestemmes for hver strækning og kryds, og derved gives god mulighed for at vurdere en uheldsbesparelse af en sortpletudbedring rimeligt korrekt. Således bliver både udpegningen af sorte pletter og prioriteringen af sortpletprojekter mere optimal.

Grundlæggende kan siges, at optimering af arbejdsprocessen med sorte pletter handler om at anvende så få ressourcer som muligt på udpegning, uheldsanalyser, prioriteringer, projektering og ombygninger i forhold til antallet af forebyggede uheld og personskader.

I Danmark udførte Vejdirektoratet og mange af de tidligere amter en modelbaseret udpegning af sorte pletter ved brug af edb-programmet VISplet (Sørensen, 2005). Vejnettet blev inddelt i et kryds- og strækningsregister med angivelse af type og variant af kryds og strækning samt trafikmængder. Uheldsmodeller blev opstillet for hver kryds- og strækningstype, altså basismodeller. Modellerne anvendes i VISplet til at beregne et forventet antal uheld i hvert kryds og på hver strækning i vejnettet. Det forventede antal uheld sammenholdes med et rapporteret antal uheld. Kryds og strækninger med signifikant flere uheld end forventet udpeges til sorte pletter. Metoden er ikke baseret på EB.

I amterne og Vejdirektoratet har man typisk udpeget kryds til sorte pletter, når der var rapporteret mindst 4 uheld over en femårig periode, og der skete flere uheld end forventet på 90 eller 95 % signifikansniveau (Sørensen, 2005). På strækninger anvendes en kompliceret fremgangsmåde, hvor der bruges en glider, som rykker sig fra uheld til uheld hen over en strækning. Længden af glideren afhænger af en række forhold, men er typisk 400-500 meter (Greibe og Hemdorff, 2001). Vejbe-

styrelser har typisk udpeget strækninger til sorte pletter, når der var rapporteret mindst 5 uheld over en femårig periode, og der skete flere uheld end forventet på et 99,9 % signifikansniveau (Sørensen, 2005). Udpegning af sorte pletter i Danmark har været karakteriseret ved:

- Vej- og krydsregistre med relevante vej- og trafikdata,
- Udpegning ud fra mange basismodeller for stræknings- og krydstyper,
- Basismodeller er opstillet med baggrund i en Poisson fordeling ved brug af NLIN procedure i statistik softwaren SAS, hvor iterativ (vægtet) mindste kvadraters metode anvendes.
- Sortpletudpegningen sker på baggrund af forskellen mellem et forventet og et rapporteret antal uheld,
- Glider metoden anvendes på strækninger,
- Minimumskrav til trafikmængder for veje i kryds,
- Minimumskrav på 4-5 uheld,
- Alvorlighed af uheld indgår ikke,
- Femårig udpegningsperiode, og
- Edb-baseret (automatiseret) udpegning.

Amterne og Vejdirektoratet i Danmark havde formentligt den mest avancerede udpegning af sorte pletter i Europa (Elvik, 2007). Men det ser ud til, at flere lande både i Europa og udenfor nu er ved at få mere pålidelige måder at udpege sorte pletter end den danske. Det skyldes bl.a., at sortpletudpegningen i Danmark ikke er baseret på state-of-the-art uheldsmodeller. I Europa og Nordamerika er der de seneste 10-15 år udført et stort antal forskningsprojekter for at forbedre uheldsmodellernes pålidelighed mv. Sammenholdes den hidtidige danske udpegning af sorte pletter med state-of-the-art kan følgende forhold forbedres:

- Uheldsmodeller vil med stor fordel kunne opstilles ud fra en Poisson-Gamma (Negativ binomial model) fordeling i langt de fleste tilfælde (se afsnit 3.2).
- Sorte pletter bør udpeges ud fra forskelle mellem et forventet og et estimeret antal uheld, hvor det estimerede antal uheld beregnes ud fra EB. Det resulterer i den mest korrekte udpegning af sorte pletter. Man gør brug af spredningsparameteren fra Negativ binomial uheldsmodeller i beregning af estimerede antal uheld (se afsnit 3.2).
- Glider metoden bør ikke anvendes, da denne metode kunstigt øger variansen i uheldsforekomsten. I stedet bør anvendes en metode, hvor vejnettet er inddelt i korte men faste strækningselementer. En sort plet kan bestå af et eller flere på hinanden følgende strækningselementer.
- Ved brug af EB bør der være en vis harmoni mellem længden af strækninger, der indgår i uheldsmodeller, og længden af de faste strækningselementer, der udpeges til sorte pletter – ellers må spredningsparameteren opfattes som en variabel, der er afhængig af strækningsslængden.

- Basismodeller er gode til udpegning af sorte pletter, når de er statistisk holdbare, dvs. sammenhængen mellem trafikmængde og uheldsforekomst er stærk. Er det ikke tilfældet, bør faktormodeller anvendes, så flere typer af kryds eller strækninger indgår, og en uafhængig variabel beskriver typen. Faktormodellen vil kun fungere fornuftigt, hvis de inkluderede typer udviser nogenlunde samme sammenhæng mellem trafikmængde og uheldsforekomst. Flere af basismodellerne i Danmark bygger ikke på stærke sammenhænge, men må anses for at være ustabile.
- I Danmark opereres med relativt høje minimumskrav til trafikmængder for sideveje i kryds. Derved bliver ”strækning” en ret varierende analyseenhed, da en strækning kan dække over stor variation i forekomsten af mindre kryds. Det kan evt. ændres ved at operere med lavere minimumskrav eller ved at lade en variabel for antal mindre kryds indgå i uheldsmodeller for strækninger.
- Udpegningsperioden kan gøres variabel, så steder med mange forventede uheld pr. år udpeges til sorte pletter ud fra en relativ kort årrække (2-5 år), mens steder med få forventede uheld pr. år udpeges ud fra en længere årrække (6-10 år). Hvis årsfaktorer indgår i uheldsmodeller, så der opereres på år-for-år basis, skal både uhelds- og trafikdata opdeles på år.
- Modellerne kan gøres afhængig af geografisk område f.eks. politidistrikt eller region. Uheldsforekomsten varierer mellem områder som følge af forskelle i f.eks. politiets praksis for registrering af uheld, trafikanters brug af sele, spritkørsel, hastighedsniveau, uddannelsesniveau, osv. Disse forskelle vil medføre, at modeller uden områdefaktorer vil udpege mange sorte pletter i nogle områder og få sorte pletter i andre områder. Brug af områdefaktorer vil medføre, at det forventede uheldsniveau er forskelligt fra område til område.

Der findes i dag ikke let anvendelige metoder, der gør det muligt på korrekt vis at lade alvorligheden af uheld eller personskader indgå ved udpegning af sorte pletter. Der findes ikke statistik software, der i dag kan håndtere de korrekte metoder. En norsk metode (*Ragnøy et al., 2002*) kan anvendes til angivelse af et forventet niveau af dræbte, alvorlige og lette skader, men den er ikke korrekt til brug for udpegning af sorte pletter med baggrund i EB.

I forbindelse med udarbejdelse af uheldsmodeller til sortpletudpegning må man tage stilling til to grundlæggende spørgsmål:

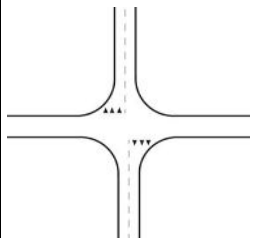
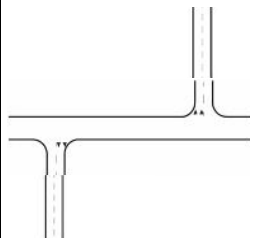
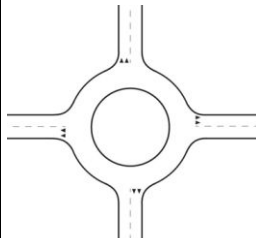
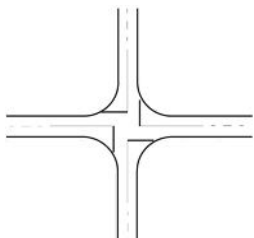
- 1) Hvad skal ”det forventede antal uheld og personskader” gælde for?
- 2) Hvad skal ”det forventede antal uheld og personskader” sammenlignes med?

At besvare disse spørgsmål og derved udarbejde en måde at foretage udpegning af sorte pletter involverer bl.a.; a) typiske tiltag til sortpletudbedring, b) inddeling i stræknings- og krydstyper, c) metoder til at opgøre et forventet antal uheld og personskader, d) metoder til at opgøre regressionseffekten og et estimeret antal

uheld og personskader, og ikke mindst e) antallet af uheld og personskader til opstilling af uheldsmodeller.

Typiske sortpletudbedringer har været f.eks. ombygning til rundkørsel, etablering af venstresvingsbaner, omplacering og forbedring af tavler og afmærkning, signaltekniske ændringer, etablering af kurveafmærkning, omprofilering, etablering af sikkerhedszoner, rabatsanering, osv. Man kan sige, at nogle projekter indebærer et skift fra én stræknings- eller krydstype til en anden, mens andre projekter er udbedringer uden skift af type. Inddelingen i stræknings- og krydstyper har betydning for, hvor godt de udpegede sorte pletter er orienteret mod projekter hhv. med og uden skift af vej- eller krydstype.

Lad os se på kryds i et plan mellem to veje i åbent land. Her kunne der f.eks. være tale om fire forskellige krydstyper, som vist i figur 2. Sikkerhedsniveauet i disse fire krydstyper er forskelligt. Antallet af uheld ved samme trafikmængde er ofte højest i prioriterede kryds og lavest i rundkørsler.

Prioriteret kryds	Forsatte kryds	Rundkørsel	Signalreguleret kryds
			

Figur 2. Fire krydstyper.

Lad os sige, at man ser på 40 sorte pletter blandt disse fire krydstyper. Og lad os for eksemplets skyld sige, at der er lige mange kryds af hver type og alle kryds har samme trafikmængde. I eksemplet vil det forventede antal uheld være gennemsnittet af uheld i de typer af kryds, som man betragter. Hvis alle krydsene betragtes under ét, altså man opfatter dem som én type, vil man hovedsageligt udpege prioriterede kryds som sorte pletter, da de generelt er mere uheldsbelastede. Lad os sige, at 30 prioriterede kryds bliver udpeget til sorte pletter. Denne udpegning sigter mest på sortpletprojekter, hvor krydstypen ændres. Man kunne også betragte krydsene som fire krydstyper og udregne et gennemsnit af uheld for hver type, og derved vil de udpegede sorte pletter blive fordelt mere ligeligt på de fire typer. Lad os her sige, at kun 10 prioriterede kryds bliver udpeget til sorte pletter. Denne udpegning sigter mere på sortpletprojekter, hvor krydstypen ikke ændres.

Der vil være et overlap af udpegede sorte pletter mellem de to udpegningsmåder, men højest 20 sorte pletter optræder i begge udpegninger. Hvis målet er at forebygge flest uheld for færrest ressourcer er spørgsmålet derfor: Hvilken udpegningsmåde giver den højeste omkostningseffektivitet? Det vil man kun kunne erfare ved at benytte begge udpegningsmåder, og så prioritere de opstillede sortplet-

projekter. I og med en del sortpletprojekter indebærer skift af stræknings- eller krydstype vil begge udpegningsmåder være relevante.

Det tænkte eksempel ovenfor kompliceres i den virkelige verden af, at der findes mange varianter af hver type. De 4-benede prioriterede kryds kunne f.eks. deles op i kryds hhv. med og uden primærkanalisering. Spørgsmålet bliver så, om det bedst kan betale sig at udpege de farligste prioriterede kryds, eller de farligste prioriterede kryds hhv. med og uden primærkanalisering, eller anvende begge udpegningsmåder. En del udbedringer af sorte pletter indebærer, at man går fra en variant af en stræknings- eller krydstype til en anden variant, og dette medfører, at begge udpegningsmåder vil være relevante.

Svaret på spørgsmål 1) ”Hvad skal det forventede antal uheld og personskader gælde for?” er derfor ikke lige til. Det kan være relevant at opgøre det forventede antal indenfor en gruppe af stræknings- eller krydstyper, som vejen eller krydset kunne tænkes at blive ændret fra og til via et sortpletprojekt – dvs. kun én model uden en variabel, der beskriver type. Det kan også være relevant at opgøre det forventede antal uheld for én stræknings- eller krydstype, men for en gruppe af varianter, som vejen eller krydset kunne tænkes at blive ændret fra eller til via et sortpletprojekt – dvs. én model, der kan opgøre et forventet antal for denne type, men uden en variabel, der beskriver variant. Og det kan slutteligt også være relevant at opgøre det forventede antal for en variant af en type for derved at finde de steder, hvor der er noget galt med det detaljerede design af varianten. Opgørelsen af det forventede antal bør tage højde for sammenhænge mellem trafikmængde og uheldsforekomst, altså bør det forventede antal opgøres med baggrund i en uheldsmodel, hvori trafikmængde optræder som en eller flere uafhængige variable. Et resultat af en sådan type sortpletudpegning kunne være:

Sted-ID	Art	Sort plet ift. art	Type	Sort plet ift. type	Variant	Sort plet ift. variant
Strækning 1	Motorvej	-	4-spor	-	Bynær	Ja
Strækning 2	Vej > 11 m	Nej	4-spor	Nej	Med midterrabat	Ja
Strækning 3	Vej < 11 m	Ja	2-spor	Ja	Uden kantbane	Nej
Kryds 1	3-ben i plan	Nej	Signal	Ja	Tidsstyret	Nej
Kryds 2	4-ben i plan	Ja	Prioriteret	Nej	Primærkanalisering	Ja
Kryds 3	5+ben i plan	Nej	Rundkørsel	Ja	1 cirkulationsspor	Ja
Kryds 4	2-plans kryds	-	Hanke	Ja	1 hank	Nej

Tabel 2. Eksempel på alternativ sortpletudpegning.

Eksemplet i tabel 2 kræver nogle bemærkninger. For strækning 1, som er angivet til at være motorvej, opereres med forskellige typer og varianter, men motorvejen udpeges ikke som sort plet ud fra arten eller typen, men alene med udgangspunkt i varianten. Det skyldes, at de forhold, der kan betragtes som udslagsgivende for opdeling i typer og varianter f.eks. afstand mellem tilkørsler, antal kørespor, forekomst af nødspor, osv., ekstremt sjældent ændres i et sortpletprojekt. Således er

det egentligt kun interessant at udpege sorte pletter på motorveje med baggrund i den enkelte variant. For en ikke-motorvej er der derimod mange forhold, som kan anvendes til opdeling i typer og varianter, der også kan ændres i sortpletprojekter f.eks. opdeling i kørespor og kantbaner, forekomst af midterrabat/-areal, osv. Det samme gør sig gældende for de fleste arter, typer og varianter af kryds. Ses på kryds 2 er det angivet til at være sort plet i forhold til arten og varianten, men ikke i forhold til typen. Det kan f.eks. skyldes, at prioriterede kryds er en ret usikker type inden for arten 4-benede kryds i plan, men primærkanalisering er en ganske sikker variant blandt de prioriterede kryds, og kryds 2 er angivet til at være usikker set i forhold til andre kryds af denne variant.

I praksis er der i Danmark så få km vej eller kryds af nogle typer og varianter, at det er statistisk uholdbart at angive et forventet antal uheld og personskader. Derfor vil inddelingen i typer og varianter i stor udstrækning være styret af omfanget af tilgængelige data.

Spørgsmål 2) er: *"Hvad skal det forventede antal uheld og personskader sammenlignes med?"* for at udpege stedet til en sort plet. Her er tre valg vigtige: a) fastlæggelse af strækning (længde og faste eller flydende start-/slutpunkter) og kryds (udbredelse) både som observationer, der indgår i uheldsmodeller, og som steder der udpeges til sorte pletter, b) sammenligne med et rapporteret eller et estimeret antal, og c) hvis der sammenlignes med et estimeret antal, hvilken uheldsmodel (art, type, variant) skal anvendes til at beregne dette.

Glider metoden øger kunstigt variansen, og det er den egentlige baggrund for, at vejbestyrelserne i Danmark har brugt et højt signifikansniveau i udpegningen af sorte pletter på strækninger. Forskning (Elvik, 2007) viser, at udpegning af sorte pletter med baggrund i et estimeret antal uheld og personskader medfører, at færre falske sorte pletter udpeges. Glider metoden bør slet ikke anvendes sammen med udpegning ud fra et estimeret antal uheld, da variansen i uheldsforekomsten blandt observationer i uheldsmodellen ikke er den samme, som variansen blandt steder, der udpeges til sorte pletter. For at kunne udpege flest sande sorte pletter er man nødt til at operere med faste start-/slutpunkter og udpege med baggrund i et estimeret antal uheld. Længden af strækningsenheder, der indgår i uheldsmodeller, kan variere, men det vil være hensigtsmæssigt, at mange enheder har en længde, der er omkring den længde, som de udpegede sorte pletter har.

For kryds er det også bedst at sammenligne det forventede antal uheld og personskader med et estimeret antal for at udpege sorte pletter. Fastlæggelsen af krydsets udbredelse som observation har stor betydning af flere årsager. For det første er det uvidt i hvilket omfang uheld, der sker uden for området, hvor trafikstrømmene krydser hinanden (konfliktområdet – selve krydsningsarealet), stedfæstes som værende sket i krydset. For det andet er den geometriske udbredelse af kryds med heller, opmarchfelter, osv. forskellig, hvilket vanskeliggør fastlæggelsen af sammenlignelige observationer. For det tredje influerer typen og varianten uheldsforekomsten op til krydset, da trafikantadfærden op til krydset afhænger af typen og

varianten. For at kunne se på tværs af typer og varianter er det måske nødvendigt at lade de sidste 100-200 m vej i alle retninger op til konfliktområdet indgå i den observation der kaldes ”kryds”.

Hvis man vælger at sammenligne det forventede antal med et estimeret antal, er det særdeles vigtigt, at opgørelsen af et estimeret antal altid udføres ved at bruge den mest specificerede uheldsmodel (og stadig statistisk holdbare) med hensyn til type og variant. Det skyldes, at jo mere vej- og krydsdesignet blandt observationerne i uheldsmodeller varierer (uden variabel for denne variation), desto mindre en andel af den systematiske variation forklares af modellen, og så vil regressionseffekten blive estimeret mere upræcist. Så selvom det forventede antal uheld er baseret på f.eks. en model for arten af kryds med 4-ben i plan, så er det mest fornuftigt, at det estimerede antal uheld findes ud fra en model for en specificeret type eller variant f.eks. 4-benede prioriterede kryds med primærkanalisering.

2.2 Valg af vej- og krydstype ved anlæg af nye veje

Ved anlæg af nye veje og kryds indgår mange overvejelser i relation til designet. En overvejelse er trafiksikkerhed. Anlægget skal helst fungere sikkerhedsmæssigt forsvarligt – og helst sikkerhedsmæssigt optimalt i forhold til prisen for anlæg og drift. Optimeringen involverer valget af vej- og krydstype og valg af én blandt de mange varianter for hver type.

For at kunne foretage optimeringen er man nødt til at vurdere, hvor mange uheld og personskader man kan forvente vil ske, hvis man vælger den ene eller anden type og variant. Det er ikke nok at kende de procentuelle forskelle i sikkerhedsniveau f.eks. hvor mange procent af de fremtidige uheld og personskader, som vejbelysning på en kommende landevej vil forebygge. Man er nødt til at vurdere det absolutte antal af fremtidige uheld og personskader som f.eks. vejbelysning kan forebygge – for at kunne vurdere, om det kan betale sig at etablere vejbelysning, altså en samstilling af økonomi og antal forebyggede uheld og personskader.

Når man skal forsøge at vurdere, hvor mange uheld man kan forestille sig vil ske på en ny vej eller i et nyt kryds, kan man af gode grunde ikke vurdere uheld, der skyldes lokale faktorer, men må alene søge at vurdere uheld, der sker som følge af generelle faktorer. Uheld, der skyldes lokale faktorer, kan ikke vurderes, da der ikke er rapporteret uheld for den kommende vej eller kryds endnu.

Derfor vil uheldsmodeller, der skal anvendes til valg af vej- eller krydstype, alene skulle beregne det forventede antal uheld og personskader – ikke det estimerede. Men det er interessant at angive en usikkerhed på det forventede antal uheld og personskader f.eks. ved et konfidensinterval eller standardafvigelse. Til angivelse af en usikkerhed er det bedst at benytte uheldsmodeller, der er specificeret på samme måde som uheldsmodeller, der muliggør EB metodikken.

Det må frarådes at anvende de uheldsmodeller, der bruges til udpegning af sorte pletter, til at beregne et forventet antal uheld og personskader for nye veje og kryds. Der er enighed i forskningskredse om, at basismodeller til udpegning sorte pletter som i Danmark ikke er brugbare til valg af vej- og krydstype, fordi modellerne ikke kan forklare forskelle i uheldsniveau mellem forskellige typer og varianter af strækninger eller kryds. Der er også enighed om, at faktormodeller, der bruges til sortpletudpegning, heller ikke er gode at anvende til valg af vej- og krydstype.

Det står i kontrast til Vejdirektoratets udsagn om, at uheldsmodeller til udpegning af sorte pletter godt kan anvendes til beregning af forventede uheld ved nyanlæg og ombygninger, dog angiver Vejdirektoratet, at vurderingen er forbundet med nogen usikkerhed og uheldstal for nye strækninger bør ganges med en faktor på 0,6 (*Greibe og Hemdorff, 2001*).

Lad os se nærmere på problemet ved at se på et eksempel. På basis af eksisterende danske uheldsmodeller kan man beregne det forventede antal af hhv. person- og materielskadeuheld i 10 krydstyper og 9 strækningstyper i det åbne land. I tabel 3 er angivet beregnede forventede uheldsniveauer pr. år i 2001-2005 for et T-kryds i åbent land med 8.000 indkørende ÅDT på primærvejen og 1.000 indkørende ÅDT på sekundærvejen ud fra danske uheldsmodeller. Modellerne er beskrevet af Hemdorff (2006).

Krydstype	Personskade-uheld	Materielskade-uheld	Uheld i alt
521: 3-ben signalreguleret	0,20	0,27	0,47
522: 3-ben primær + sekundær vej kanaliseret	0,14	0,18	0,32
523: 3-ben primær vej kanaliseret	0,08	0,11	0,19
524: 3-ben sekundær vej kanaliseret	0,19	0,19	0,37
520: 3-ben ingen kanalisering	0,07	0,09	0,16

Tabel 3. Beregnede forventede uheldsniveauer pr. år i 2001-2005 for T-kryds med 8.000 indkørende ÅDT på primærvej og 1.000 på sekundærvej.

Af tabel 3 kunne man foranlediges til at tro, at et vigepligtsreguleret T-kryds uden kanalisering er den sikreste og et signalreguleret T-kryds er den usikreste af de fem krydstyper. Det harmonerer ikke med den troværdige viden, om den sikkerhedsmæssige betydning af kanalisering og signalregulering af kryds fundet ud fra før-efter uheldsevalueringer. Disharmonien kan skyldes flere ting. Nedenfor er kort nævnt nogle potentielle muligheder, og disse er efterfølgende diskuteret:

- Uheldsmodellernes gyldighedsområde
- Variablenes forklaringskraft – usikkerhed
- Type af uheldsmodel – modelspecifikation
- Variable og observationer i uheldsmodellerne
- Endogene uafhængige variable

Vejdirektoratet har ikke angivet for hvilke trafikmængder, at uheldsmodellerne er gyldige. Det er således ikke sikkert, at alle uheldsmodellerne, der er anvendt til at beregne resultaterne i tabel 3, er gyldige ved de anvendte trafikmængder. Med gyldighedsområde menes på en simpel facon modellens indgangsværdier på nær ekstreme værdier (outliers), mens man på en mere avanceret facon kan fastlægge gyldighedsområde ud fra analyser af residualer, der er forskelle mellem rapporterede og forventede uheldstætheder.

Variablenes forklaringskraft eller usikkerheden på resultatet kan være vigtig at have for øje i en vurdering af forventede uheld og personskader. Det er ganske varierende i hvilket omfang de uafhængige variable (trafikmængde på primær- og sekundærvej) i uheldsmodellerne, der er grundlaget for tabel 3, kan forklare den systematiske variation i uheldstætheden. Faktisk synes usikkerheden på resultatet at være relativ beskeden for model nr. 521, 522 og 524, mens den er stor for nr. 520 og 523. Usikkerheden beskrives f.eks. ved at angive et konfidensinterval.

Modelspecifikationen er også vigtig. I den aktuelle sammenligning i tabel 3 er resultaterne baseret på 5 basismodeller. Hvis den i stedet var baseret på kun én faktormodel, hvor krydstype (regulering og kanalisering) var en eller flere uafhængige variable, så ville sammenhængen mellem trafikmængde og uheldsforekomst blive bedre bestemt. Det vil gøre sammenligningen mindre usikker. I faktormodeller er samvariation mellem uafhængige variable dog ofte et problem, og man kan være nødsaget til at dele faktormodellen op i flere modeller eller at fjerne en eller flere af uafhængige variable. Et andet problem kan være, at trafikmængder er kontinuerte variable i en potensfunktion. Sammenhænge mellem trafikmængde og uheldsforekomst kan være mere kompleks med lokale maksima, minima, spring, osv. En sammenligning som i tabel 3 vil måske være mere pålidelig, hvis man alene lod kryds, der blev anvendt til at opstille modeller og havde f.eks. 7.000-9.000 indkørende ÅDT på primærvejen og 750-1.250 på sekundærvejen, indgå. På den måde kan problemer med gyldighedsområde og forklaringskraft undgås, hvis der opstilles krav til uheldsforekomsten, som sammenligningen baseres på – og man lader de krydstyper udgå, som ikke imødekommer kravene.

Præcis hvilke variable og observationer (strækninger eller kryds), der indgår i uheldsmodellerne – og hvilke der er udeladt, er vigtigt for, om modellerne kan anvendes til valg af vej- og krydstype. De trafikale og vejtekniske forskelle blandt observationer, der indgår i uheldsmodeller og ikke er beskrevet ved uafhængige variable, kan være så store, at det er umuligt at sige, hvordan den enkelte vej- eller krydstype ser ud eller hvordan trafikken er sammensat og opdelt. Typologien er så at sige ikke specificeret nok til, at den kan anvendes til valg af vej- eller krydstype – og slet ikke til valg af variant. En anden ting er, at en række forskelle i design og trafik helt sikkert samvarierer stærkt med trafikmængden, der indgår som variable i modellen. Det vil sige, at et kryds af én type – uden tvivl – ser vidt forskelligt ud afhængig, om det har 2.000 eller 15.000 indkørende motorkøretøjer pr. døgn. Det kan hjælpe at lade flere variable beskrive de store forskelle i design og trafik, men man har sjældent data til at opstille yderligere variable. Samtidig er det ikke sik-

kert, at det hjælper, dels fordi variablene kan udvise stor samvariation, dels fordi de muligvis blot er en proxy for en eller flere udeladte variable, som faktisk er dem, der betyder noget for uheldstætheden.

I modeller bør uafhængige variable så vidt muligt være eksogene. Er de ikke det, så er der tale om endogene uafhængige variable. En uafhængig variabel er endogen, hvis der er samvariation med andre af de uafhængige variable, eller den uafhængige variabel afhænger af den afhængige variabel (uheldstætheden). I eksemplet med 3-benede kryds er der – uden tvivl – tale om endogene variable. Mængden af trafik på primærvejen samvarierer med mængden af trafik på sekundærvejen. Hvis denne samvariation ikke er særlig stærk – og i øvrigt nogenlunde ens på tværs af de 5 basismodeller – skulle det ikke betyde så meget for sammenligningen af resultater i tabel 3, men det er næppe tilfældet. Komplekset af modeller bag resultaterne i tabel 3 kan også siges at indeholde endogene variable. Det ene er, at mængden af trafik afhænger af modellen – altså gennemsnittet af trafik i modellen for lyskryds er højere end i modellen for ukanaliserede vigepligtsregulerede kryds. Når vi ser på tværs af modellerne, så afhænger den uafhængige variabel, trafik, altså af den afhængige variabel, uheldstæthed. Den anden ting er, at nogle af krydsene måske er blevet kanaliseret eller signalreguleret, fordi der tidligere skete mange uheld. Det vil sige, at typen af kryds er betinget af en tidligere uheldstæthed. Da typen af kryds ikke kan forklare hele forskellen i uheldstæthed (sammen med trafikmængden), så må der være uforklarede forhold (ikke beskrevne forhold, der udgør et uheldsproblem), der samvarierer med både med krydstypen og uheldstætheden. Derved kommer den uafhængige variabel i modelkomplekset, nemlig krydstypen, til at afhænge af uheldstætheden.

Men hvis de sædvanlige uheldsmodeller til sortpletudpegning ikke kan anvendes til valg af vej- og krydstype, hvad gør man så? De nævnte problemer med gyldighedsområder og usikkerheder på resultater kan håndteres ved oplysning om disse og / eller ved inddeling i små trafikintervaller. Problemer med modelspecifikation og samvariation mellem uafhængige variable kan løses ved at overgå til færre modeller – faktormodeller – og at udvælge mere ortogonale sæt af observationer, hvorved samvariation mellem uafhængige variable kan undgås. Derved bliver datasættet reduceret, men forklaringskraften bliver bedre.

Problemer med udeladte men betydningsfulde variable samt endogene variable er dog ikke lige til at løse. Det vil kræve, at variable for væsentlige uheldsproblemer opstilles og data herfor indsamles for alle observationer. Derved bliver problemet med udeladte variable løst. Problemet med endogene variable kan løses ved en metode angivet af Kim og Washington (2006), hvor uheldsmodellens konstanter estimeres simultant med en model for f.eks. forekomsten af kanalisering. Men det vil blive en stor indsats, der ikke står mål med udbyttet. Alternativt må fuldstændigt anderledes former for og sæt af modeller opstilles.

Et sådant alternativ har man opstillet i USA i projektet Highway Safety Manual forkortet HSM (AASHTO, 2010). Grundfilosofien i HSM er, at man udarbejder en

grundlæggende uheldsmodel (Safety Performance Function) for en udspecificeret variant af en kryds- eller strækningstype. Udarbejdelsen af denne **grundmodel** bygger alene på observationer, der svarer til den udspecificerede variant. Grundmodeller opstilles for at få et tidssvarende uheldsniveau og en "ren" sammenhæng mellem uheldsforekomst og trafikmængde. Trafikken kan være mere detaljeret beskrevet end i danske uheldsmodeller til sortpletudpegning.

Beregningen af forventede uheld og personskader for kommende veje og kryds foretages efterfølgende ved at anvende grundmodellen samt ingen, en eller flere sikkerhedsfaktorer (Crash Modification Factors). En sikkerhedsfaktor (**SF**) er odds ratio værdien for ændringen i sikkerhedsniveauet, som en ændring af vej- eller krydsdesign vil medføre. Hvis en ændring af designet f.eks. etablering af 1 m brede kantbaner i vejsiderne, medfører en forbedring af sikkerheden på 25 %, så vil SF'en være $1 - 0,25 = 0,75$. SF'er stammer fra pålidelige studier, f.eks. gode før-efter uheldsevalueringer eller analyser af designelementer. I det amerikanske modelkompleks med grundmodeller og SF'er opereres også med usikkerhed på det forventede antal uheld.

Eksempel: En grundmodel gælder for landeveje med 2 kørespor á 3,5 m bredde med midt- og kantlinjer, ingen kantbaner, ingen cykelstier og fortove, 80 km/t hastighedsbegrænsning, ingen vejbelysning, samt kurveradier over 1000 m. Lad denne grundmodel være lig den i figur 1, altså $UHT = 0,0015 \times \text{ÅDT}^{0,7}$. Bruges filosofien i HSM, så kan det forventede antal uheld beregnes for en landevej med 5.000 i ÅDT, der ligner landeveje i grundmodellen, men med 1 m brede kantbaner ved at multiplicere grundmodellen med sikkerhedsfaktoren, altså de forventede uheld bliver $0,0015 \times 5000^{0,7} \times 0,75 = 0,44$ uheld pr. år.

Hvis den landevej man ønsker at beregne det forventede antal uheld for, er endnu mere forskellig fra landeveje i grundmodellen, så anvendes flere SF'er. Hvis det forventede antal uheld skal beregnes for en landevej, som til forskel fra landevejene i grundmodellen har 3,0 m brede kørespor, 1 m brede kantbaner og vejbelysning, så multipliceres grundmodellen med SF'er for hhv. kørespor, kantbaner og vejbelysning.

Metodikken i HSM har givet anledning til kritik og kommentarer. Herunder er angivet nogle af disse:

- Der findes ikke SF'er for alle typer og varianter.
- Flere SF'er afhænger sandsynligvis af trafikmængden (eller påvirker trafikmængden), men dokumentationen for denne afhængig er meget usikker eller ikke-eksisterende, og derfor opereres med SF'er uden en sådan sammenhæng.
- Der er sandsynligvis synergieffekter mellem en lang række designelementer. I så fald vil metoden med at multiplicere SF'er (også kaldet **restfaktormetoden** i Danmark) give afvigende resultater. Forskere har eftervist, at metoden med blot at multiplicere SF'er giver lettere afvigende resultater, så der er altså sikkerhedsmæssig synergi mellem designelementerne (Lord et al., 2010).

Med nye studier vil flere typer og varianter ad åre blive beskrevet med SF'er, og SF'ers afhængighed af trafikmængde vil formentligt også blive beskrevet for et tiltagende antal SF'er.

Men problemet med synergieffekter er alvorligt. Det fører frem til spørgsmålet: *Hvor meget kan et opstillet alternativ adskille sig fra veje eller kryds, der indgik i grundmodellen, før en forventet uheldstæthed ikke længere er troværdig?* Hvis man kendte svaret på spørgsmålet, vil man kunne angive, hvor mange grundmodeller der må opstilles for at kunne afdække de typer og varianter, der opereres med i praksis. Men man kender ikke svaret, og derfor har man iværksat store forskningsprojekter i USA for at få det belyst.

Den amerikanske metodik vurderes at være den mest hensigtsmæssige til valg af vej- og krydstype i forbindelse med anlæg af nye veje. Metodikken kan bruges på to forskellige måder, nemlig i relation til:

- Valg af vej- og krydstype for nye veje, og
- Opstilling af manualer og vejregler i relation til planlægning af nye veje.

Det er oplagt at anvende den amerikanske metodik til valg af vej- og krydstype for nye veje i Danmark. Det vil kræve opstilling af danske grundmodeller. Ved opstilling af forskellige alternative strækings- eller krydsdesign for ny vejinfrastruktur er det også muligt at give overslag på prisforskelle for anlæg og drift for enkelte alternativer. Disse prisforskelle kan så sammenstilles med forskelle i trafikssikkerhed for alternativerne.

Metodikken kan også give input til, hvordan fremtidige veje og kryds mere generelt kan anbefales at blive designet. Her kan man ved at sammensætte et sæt af foretrukne designs analysere, hvordan disse designs fungerer sikkerhedsmæssigt i forhold til hinanden ved forskellige trafikmængder, hastighedsniveauer, osv. Og dette kan give input til anbefalinger om f.eks. køresporsbredde, kantbaner og vejbelysning.

Amerikanerne ønsker også at anvende deres modelkompleks med grundmodeller og SF'er til at udpege sorte pletter, hvor de så vil opgøre det forventede antal uheld ud fra grundmodellen og de designkarakteristika med tilhørende SF'er, som hver strækning og kryds har. Det er en interessant indgangsvinkel, men nærmest uoverkommelig opgave. For det første vil det kræve langt flere oplysninger, om hver strækning og kryds i forbindelse med udpegningen. Samtidig er det eksisterende vejnet langt mere varieret end kommende nye veje og kryds, og derfor vil der være brug for yderligere SF'er, hvis modellerne skal anvendes til udpegning af sorte pletter. En udpegning af sorte pletter med brug af grundmodeller og SF'er vil således være betydeligt mere ressourcekrævende end en udpegning med brug af de uheldsmodeller, som er angivet i afsnit 2.1.

2.3 Vurdering af forventet uheldsbesparelse

I forrige afsnit indgik uheldsbesparelser ved at overgå fra et alternativ til et andet i relation til anlæg af nye veje. I nærværende afsnit ses på, hvilken uheldsbesparelse der kan forventes ved ombygning af eksisterende veje på to måder:

- Forventet uheldsbesparelse ved et tiltag ét konkret sted.
- Analyse af forventede uheldsbesparelser ved samme tiltag for en lang række steder til brug for trafiksikkerhedsplaner, designmanualer eller vejregler.

For at kunne vurdere størrelsen af en forventet uheldsbesparelse for eksisterende strækninger eller kryds må der foreligge oplysninger, der kan bruges til at angive; a) hvor mange uheld og personskader der kan estimeres at ske på lang sigt med den nuværende udformning, og b) hvilken andel af det estimerede antal uheld og personskader tiltaget kan forebygge.

I Danmark har det været og er det mest almindeligt at bruge det rapporterede antal uheld og personskader til vurdering af forventet uheldsbesparelse. Det leder ofte til en overvurdering af uheldsbesparselsen, da der ikke tages højde for regressions-effekt. Det giver mere præcision i vurderingen af forventet uheldsbesparelse, hvis man benytter det estimerede antal frem for det rapporterede.

Der er mange metoder til at vurdere, hvilken andel af uheldene, som et tiltag kan forebygge. En metode er at bruge sikkerhedseffekter, altså SF'er der stammer fra gode før-efter uheldsevalueringer eller analyser af designelementer. En anden metode bruges i stor udstrækning ved forhåndsvurdering af sortpletprojekter i Danmark. Her vurderes hvert uheld og hver personskade for sig med hensyn til, om det kan forebygges med tiltaget og i hvilken udstrækning. Sortpletmetodikken er relateret til de uheldsproblemer, som en uheldsanalyse identificerer, samt de erfaringer man har med de enkelte tiltag og deres indvirkning på uheldsproblemerne. Metodikken med SF'er er nemmere og mindre tidskrævende at bruge end sortpletmetodikken, og kræver ikke så stor ekspertviden, som er nødvendig for at mestre sortpletmetodikken.

I Danmark har man også brugt en metode med at sammenligne resultater fra de uheldsmodeller, der anvendes til sortpletudpegning. Det må frarådes af årsager nævnt førhen.

Det kan dog være en idé i at benytte uheldsmodeller til at vurdere en forventet uheldsbesparelse, når der er tale om en omfattende ombygning. Her må man bruge den amerikanske metodik med grundmodeller og SF'er.

Hvis man skal bygge et signalreguleret kryds om til en rundkørsel, så kunne man beregne det forventede antal uheld og personskader for hver af disse design, og så bruge den procentuelle forskel mellem disse forventede tal til at vurdere, hvilken andel af uheldene, som tiltaget kan forebygge. Den andel ganges så med det esti-

merede antal uheld for det signalregulerede kryds, og således fås den forventede uheldsbesparselse. Med dette angives indirekte, at den absolutte forskel i de forventede uheld ikke vil svare til den uheldsbesparselse, der kan forventes. Ej heller vil den absolutte forskel mellem estimerede uheld for det signalregulerede kryds og forventede uheld for rundkørslen svare til den uheldsbesparselse, der kan forventes. I begge tilfælde skyldes det, at det er uvist, hvilken betydning ombygningen vil have for de uheld, der skyldes lokale faktorer (uforklarede faktorer). Den bedste antagelse er derfor, at den andel af de uheld, der kan skyldes generelle faktorer, som tiltaget beregnes at forebygge (forskelle mellem to forventede tal) – den andel antages også at blive forebygget blandt uheld, der skyldes lokale faktorer.

Baggrunden for at bruge ovenstående metodik til at vurdere en forventet uheldsbesparselse, når der er tale om en omfattende ombygning er, at der findes mange forskellige slags signalregulerede kryds og rundkørsler. Sikkerhedseffekten af at ombygge fra et signalreguleret kryds til en rundkørsel kan ikke siges altid at være den samme uanset design før og efter. Kun i tilfælde, at der findes grundmodeller og SF'er, der faktisk kan belyse forskelle i sikkerhedsniveau ved diverse design af hhv. signalregulerede kryds og / eller rundkørsler – er det interessant at bruge ovenstående metodik, da man ellers blot kan bruge en SF, der beskriver den sikkerhedsmæssige forskel på de to krydstyper.

En anden situation, hvor man gerne vil beregne en forventet uheldsbesparselse, er i forbindelse med ændringer i trafikmængder på influensvejnettet ved anlæg af nye veje. Eksempelvis vil anlæg af en omfartsvej medføre, at antallet af uheld på den eksisterende gennemfartsvej vil falde, da noget trafik flyttes til omfartsvejen. Spørgsmålet er, hvordan man bedst kan beregne en forventet uheldsbesparselse for gennemfartsvejen. Umiddelbart anses ovenstående metodik med beregning af estimerede uheld samt forventede uheldstal, dog her baseret på én grundmodel og tilhørende SF'er ved to forskellige trafikmængder (i stedet for to grundmodeller ved samme trafikmængde), som den mest korrekte metode. Det skyldes, at sammenhængen mellem trafikmængde og uheldsforekomst anses for at være mere korrekt i en grundmodel end i en model til sortpletudpegning, da grundmodellen er "renset" for indvirkningen af varierende design.

I rapporten "Effektkatalog – Viden til bedre trafiksikkerhed" (Jensen, 2008) er der en række eksempler på analyser af forventede uheldsbesparelser ved samme tiltag for en lang række af steder. Deri er brugt pålidelige SF'er sammen med sortpletmodeller og en speciel metode, hvor der tages højde for regressionseffekt på tværs af steder med samme uheldstæthed, til at beregne de forventede uheldsbesparelser. Den anvendte metode i Effektkataloget kunne forbedres, hvis der var beregnet estimerede uheldstal for hvert sted. Men det var umuligt, da uheldsmodellerne ikke fandtes. Hvis estimerede uheldstal var anvendt, ville præcisionen i den forventede uheldsbesparselse for det enkelte sted være større.

Et eksempel på analyse af forventede uheldsbesparelser kunne være en vurdering af nytten af at udføre rumleriller i vejmidten på landeveje. Med udgangspunkt i

rumlerillers sikkerhedseffekt og pris samt samfundsøkonomiske enhedspriser for uheld og personskader kan man med baggrund i forventede uheldsbesparselser på de analyserede landeveje vise ved hvilken trafikmængde, at det er økonomisk fordelagtigt at udføre rumleriller.

Vurdering af forventet uheldsbesparselse er forbundet med usikkerhed. Det er ikke muligt at angive denne usikkerhed præcist, da kilder til usikkerhed er forskellige og ikke kan beskrives fuldt ud. Usikkerheden på den forventede uheldsbesparselse kan dog angives på en rimelig facon ved at samle usikkerheden på det estimerede uheldstal og usikkerheden på sikkerhedseffekten. Den almindelige måde at angive en samlet usikkerhed er "kvadratroden af kvadratsummen". Det vil sige, hvis det estimerede uheldstal er $10 \pm 10 \%$ (dvs. der sker mellem 9 og 11 uheld – de $\pm 10 \%$ er standardafvigelsen) og sikkerhedseffekten (angivet som SF – odds ratio) er $0,75 \pm 0,05$ (dvs. fald i uheld på 20-30 procent), så vil den samlede usikkerhed på den forventede uheldsbesparselse på $10 \times 0,25 = 2,5$ være

$$\sqrt{(10 \%)^2 + \left(\frac{0,05}{1-0,75}\right)^2} = \sqrt{0,1^2 + 0,2^2} = \pm 22,4 \%$$

Uheldsbesparselsen er således $2,5 \pm 22,4 \%$, altså mellem 1,94 og 3,06 uheld.

2.4 Før-efter uheldsevalueringer

Nogle år efter udbedringer af sorte pletter eller andre uheldsbelastede steder eller noget helt tredje er udført, så kan en før-efter uheldsevaluering vise, hvad man reelt fik ud af tiltagene. Her skelnes mellem to slags uheldsevalueringer. En slags er en uheldsevaluering af et enkelt eller få steder, der er blevet ombygget. Her vil usikkerheden på resultatet – den fundne ændring i trafiksikkerheden – være så stor, at brug af en relativ simpel metode er at foretrække. Den anden slags uheldsevaluering bruges til at dokumentere et tiltags sikkerhedseffekt (SF), og derfor ses på mange steder, der er ombygget på nogenlunde samme måde, så usikkerheden på resultatet på tværs af disse steder er lille. Her må anvendes avancerede metoder for at kunne tage højde for skævheder, opføre statistiske resultater, osv.

I begge tilfælde – simpel og avanceret uheldsevaluering – er det en fordel at kunne gøre brug af uheldsmodeller. Både i den simple og den avancerede evaluering er man nødt til at beregne det antal uheld, som man kunne forvente ville være sket i den undersøgte periode efter tiltagets implementering, hvis stedet ikke var blevet bygget om. For kunne gøre det, må man have et uheldstal for perioden før tiltagets implementering, der kan fremskrives til efter-perioden. Her kan uheldsmodeller bruges både til at tage højde for regressionseffekt i før-perioden, og derved beregne et estimeret antal uheld for før-perioden (EB baseret uheldsmodel), og til at fremskrive til efter-perioden ved at sætte to uheldsmodeller gældende for samme type og variant i hhv. før- og efter-periode i forhold til hinanden – evt. en model med årsfaktorer. Uheldsmodeller kan også anvendes til at tage højde for regressi-

onseffekt i efter-perioden. Endelig kan uheldsmodeller anvendes ved anlæg af nye veje og kryds til at beregne et forventet antal uheld for efter-perioden ved brug af grundmodeller og SF'er.

Før-efter uheldsevaluering kan udføres uden brug af uheldsmodeller. I de enkelte evalueringer uden modeller afgør diverse omstændigheder, om det er muligt at tage højde for regressionseffekt på behørig vis. Før-efter uheldsevalueringer, der ikke tager højde for regressionseffekt, kan udvise meget upålidelige resultater. Det er ikke muligt at udføre en uheldsevaluering af nye vejanlæg uden brug af uheldsmodeller.

2.5 Opsummering

Uheldsmodeller kan anvendes til flere formål f.eks. udpegning af sorte pletter, valg af vej- og krydstype, vurdering af forventet uheldsbesparselse samt før-efter uheldsevaluering.

Eksisterende danske uheldsmodeller kan kun anvendes til sortpletudpegning og med modifikationer også til visse vurderinger af forventede uheldsbesparselser. Uheldsmodellerne vil blive mere præcise og kunne anvendes til flere formål, hvis de blev baseret på en Poisson-gamma fordeling (Negativ binomial model) med en efterfølgende empirical Bayes (EB) metodik. Med en EB tilgang vil udpegningen af sorte pletter blive forbedret lidt, mens vurderingen af forventet uheldsbesparselse vil blive forbedret meget, og brug af uheldsmodeller i før-efter uheldsevaluering vil blive muliggjort.

Hvis uheldsmodeller skal anvendes til valg af vej- og krydstype, er det nødvendigt med en anden type af modeller end de modeller, der anvendes til udpegning af sorte pletter. Den anden type af modeller, som er benævnt grundmodeller, er til forskel fra sortpletmodeller, udarbejdet på grundlag af data om en udspecificeret variant af en stræknings- eller krydstype. For at kunne beskrive det forventede uheldsniveau for alternative stræknings- og krydsdesign, så anvendes en eller flere grundmodeller i kombination med sikkerhedsfaktorer (SF'er). En SF beskriver den sikkerhedsmæssige betydning af et designelement f.eks. køresporsbredden. Vurderinger af forventet uheldsbesparselse af større ombygninger af strækninger og kryds vil også bedst kunne foretages med brug af grundmodeller og SF'er.

Hvis uheldsmodeller skal indgå på alle anførte anvendelsesområder, så vil det være bedst at opstille to sæt af modeller. Disse er:

- Sortpletmodeller med mulighed for beregning af estimerede uheld, dvs. en Negativ Binomial model med tilhørende Empirical Bayes metodik. De modeller skal dække eksisterende kryds og strækninger, som kan være sorte pletter, eller hvor det er relevant at beregne en forventet uheldsbesparselse, altså hvor

tiltag er relevante. Modellerne bør kunne angive usikkerhed på bestemmelsen af forventede og estimere uheld.

- Grundmodeller med tilhørende SF'er, hvor der skal være en lav usikkerhed på bestemmelsen af forventede uheld, hvilket bedst gøres ved brug af en model-specifikation, der også muliggør en empirical Bayes metodik. Modeller og SF'er bør kunne dække typer og varianter af kryds og strækninger, som kan forventes etableret i nærmeste fremtid enten ved nyanlæg eller omfattende ombygninger.

I nærværende kapitel er alvorlighed af uheld og personskader ikke behandlet nærmere. Det er muligt, at sortplet- og grundmodeller på en eller anden måde kan opgøre forventede og estimerede tal for alvorligheden. Og man kan lade tallene indgå i udpegning af sorte pletter, valg af vej- og krydstype, vurdering af forventet uheldsbesparelse samt før-efter uheldsevaluering.

3. Uheldsmodeller, et indblik

Der findes mange typer af uheldsmodeller, der kan forklare forskellige sammenhænge. Interessen i denne rapport falder på uheldsmodeller, der kan angive, hvor mange uheld og personskader, der kan forventes at være sket i en given periode på en strækning eller i et kryds med given udformning og trafikmængde, samt kan estimeres for samme sted, når der haves oplysninger om rapporterede uheld og personskader. Med et estimeret antal uheld menes uheldsniveauet, når der er taget højde for den tilfældige variation i de rapporterede uheld.

3.1 Nuværende danske uheldsmodeller

Som nævnt i kapitel 1 er der i Danmark udarbejdet uheldsmodeller, der normalt benævnes basismodeller, og disse har følgende udtryk:

$$\text{Strækning: } UHT = a \times N^P \quad (1)$$

$$\text{Kryds: } UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2}, \quad (2)$$

hvor UHT er uheldstætheden, N er ÅDT på vejen i begge retninger, N_{pri} er ÅDT ind i krydset fra primærvejen, N_{sek} er ÅDT ind i krydset på sekundærvejen, og a , P , P_1 og P_2 er konstanter.

Primærvejen i krydset er de to vejben uden vigepligt i vigepligtsregulerede kryds, eller de to vejben i signalregulerede kryds, der har den største trafik. Hvis sekundærvejens ÅDT i 3-benede kryds er mindre end 250, så udelades krydset af uheldsmodeller for kryds, og indgår i stedet i modeller for strækninger. Tilsvarende for kryds med 4 eller flere ben udelades krydset af uheldsmodeller for kryds, hvis sekundærvejens ÅDT er mindre end 500.

De seneste danske uheldsmodeller er baseret på uheldsdata for årene 2001-2005 og trafikdata for 2005 (*Hemdorff, 2006*). Modellerne er opdelt i lokaliteter med og uden randbebyggelse samt i kryds og strækninger. I tabel 4 og 5 på næste side er vist vej- og krydstypeinddelingen, der er anvendt for de danske uheldsmodeller. Modellerne er udarbejdet ud fra hhv. personskadeuheld samt person- og materiel-skadeuheld. Nye eller ændrede strækninger og kryds, hvor der foreligger færre end 3 års uheldsdata er udeladt.

Desuden benyttes en vægtning, så ekstreme værdier (outliers) ikke får lov at tælle så meget med, da modeller skal udtrykke ”det gennemsnitlige billede” (*Hemdorff, 2006*). Som vægtning er valgt den reciprokke værdi af variansen for modeludtrykket (*Hemdorff, 1993*). Der nævnes diverse problemer med at opdele strækninger i vejtyper, da veje kan ændre udformning over korte strækninger og være forskellig i de to vejsider.

Strækningstype	Randbebyggelse	
	Med rand	Uden rand
Motorvej	-	121
Motortrafikvej	-	122
Ramper ved motorvej o. lign.	-	123
2-spor med cykelsti	-	220
2-spor uden kantbane og uden cykelsti	-	221
2-spor med kantbane og uden cykelsti	-	222
2-spor med kantbane og med cykelsti	211	-
2-spor uden kantbane og med cykelsti	212	-
2-spor uden cykelsti	213	-
3-spor	310	320
4-spor	410	420
Øvrige veje	910	920

Tabel 4. Strækningstypeinddeling i danske uheldsmodeller.

Krydstype	Regulering	Randbebyggelse	
		Med rand	Uden rand
3-ben	Signalreguleret	511	521
	Primær + sekundær kanaliseret	512	522
	Primær kanaliseret	513	523
	Sekundær kanaliseret	514	524
	Ingen kanalisering	510	520
4-ben	Signalreguleret	611	621
	Primær + sekundær kanaliseret	612	622
	Primær kanaliseret	613	623
	Sekundær kanaliseret	614	624
	Ingen kanalisering	610	620
Rundkørsel		700	
Flettestrækning		800	
Andet		900	

Tabel 5. Krydstypeinddeling i danske uheldsmodeller.

Udover tal for ap-værdier beskrives modeller ved en korrelationskoefficient, R^2 , og gennemsnittet af uheld pr. kryds eller km vej. En ”traditionel” korrelationskoefficient defineres til at være andelen af den samlede variation, som den pågældende model eller variabel forklarer. R^2 kan derfor være mellem 0 og 1. Flere har fundet, at de traditionelle korrelationskoefficienter ikke er hensigtsmæssige mål for, hvor god eller præcis en uheldsmodel er (Fridstrøm et al., 1995; Maher og Summersgill, 1996; Miaou et al., 1996). I tabel 6 er givet et par eksempler på modelresultater for danske uheldsmodeller.

Kryds / strækninger Uden rand	AP- Type	Antal kryds eller km vej	Antal uheld 2001-2005	A	P	P ₁	P ₂	Korrel- ation	Uheld pr. år pr. kryds eller km vej
4-ben signalreguleret	621	263	1.489	0,000382		0,31	0,60	0,68	1,14
Motorvej	121	1.017	3.266	0,000022	1,00			0,91	0,64

Tabel 6. Eksempler på parameterværdier for uheldsmodeller hhv. 4-benede signalregulerede kryds og motorveje uden rand, baseret på person- og materielskadeuheld 2001-2005 (Hemdorff, 2006).

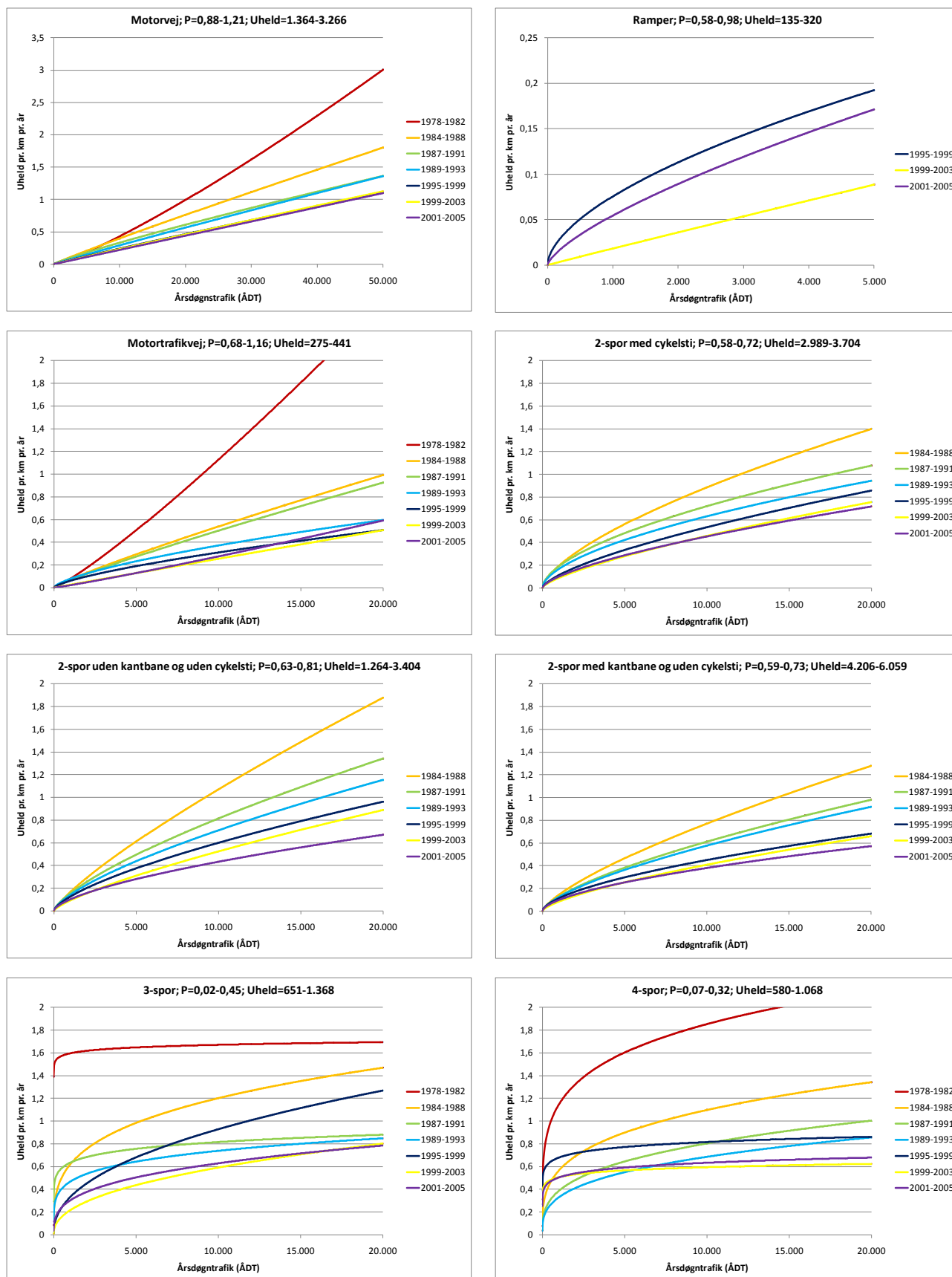
De danske uheldsmodeller er lette at forstå og nemme at bruge, idet nogle få tryk på en lommeregner kan give et resultat for det forventede uheldsniveau. Hvis det forventede uheldsniveau ønskes beregnet for et større vejnet vil en pc lette arbejdet med beregningerne.

I figur 3 og 4 på de næste sider er resultater for det forventede uheldsniveau ved brug af danske modeller for forskellige uheldsperioder. Ses på strækningstyper i figur 3, så er det tydeligt, at modeller for nogle strækningstyper er mere ”stabile” end andre. I hovedtræk kan man sige, at jo flere strækninger og uheld, som modellerne er bestemt ud fra, desto mere ”stabile” er de. Ses på modeller, der er baseret på mange km strækninger og mange uheld, f.eks. de tre modeller for 2-sporede veje, kan man ved sammenligning med andre modeller konstatere, at P-værdien for modeller for 2-sporede veje kun varierer relativt lidt fra uheldsperiode til uheldsperiode. P-værdien varierer med en faktor på 1,2-1,3 for 2-sporede veje, 4,6 for 4-sporede veje og 23 for 3-sporede veje. For motorveje varierer P-værdien med en faktor 1,4, mens den er 1,7 for motortrafikveje og ramper.

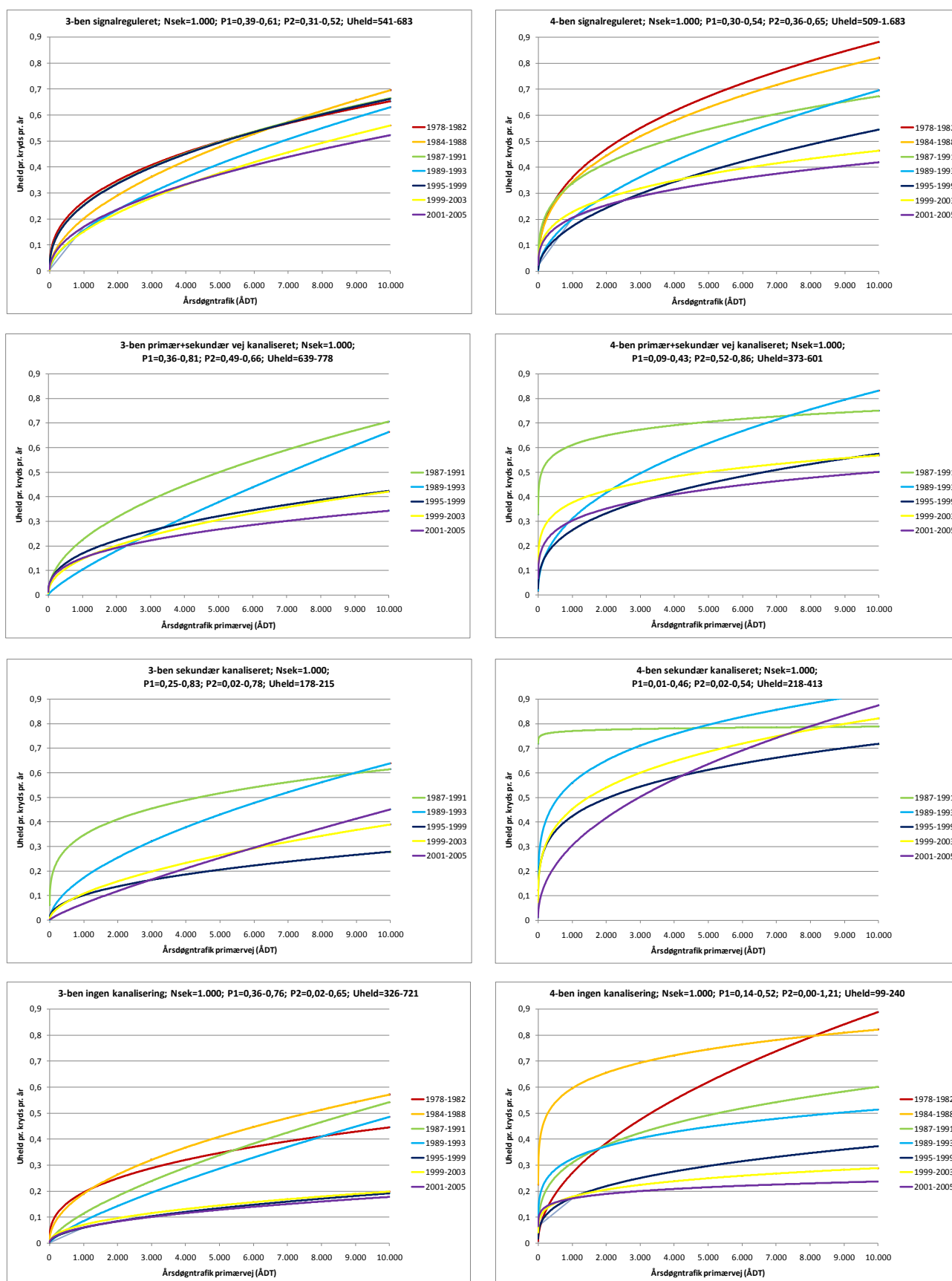
Særligt for 3-4 sporede veje er angivelse af et forventet antal uheld forbundet med stor usikkerhed. Den store variation i P-værdien skyldes måske en kombination af stor variation i geometrisk udformning og trafikregulering samt relativt få strækninger og uheld. En ny måde at modellere 3-4 sporede strækninger er nødvendig. Det kan også være relevant at inddrage andre variable i modeller for strækninger, f.eks. hastighedsbegrænsning, forekomst af midterrabat eller midterafmærkning samt adgangsforhold.

Ses på krydstyper i figur 4, så fås indtrykket, at den mest sikre krydstype er forskellig fra uheldsperiode til uheldsperiode. Det skyldes, at P-værdierne varierer meget og modellerne er baseret på relativt få uheld. P-værdien (den der varierer mest af p₁ eller p₂ fra uheldsperiode til uheldsperiode) varierer med 1,7-1,8 for signalregulerede kryds, 2,2-4,8 for kryds med kanalisering på både primær- og sekundærvej, og med 30-50 for andre kryds.

Ligesom med 3-4 sporede veje er det nødvendigt at modellere ej signalregulerede kryds på en anden måde end hidtil. Det kan være også relevant at inddrage andre variable i modellerne.



Figur 3. Forventede uheldsniveauer ifølge danske uheldsmodeller for strækninger i åbent land (uden rand) baseret på uheld fra 1978-2005.



Figur 4. Forventede uheldsniveauer ifølge danske uheldsmodeller for kryds i åbent land (uden rand) baseret på uheld fra 1978-2005.

Hvis man skulle angive nogle simple måder at opnå relativt store forbedringer af de ovenstående danske uheldsmodeller, så kunne man:

- Bruge en Negativ binomial (Poisson-gamma fordeling) modellering og empirical Bayes (EB). Det vil ændre modeludtrykket og betyde, at residualerne (forskellen mellem forventet og rapporteret uheldstæthed) vil blive beskrevet bedre, og et estimeret antal uheld kan beregnes.
- Angive en gennemsnitlig uheldsalvorlighed, f.eks. dræbte, alvorlige og lette skader pr. uheld, for de uheld, som uheldsmodellen er baseret på. Derved skal særskilte personskademodeller ikke opstilles, da det antages, at den gennemsnitlige uheldsalvorlighed ikke afhænger af trafikmængde.
- ændre modelstrukturen, så geometriske og reguleringsmæssige forhold indgår som variabel. Derved overgår man fra basismodeller til faktormodeller, f.eks. ved at de fire forskellige varianter af kanalisering i hhv. tre- og firebenede kryds optræder som en variabel i én faktormodel frem for, at der opereres med fire forskellige basismodeller. Det vil medføre mere stabile P-værdier, da sammenhængen mellem trafikmængde og uheldsforekomst vil blive bedre beskrevet, og det vil give mere pålidelige modeller.

Til de ovenstående simple forbedringer bør tilføjes, at den gennemsnitlige uheldsalvorlighed faktisk afhænger af trafikmængden, da stigende trafikmængder typisk giver mindre alvorlige uheld. De ovenfor nævnte faktormodeller kaldes også for kovariatmodeller og ser således ud:

$$\text{Strækning: } UHT = a \times N^P \times e^{\sum_{i=1}^n b_i x_i} \quad (3)$$

$$\text{Kryds: } UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2} \times e^{\sum_{i=1}^n b_i x_i} \quad (4)$$

I faktormodeller indgår konstanterne b_i og variablene x_i , der beskriver stedets udformning (f.eks. kanalisering, forekomst af kantbane og cykelsti, osv.). Det anses for uhensigtsmæssigt at operere med vej- eller krydstyper med meget forskellige P-værdier i samme faktormodel. Faktormodeller kan desuden være problematiske, idet der kan være betydelig samvariation mellem de ellers uafhængige variable.

3.2 State-of-the-art uheldsmodeller

I nærværende afsnit beskrives, hvordan man kan udarbejde uheldsmodeller på en relativ optimal facon. Forskningsområdet med uheldsmodeller er under stadig udvikling og nye idéer til modeludtryk osv. afprøves jævnligt. Afsnittet inddrager uheldsmodeller mv., der er velafprøvet og som synes at fungere godt i praksis.

3.2.1 Variable

Antallet af uheld (uheldstæthed) bruges i de fleste moderne uheldsmodeller som den afhængige variabel. Uheldsfrekvens bruges (uheld pr. eksponeringsenhed f.eks. kørt km) derimod sjældent i dag som den afhængige variabel. Korrelationen mellem trafikmængde og uheldstæthed er ikke-lineær. Stiger trafikmængden med 10 procent, så stiger uheldstætheden ikke med 10 procent, men ofte med en lavere procentsats. Derfor er det uhensigtsmæssigt at anvende uheldsfrekvensen som afhængig variabel.

Eksponering indgår i stedet som en uafhængig variabel. Eksponering kan opgøres på ufattelig mange måder, men det mest almindelige er årsdøgntrafik (ÅDT) enten på strækningen eller indkørende til krydset. I realiteten er det sjældent muligt at opgøre eksponering anderledes, fordi andre eller mere detaljerede data ikke forefindes eller ikke er gjort tilgængelige. Eksponeringen kunne ellers opdeles på tid (måned, dag, time, osv.), trafikstrøm (retning, manøvre, osv.), transportform (bil, motorcykel, knallert, cykel, fodgænger, osv.), mv. Det er i dag kun almindeligt at opdele eksponeringen på år samt ben i kryds.

Det har vist sig, at sammenhængen mellem uheldstæthed og ÅDT oftest bedst lader sig beskrive ved en potensfunktion. Derfor, som angivet tidligere, arbejdes oftest med basis- og faktormodeller, der ser således ud:

Basismodeller:

$$\text{Strækning: } UHT = a \times N^P \quad (1)$$

$$\text{Kryds: } UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2}, \quad (2)$$

Faktormodeller:

$$\text{Strækning: } UHT = a \times N^P \times e^{\sum_{i=1}^n b_i x_i} \quad (3)$$

$$\text{Kryds: } UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2} \times e^{\sum_{i=1}^n b_i x_i} \quad (4)$$

hvor UHT er uheldstætheden, N er ÅDT på strækningen, N_{pri} er ÅDT ind i krydset fra primærvejen, N_{sek} er ÅDT ind i krydset på sekundærvejen, variablene x_i beskriver stedets udformning, samt a, P, P_1 , P_2 og b_i er konstanter.

Til forskel fra basismodeller indgår andre uafhængige variable end ÅDT i faktormodeller. De andre uafhængige variable kaldes ofte for risikofaktorer og må nødvendigvis have en stor betydning for trafiksikkerheden for at være relevante at medtage. Forskningen viser, at der er mange forhold, som har stor betydning for trafiksikkerheden (*Elvik et al., 2009*). Det er eksempelvis a) vejens eller krydsets tværsnit, tracé, regulering og kontrol, b) forhold som vejr, føre og lysforhold, c) trafikantadfærd som hastighed, træthed, selebrug og påvirkning af alkohol, medicin og stoffer, samt d) andre mål for eksponering. Men valget af uafhængige variable afgøres næsten udelukkende af forhåndenværende datatilgængelighed.

Basismodeller kan være fornuftige at opbygge, men hvis der ikke er tilstrækkeligt med data for den aktuelle designvariant til at få en rimelig sikker bestemmelse af konstanterne i basismodellen, så er man nødt til at lade flere varianter indgå. Her kan det være hensigtsmæssigt med en faktormodel, hvor variablene, x_i , kan beskrive forskelle mellem varianterne.

Uafhængige variable må nødvendigvis kunne måles eller klassificeres pålideligt og ensartet. Det er vigtigt at undgå, at uafhængige variable udviser en stor grad af samvariation eller endogenitet. Alternativt må samvariationen eller endogeniteten håndteres. Kraftig samvariation kan håndteres på to måder. Enten udelades én af de samvarierende uafhængige variable eller observationer udelades, så samvariationen reduceres eller helt undgås. En metode med vægtning af observationer kan ikke anbefales til at håndtere samvariation i uheldsmodeller. Endogene uafhængige variable kan også håndteres på flere måder f.eks. ved eksogene instrumentvariable, vektorautoregressive og strukturelle modeller. Kim og Washington (2006) opstillede en uheldsmodel og en model for den endogene variabel (forekomst af venstresvingsbane) samtidigt (simultaneous equation models), hvorved uheldsmodellen tager højde for den endogene variabels afhængighed af uheldstætheden. Simultan modellering kan ikke udføres med almindelig statistik software, men skal programmeres særskilt. I stedet for denne tilgang kunne en eksogen instrumentvariabel, f.eks. mængden af venstresvingende trafik, været anvendt.

Et eksempel på en stærk samvariation kunne være mellem ÅDT og antallet af kørespor. Hvis begge variable indgår i en model, så vil UHT være pålidelig, når ÅDT er i nærheden af gennemsnittet for hhv. veje med 2, 3, 4, osv. kørespor. Er ÅDT nær kapacitetsgrænsen for en 2-sporet vej eller meget lav for en 4-sporet vej, så kan modellen med begge variable give meget upålidelige resultater, og de kan være langt mere upålidelige end modellen, hvor kun ÅDT indgår. Derfor er man nødsaget til at håndtere stærk samvariation mellem uafhængige variable.

Et eksempel på en særdeles endogen variabel kunne være advarselstavlen ”Farligt vejkryds” (A11). Hvis den variabel indgår i en uheldsmodel for kryds, så vil man få opfattelsen, at tavlen betyder en voldsom stigning i uheldstætheden. Det vil samtidig betyde, at ÅDT’s betydning for UHT bliver undervurderet (P-værdi for lav). Det vil betyde, at kryds med stor trafikmængde og uden advarselstavlen vil have en undervurderet UHT, mens kryds med beskeden trafikmængde og uden advarselstavlen vil have en overvurderet UHT.

Enheden for uheldstætheden, UHT, er typisk uheld pr. år pr. km for strækninger, mens den for kryds er uheld pr. år pr. kryds.

Hvis strækningerne til opstilling af uheldsmodellen har varierende længde, tilføjes længden i km til formeludtrykket. Derved ændres udtrykket for formel 1 til: $UHT_L = a \times L \times N^P$, hvor L er strækningens længde. For kryds opereres med længden én uanset, hvor store kryds defineres til at være.

Strækninger og kryds til opstilling af uheldsmodeller kan have varierende eller meget lange uheldsperioder. Det kan håndteres på flere måder, og er samtidigt et spørgsmål om, hvordan modellen skal anvendes i praksis.

Hvis man har tilstrækkeligt med uheldsdata i en relativ kort tidsperiode, f.eks. 2007-2009, og uheldsdata for hele perioden indgår for alle observationer, samt den forventede uheldstæthed skal repræsentere hele perioden, så kan man blot tilføje periodelængden (T) på samme måde som strækningslængden. Formel 1 får derved udtrykket: $UHT_{L,T} = a \times L \times T \times N^P$.

Hvis man derimod bruger en relativ lang periode, og ønsker at kunne angive forventede uheldstal for et enkelt eller få år, så kan man lade en årsfaktor, a_t , for hvert år indgå. Det vil så kræve, at modellen bygges op med uheldsdata og ÅDT opdelt efter år. Der skal foreligge tal for trafik hvert år – målt, interpoleret eller modelberegnet. Formel 1 kan så formuleres til f.eks.: $UHT_{L,t} = a \times L \times a_t \times N^P$. Observationer opstilles derved som paneldata, hvor både forskelle mellem år for hver strækning eller kryds samt forskelle mellem strækninger eller kryds indgår i estimering af konstanter.

I en model med årsfaktorer antages, at konstanten P er ens for alle årene. Det kan være en dårlig antagelse. Eksempelvis vil overvægt af trafiksikkerhedstiltag på veje med store trafikmængder medføre, at P bliver lavere med årene. Man kan også forestille sig en ændret trafiksammensætning eller trafikantadfærd kan føre til en anden P -værdi. Med baggrund i danske uheldsmodeller er der ingen grund til at tro, at P -værdierne reelt har ændret sig gennem tiden. Man skal dog være påpasselig med modeller med årsfaktorer, da observationer er korrelerede over tid (den enkelte strækning indgår flere gange), og samtidig kan metodikken medføre masser af observationer uden uheld, hvilket kan føre til lavere præcision eller fejl i estimeringen af konstanterne.

På samme måde som med årsfaktorer kan områdefaktorer indgå. En områdefaktor kan tage højde for, at f.eks. kryds i en region generelt er mere sikre end i en anden region. Det kan f.eks. skyldes en anden rapporteringsgrad for uheld, andre drifts-rutiner og anden sammensætning af køretøjer eller førere. En anbefaling ved estimering af årsfaktorer og områdefaktorer er, at der i hvert år eller hvert område mindst bør være indtruffet 75-100 uheld (AASHTO, 2010).

For at håndtere korrelation mellem observationer over tid eller i samme område kan man ”ligge en model ind over sin model”. Det kan eksempelvis gøres med Generalized Estimating Equation (GEE) modeller (Lord og Mannering, 2010), som indgår i GENMOD og en række andre procedurer i SAS. GEE modeller kan kun håndtere få eller ingen udeladte data, hvorfor f.eks. kryds ikke kan udelades i et antal år pga. nyanlæg, vejarbejde, osv. GEE er metodisk beskrevet af Lord og Persaud (2000). En god gevinst ved at håndtere korrelation mellem observationer er, at en større del af den systematiske variation bliver beskrevet, og en model med årsfaktorer vil typisk have lidt anderledes a - p -værdier end en model uden

årsfaktorer, som er baseret på samme data for et gennemsnit år (*Mountain et al., 1998; Lord og Persaud, 2000*). Vistisen (2002) udførte flere uheldsmodeller baseret på hierarchical generalized linear models (HGLM) med Poisson-Gamma fordeling, hvor tid blev behandlet som en kontinuer variabel, $a^{(\text{årstal} - \text{årstal for basisår})}$, og a blev estimeret til mellem 0,97 og 1,11 i modellerne. Den måde at håndtere årsvariation kan ikke anbefales, da trafikikkerheden svinger anderledes fra år til år. Årsvariationen bør håndteres med brug af dummy-variable. Men HGLM er et godt alternativ til GEE modeller, da udeladte data ikke synes at være et problem. Dog er der sået tvivl om den teoretiske basis og præcision ved HGLM (*Lee et al., 2006*).

Da datamængden i Danmark er relativ beskeden, kan det være relevant at foretage modeltekniske valg, der kan øge antallet af observationer. Det mest oplagte er at overgå fra basis- til faktormodeller samt øge tidsperioden med data fra 5 til flere år og samtidigt indføre årsfaktorer. Omfanget af uafhængige variable vil i praksis blive afgjort af tilgængeligheden til data.

3.2.2 Funktionsudtryk

Uheldstætheden bør aldrig kunne blive nul eller negativ i en uheldsmodel, hvis trafikanter er eksponeret, altså N større end nul. Det skyldes, at der altid er risiko ved at færdes på vejene. De tidligere nævnte modeller 1 – 4 er multiplikative, og de enkelte led er positive, så de kan ikke give nul eller negative resultater. Disse modeller er derfor logiske.

Der findes andre funktionsudtryk, som på samme eller næsten samme måde er logiske. Miaou og Lord (2003) afprøvede seks funktionsudtryk på data for 4-benede lyskryds i Toronto, Canada, i alt 868 kryds og 54.989 uheld, hvor ÅDT på primærvejen varierede mellem 5.300 og 72.300 og mellem 52 og 42.600 på sekundærvejen (bemærk model B er lig den tidligere nævnte model 2):

$$UHT = a \times (N_{\text{pri}} + N_{\text{sek}})^P \quad (\text{A})$$

$$UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2} \quad (\text{B})$$

$$UHT = a \times (N_{\text{pri}} \times N_{\text{sek}})^P \quad (\text{C})$$

$$UHT = a \times (N_{\text{pri}} + N_{\text{sek}})^{P_1} \times \left(\frac{N_{\text{sek}}}{N_{\text{pri}}}\right)^{P_2} \quad (\text{D})$$

$$UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2} \times e^{P_3 \times N_{\text{sek}}} \quad (\text{E})$$

$$UHT = N_{\text{pri}} \times e^{(a_1 + P_1 \times N_{\text{sek}})} + N_{\text{sek}} \times e^{(a_2 + P_2 \times N_{\text{pri}})} \quad (\text{F})$$

Disse seks modeller peger på nogle væsentlige egenskaber ved uheldsmodeller for kryds. Model B, C, D og E kan ved en trafik på 0 på en af vejene kun give nul uheld. Det opfattes oftest som en god egenskab, men Miaou og Lord påpeger, at krydset har en udbredelse og altså en længde på begge veje. Derfor vil det være rimeligt ved en trafik nær 0, at antallet af uheld svarede til en model for en strækning med den længde, som krydset har. Model A og F giver en uheldstæthed over nul. Miaou og Lord (2003) viser, at model A overestimerer uheldstætheden voldsomt ved lave trafikmængder på sekundærvejen, mens model F synes at give en god estimering.

Ved almindelige trafikmængder, altså hverken meget lave eller meget høje, synes model A, B, C, D, E og F ikke at afvige meget fra hinanden, og derfor vurderes de næsten ens ud fra et statistisk synspunkt. Ved meget høje trafikmængder på en eller begge veje er der lige som ved meget lave trafikmængder en større afvigelse mellem modellerne. Dataanalysen kunne tyde på, at model B afviger mindst ved meget høje trafikmængder, men det er lidt uklart i artiklen af Miaou og Lord.

Modeller af type B vil nærmest altid fungere bedre end type A og C (se f.eks. *Miaou og Lord, 2003; Bélanger, 1994*).

Artiklen af Miaou og Lord giver indblik i to problemstillinger ved de almindelige uheldsmodeller (model B, som også er model 2). Den ene er lave trafikmængder, som denne model ikke er god til at håndtere, hvorfor en nedre grænse for trafik er væsentlig. Det problem må nødvendigvis betragtes sammen med definitionen på krydsets udbredelse, som derfor bliver et andet problem. Man kan sige, at jo større kryds defineres til at være, desto højere bør grænsen for trafik være. Så en løsning er at operere med nedre grænser for trafik og definitioner for krydsudbredelse ved uheldsmodeller for kryds.

En anden løsning kunne være additive modeller som model F, som samtidig er logisk, da leddene i modellen kun kan være nul eller positive. Egentligt kunne man blot addere model 1 og 2, og derved løse de to nævnte problemer:

$$UHT = a_1 \times N_{pri}^{P_1} \times L_1 + a_2 \times N_{sek}^{P_2} \times L_2 + a_3 \times N_{pri}^{P_3} \times N_{sek}^{P_4} , \quad (5)$$

hvor L_1 og L_2 er længden af strækningen med krydset på hhv. primær- og sekundærvejen, mens a_1 , a_2 , a_3 , P_1 , P_2 , P_3 og P_4 er konstanter. Ved at anvende log-lineær modellering (der tages logaritme på begge sider af lighedstegnet) sikre man sig, at a_1 , a_2 og a_3 kun kan være positive.

Det vil være problematisk at estimere konstanterne i model 5, hvis der er mange typer af strækninger og kryds blandt observationer. Et andet problem er, at trafikmængden sjældent er den samme på begge sider af krydset.

I nærværende rapport er der særligt fokus på uheldsmodeller for veje i åbent land. Her kunne en løsning være at estimere modeller for "rene" strækninger først. De

”rene” strækninger skulle være beliggende med en vis minimumsafstand fra kryds. En minimumsafstand kunne være 50-200 meter ved almindelige kryds, og f.eks. 100-400 meter ved flettestrækninger på motorveje. Herefter modelleres kryds ved at benytte model 5 (alternativt model 2), hvor der ses på krydset og strækninger op til krydsene indenfor den førnævnte minimumsafstand. Men i stedet for at estimere konstanter for ”strækningsled”, altså a_1 , a_2 , P_1 og P_2 , så bruges de allerede estimerede modeller for ”rene” strækninger til at beregne forventede antal uheld for strækningerne op til krydsene. Model 5 bruges derved til at estimere, hvor mange uheld der sker yderligere i og ved kryds set i forhold til strækninger.

Additive modeller giver også mening at anvende i andre sammenhænge, selvom de kan give anledning til ulogiske resultater (Hauer, 2004; Hauer et al., 2004). Det skyldes, at kryds og strækninger inkluderer forhold, der øger eller reducerer antallet af uheld, hvis antal delvist er uafhængig af uheldsfrekvensen i øvrigt. Det kan f.eks. være ud-/indkørsler, faste genstande, ”Din fart” tavler, osv. på strækninger, og i kryds kan det f.eks. være separate svingfaser, oversigtsforhold, osv. De forhold kan behandles i faktormodeller (model 3-4) ved, at f.eks. antallet af ud-/indkørsler på strækninger indgår som en variabel. Ved at gøre det antages, at uheld som følge af ud-/indkørsler har samme relation til trafikmængden og de andre uafhængige variable, f.eks. køresporsbredde, som de andre uheld på strækningen. Det er næppe tilfældet. I stedet kan et additivt led tilføjes:

$$UHT = a \times N^P + a_{ud-ind} \times N^{P_{ud-ind}} \times x_{ud-ind} ,$$

hvor a_{ud-ind} er konstanten for det additive led med ud-/indkørsler, mens x_{ud-ind} repræsenterer antallet af ud-/indkørsler, og $N^{P_{ud-ind}}$ angiver det additive leds særlige relation til strækningens trafik.

Additive led kan også være relevante til at få uheld med bestemte trafikantgrupper bedre estimeret. Det kan også være additive led for specifikke uheldstyper. Hvis man har trafiktal for cyklister på strækninger, så kan de indgå på flere måder f.eks. følgende fire:

I) En multiplikativ model:
$$UHT_{alle} = a \times N_{bil}^{P_1} \times N_{cykel}^{P_2}$$

II) En additiv model:
$$UHT_{alle} = a_1 \times N_{bil}^{P_1} + a_2 \times N_{cykel}^{P_2} \times N_{bil}^{P_3}$$

III) To multiplikative modeller:
$$\begin{aligned} UHT_{alle} &= UHT_{cykel} + UHT_{andre} \\ UHT_{cykel} &= a \times N_{bil}^{P_1} \times N_{cykel}^{P_2} \\ UHT_{andre} &= a \times N_{bil}^P \end{aligned}$$

IV) En nominal bivariat model:
$$UHT_{alle}(UHT_{cykel}, UHT_{andre}) = \dots$$

De to første af de ovenstående modeller har hver en væsentlig ulempe. Model I) er problematisk, fordi når cykeltrafikken er nul, så forudses ingen uheld, selvom der er masser af biltrafik. Man står altså i de samme problemer som med modeller for kryds, og derfor er der brug for nedre grænser af trafik eller at omdanne en trafik-art til en kategorisk variabel. Model II) kunne give ulogiske resultater, hvis f.eks. cykeltrafik gav anledning til færre uheld. Men det gør cykeltrafik ikke, så den model kunne godt være logisk. Det er dog vigtigt at sørge for, at leddene i model II) ikke kan blive negative.

Model III) og IV) er forskellige fra I) og II) ved, at uheld opsplittes i cykeluheld og andre uheld. Det vil sige, at model III) og IV) er baseret på dobbelt så mange observationer som I) og II), hvilket kan føre til en bedre estimering, men ikke behøver at gøre det. Bi- og multivariate modeller er beskrevet i afsnit 3.2.5. Til forskel fra III), så indgår fordelingen mellem cykeluheld og andre uheld i model IV), hvilket kan forbedre estimeringen, men ikke behøver at gøre det.

Geedipally og Lord (2010) har sammenlignet modeltype I), III) og IV) dog for en opsplitning i ene- og flerpartsuheld samt kun med eksponering af motoriserede køretøjer. Deres resultater viser, at det ikke er muligt på forhånd at sige, hvilken type model der er bedst.

At operere med additive led giver flere muligheder for at sammensætte modeller, men samtidig øges mulighederne for fejl i estimeringen og fejl ved anvendelse af modellen.

En anden mulighed er at operere med synergieffekter. Eksempelvis viser studier, at sikkerhedseffekten af køresporsbredden afhænger af trafikmængden og sikkerhedseffekten af faste genstande afhænger af hastighedsbegrænsningen. En sådan synergieffekt kan modelleres på mange måder, f.eks.:

$$UHT = a \times N^P \times e^{\Sigma(b_1 x_{ks} + b_2 x_{ks} \times \ln(N^{P_{ks}}))},$$

hvor b_1 og b_2 er konstanter, x_{ks} er køresporsbredden og $N^{P_{ks}}$ er trafikmængden opløftet til en særlig konstant for synergiledet. Med ovenstående model opnås, at sikkerhedseffekten af køresporsbredde kan afhænge af trafikmængden.

Ligesom med additive led, så øges mulighederne for at sammensætte modeller markant med led for synergieffekter. Men den større kompleksitet i opbygningen af modellen kan føre til flere fejl i estimeringen og anvendelsen.

En måde at simplificere modeller samt undgå uheldige estimater og anvendelse heraf er ved at omdanne kontinuere variable til kategoriske variable. Kontinuere variable kan være lineære, ikke-lineære og have maksima og minima. Et problem med model 1-4 er, at der kun indgår potens- og eksponentielle funktioner, som er monotone og ikke tillader lokale maksima og minima. Hvis køresporsbredden atter bruges som eksempel, så vides, at ved høje trafikmængder er sikkerheden

optimal ved en bredde på 3,5 – 3,7 meter. I stedet for at bruge f.eks. et uheldigt andengradsled, $b_3(x_{ks})^2$, så kan køresporsbredde ændres til en kategorisk variabel, så konstanten b_i antager en værdi for hver breddekategori.

Som nævnt kan det være et problem, at der i model 1-4 kun indgår potens- og eksponentielle funktioner, som er monotone, hvorved lokale maksima og minima ikke er mulige. I de senere år er almindelig pc'er blevet så kraftige og statistik software så avanceret, at man kan opbygge modeller ved brug af transformationer og spline-teknikker, der giver rige muligheder for lokale maksima og minima samt nærmest alle mulige funktionsudtryk (*Elvik, 2007*). Problemet med disse transformationer og spline-teknikker er ligesom med additive- og synergiled, at det øger modellens kompleksitet, der kan give anledning til uforståelige sammenhænge, men samtidig øges chancerne for fejl i estimering og anvendelse.

Umiddelbart anbefales at afholde sig fra brug af additive led og synergieffekter, såfremt der ikke forefindes pålidelig dokumentation om, hvordan de bør se ud. Det anbefales at gøre brug af model 1-4, hvor faktormodellerne kan indeholde både kontinuere og kategoriske variable. Det bør undersøges nærmere, om det er bedst med a) model 2 (eller 4) med tilhørende nedre grænser for trafik samt klare definitioner for krydsets udbredelse, eller b) model 5 med en anderledes definition af kryds og strækninger.

3.2.3 Residualer og empirical Bayes

For at foretage regressionsanalyse, der resulterer i bestemmelse af konstanterne i uheldsmodellen, så tages først logaritmen af udtrykket på begge sider af lighedstegnet. Formel 1-4 omskrives således til log-lineære modeller:

Basismodeller:

Strækning: $Ln(UHT) = a + P \times \ln(N)$

Kryds: $Ln(UHT) = a + P_1 \times \ln(N_{pri}) + P_2 \times \ln(N_{sek})$

Faktormodeller:

Strækning: $Ln(UHT) = a + P \times \ln(N) + \sum_i^n b_i x_i$

Kryds: $Ln(UHT) = a + P_1 \times \ln(N_{pri}) + P_2 \times \ln(N_{sek}) + \sum_i^n b_i x_i$

Regressionsanalysen tager udgangspunkt i en fordeling af residualer. Almindelige regressionsmodeller, både lineære og ikke-lineære, hvor der er antaget en normalfordeling og homoscedastiske variable (homogene varianser), er meget ringe til at modellere uheldstæthed. Flere har fundet, at estimering med almindelige regressionsmodeller ikke er hensigtsmæssig, da uheldsgenereringsprocessen synes at have en slags Poisson fordeling (se f.eks. *Maycock og Hall, 1984; Jovanis og Chang, 1986; Joshua og Garber, 1990; Miaou og Lum, 1993; Turner og Nicholson, 1998*). Det er især ved forventede uheldstætheder under 1 eller ved modeller baseret på relativt få observationer, at almindelige regressionsmodeller kan være meget misvisende.

Ashton (1966) forsøgte sig med en Poisson-gamma fordeling – dvs. en Poisson fordeling for uheldsforekomsten det enkelte sted og med en gamma (negativ binomial) fordelt variation mellem steder. Satterthwaite (1976) og Abbess et al. (1981) viste, at Poisson-gamma fordelingen er god til at beskrive variationen i uheldsforekomsten. I disse tidlige studier opstillede man ikke uheldsmodeller med baggrund i stedspecifikke data som ÅDT, geometri, osv. Maycock og Hall (1984) opstillede modeller for rundkørsler med baggrund i en Poisson fordeling. Derved beskrev de ikke den uforklarede systematiske variation mellem stederne. Maycock og Hall brugte generalized linear models (GLM), se evt. McCullagh og Nelder (1989). Maycock og Hall antog så at sige, at deres modeller var ”perfekte” – altså beskrev al systematisk variation. Hauer og Persaud (1987) opstillede en Negativ binomial (NB) model, altså en Poisson-gamma fordeling, igen baseret på GLM. NB modeller er siden da blevet til den mest accepterede og benyttede type af uheldsmodel (se f.eks. Hauer, 1997; Lord og Mannering, 2010).

I Danmark er Vejdirektoratets uheldsmodeller blevet estimeret med baggrund i en Poisson fordeling (Vistisen, 2002). Der er også opstillet modeller for bygader med baggrund i en Poisson fordeling (Greibe og Hemdorff, 1995 og 1998). Vistisen (2002) opstillede NB uheldsmodeller for strækninger og kryds med baggrund i en Poisson-Gamma fordeling.

Forud for opstilling af modeller må man vide, om variationen i uheldsforekomsten over- eller underspredt, dvs. om den statistiske variation er større eller mindre end man kunne forvente i en simpel model. En simpel model har blot en Poisson fordeling, hvor middelværdien sættes lig med variansen (spredningsparameteren er lig 1). Hvis middelværdien er noget forskellig fra variansen, må en anden fordeling end Poisson vælges, da modellens konstanter ellers bliver fejlestimeret. Samtidig vil resultater af de statistiske analyser såsom konfidensintervaller og testværdier blive ugyldige eller misvisende. Det er særligt ved lave forventede uheldstætheder, at Poisson modeller kan være misvisende.

Ved beregning af middelværdi og varians tages højde for antal år og længde af strækning, hvis disse er forskellige blandt observationer. Middelværdien (måske bedre kendt som gennemsnit) beregnes ved:

$$\mu = \frac{\sum y}{n}, \text{ og variansen beregnes ved: } Var(Y) = \frac{\sum (y - \mu)^2}{n},$$

hvor μ er middelværdien, y er den rapporterede uheldstæthed for den enkelte observation, n er antallet af observationer, og $Var(Y)$ er variansen. Når man har opstillet en model, så er middelværdien, μ , lig den forventede uheldstæthed (UHT).

Hauer (1997) angiver, at variation i uheldsforekomst kan opdeles i en tilfældig og en systematisk variation. Der er systematisk variation i uheldsforekomsten, når variansen er større end middelværdien, altså når variationen er overspredt. I langt de fleste tilfælde er fordelingen i uheldsforekomsten overspredt, både før og efter

uheldsmodellen er opstillet. Det betyder, at i langt de fleste tilfælde må man gøre brug af en anden fordeling end Poisson.

Der kan være mange baggrunde for overspredning, og mange har peget på flere muligheder. Der er tale om, at man ikke kan opstille den ”perfekte model”, som kan forklare hele den systematiske variation mellem observationer, altså mellem de kryds eller strækninger, der indgår. Det kan f.eks. skyldes:

- Udeladte uafhængige variable, f.eks. indgår ikke alle data for eksponering (fodgængere, påvirkede bilførere, osv.) og data for vej- eller krydsdesign.
- Fejl i uafhængige variable, f.eks. trafik talt i et ej repræsentativt tidsrum og vej- eller krydsdesign er fejlregistreret.
- Den afhængige variabel (uheldstætheden) afhænger af Politiets rapporteringsgrad for uheld – altså mørketallet – og den kan være forskellig fra sted til sted, og f.eks. afhænge af afstande til skadestuer og politistationer (altså potentielle udeladte uafhængige variable).
- Modellen kan være misspecificeret, f.eks. forkert funktionsudtryk, udeladelse af additive led og synergieffekter, osv.

Helt grundlæggende kan det siges, at overspredning er et udtryk for uheld, der sker som følge af uforklarede forhold. Overspredning er altså den del af variationen, der ikke kan forklares med tilfældig variation i uheldsforekomsten, også kaldet regressionseffekt.

Forskellige fordelinger og modeltyper har været anvendt og undersøgt i relation til uheldsmodeller (*Lord og Mannering, 2010*). Når overspredning er til stede i data, så er en Negativ binomial (Poisson-gamma fordelt) regressionsmodel oftest blevet anvendt. Der findes flere andre modeller såsom Quasi-Poisson, Conway-Maxwell-Poisson, Poisson-lognormal, Zero-inflated Poisson og Zero-inflated Negativ binomial, der kan også håndtere overspredning. Men en Negativ Binomial (NB) model synes i de fleste tilfælde at være nemmere, bedre og / eller mere korrekt at anvende (*Maher og Summersgill, 1996; Lord og Mannering, 2010; Geedipally og Lord, 2011; Lord et al., 2005; Elvik, 2007*).

I de ekstremt få tilfælde, hvor data er underspredt, er en NB model derimod ugyltig. Så kan anvendes f.eks. Conway-Maxwell-Poisson eller Gamma modeller. Underspredte data forekommer som regel, når de fleste observationer har nul uheld, få observationer har ét uheld og ingen observationer har mere end ét uheld. Det kan f.eks. være dødsulykker eller uheld ved usikrede jernbaneoverkørsler. Ellers forekommer sådanne datasæt næsten kun, når man betragter strækninger eller kryds med meget lidt trafik eller betragter steder i en ganske kort periode. Underspredte data kan som regel undgås ved at udelade meget svagt trafikerede steder eller øge periodelængden.

Conway-Maxwell-Poisson modellen bør fremhæves, fordi den kan håndtere både over- og underspredning, og faktisk kan en shapeparameter modellens, så den kan

håndtere under- og overspredning samtidigt. NB modellens spredningsparameter kan også modelleres, men altså kun varierende overspredning.

Det antages oftest, at uheld sker med en konstant frekvens pr. tidsrum i en given periode. Denne frekvens kan variere fra sted til sted, og kan variere fra periode til periode. Modeller baseret på antagelsen om en sådan uheldsgenereringsproces kaldes for single-state modeller. Poisson, NB, Poisson-lognormal og Conway-Maxwell-Poisson er single-state modeller.

I dual-state modeller, som f.eks. Zero-inflated Poisson, Zero-inflated Negativ binomial og Gamma modeller, antages derimod, at uheldsgenereringsprocessen har to tilstande, dels den almindelige hvor der sker uheld med en konstant frekvens, dels en sikker tilstand hvor der ikke sker uheld. Dual-state modeller skulle efter sigende fungere bedre i forhold til single-state modeller, når langt hovedparten af observationer har nul uheld. To problemer er dog væsentlige at bemærke ved dual-state modeller. Teoretisk set findes der ikke en sikker tilstand. Nul uheld er blot et udtryk for uhensigtsmæssig valg af periodelængde. For det andet er dual-state modeller mere komplekse, da de to tilstande får hver sine estimater af konstanter i en bivariat model, der angiver sandsynligheden for den ene og den anden tilstand.

Det anbefales at anvende NB modeller, når data er blot en smule overspredt, f.eks. når variansen er mere end 1,05 gange middelværdien, når modellen er opstillet. I de få tilfælde, hvor variansen er mindre end 1,05 gange middelværdien kan man vælge at benytte Poisson eller Conway-Maxwell-Poisson modeller.

I forbindelse med NB modeller kan omfanget af overspredning opgøres ud fra spredningsparameteren, k , der er defineret ud fra følgende formel:

$$\text{Var}(Y) = \mu \times (1 + k\mu) = \mu + k\mu^2$$

Bemærk, at når k går mod nul, så nærmer variansen sig middelværdien og derved en Poisson fordeling. Spredningsparameteren er lig med:

$$k = \frac{\frac{\text{Var}(Y)}{\mu} - 1}{\mu} = \frac{\text{Var}(Y) - \mu}{\mu^2}$$

Ved beregning af den estimerede uheldstæthed med en empirical Bayes metode (Hauer, 1997) gøres der brug af spredningsparameteren. Den estimerede uheldstæthed, λ , for et sted kan opgøres ud fra:

$$\lambda = \omega \times \mu + (1 - \omega) \times y,$$

hvor ω er en vægtfaktor, μ er den forventede uheldstæthed, der er beregnet via modellen, og y er den observerede uheldstæthed på stedet. Den reciprokke eller

inverse spredningsparameter kaldes ofte for ϕ , og $\phi = 1 / k$. Det bedste estimat for vægtfaktoren er:

$$\omega = \frac{1}{1 + k\mu} = \frac{\phi}{\phi + \mu}$$

Vil man undgå opgørelsen af vægtfaktoren, så kan den estimerede uheldstæthed opgøres ud fra:

$$\lambda = y - \frac{y}{1 + k\mu} + \frac{\mu}{1 + k\mu} = \mu \times s = \mu \times \frac{\phi + y}{\phi + \mu},$$

hvor s kaldes for spredningseffekten, og er den værdi man skal multiplicere den forventede uheldstæthed med for at få den estimerede uheldstæthed. Af ovenstående formel ses, at når k går mod nul, så nærmer den estimerede uheldstæthed, λ , sig den forventede uheldstæthed, μ . Som nævnt førhen så er k nul, når der kun er tilfældig variation i uheldsforekomsten, altså når uheldsmodellen er perfekt. Jo større k er, desto nærmere er den estimerede uheldstæthed på den rapporterede uheldstæthed. Af formelen ovenfor kan umiddelbart også udledes, at λ nærmer sig y , når μ bliver større og $y - \mu$ eller y/μ holdes konstant, hvilket vil sige, at den uforklarede systematiske variations andel aftager med stigende uheldstæthed.

Miaou og Lum (1993) opstillede flere Quasi-Poisson modeller, hvor ϕ var antaget at være lineært afhængig af den forventede uheldstæthed. Maher og Summersgill (1996) fandt, at ϕ synes at være afhængig af den forventede uheldstæthed opløftet til en potens af ca. 0,25. Det vil sige, at k falder med stigende uheldstæthed. Uheldsmodellen synes altså at blive mere ”perfekt” med stigende uheldstæthed.

Forskere har dog alligevel i mange år opfattet spredningsparameteren, som en præcist estimeret konstant uden usikkerhed. Den opfattelse er blevet udfordret, da det er blevet let de sidste ca. 5-10 år via statistik software at analysere, om og hvordan spredningsparameteren varierer. Og flere studier viser, at k næppe er konstant og i hvert fald må betragtes som usikkert bestemt (se f.eks. Heydecker og Wu, 2001; Hauer, 2001; Vistisen, 2002; Miaou og Lord, 2003; Agrawal og Lord, 2006; Geedipally et al., 2009; Geedipally og Lord, 2011).

Hauer (2001) sætter ϕ til at være afhængig af strækningens længde. Dette er selvfølgelig kun muligt at modellere, hvis strækninger, der indgår som observationer, faktisk har varierende længde. Hauer angiver, at en spredningsparameter, der varierer med strækningens længde, er problematisk, da summen af estimerede uheld for delstrækninger ved brug af EB ikke er lig med det estimerede antal uheld for hele strækningen. Hauer argumenterer, at k nødvendigvis må falde med stigende strækningens længde. Ved sortpletudpegning vil det derfor være smart enten med en spredningsparameter, der varierer med strækningens længde, eller udpegede sorte pletter har nogenlunde samme længde som de strækninger, som uheldsmodeller er baseret på.

Miaou og Lord (2003) viser, at ϕ varierer mellem 3 og 18, altså varierer k mellem ca. 0,05 og 0,35. Ved at sammenligne uheldsmodeller med hhv. fast og varierende spredningsparameter kan de fastslå, at uheldsmodellens a-p-værdier næsten ikke påvirkes. De finder, at spredningsparameteren falder med stigende trafikmængde. De angiver, at forskellen i estimerede uheld ved brug af empirical Bayes kan være op til 35 procent mellem uheldsmodeller med hhv. fast og varierende spredningsparameter.

Mitra og Washington (2006) anfører, at det er svært at tro, at nogle variable skulle påvirke variansen uden samtidig at påvirke middelværdien. De angiver, at når uheldsmodellen er veldefineret (ingen større ubeskrevet variation i vej- eller krydsdesign), så må en konstant spredningsparameter være tilstrækkelig til at opfange uforklaret systematisk variation i uheldsforekomsten. De viser, at når uheldsmodellen er velbeskrevet, så er modellen med fast spredningsparameter faktisk bedst. I øvrigt er estimeringen af de forskellige udtryk for den varierende spredningsparameter ikke statistisk signifikant. Men i modeller, hvor kun trafikmængder indgår som uafhængige variable, er det en model med varierende spredningsparameter, der er bedst. Alle deres modeller med varierende spredningsparameter viser, at k falder med stigende trafikmængde.

ϕ konvergerer med stadig flere observationer. Det vil sige, at k går ikke mod nul men mod en bestemt værdi, når stadig flere uheld, strækninger og kryds indgår i opbygningen af modellen (Lord, 2005; Agrawal og Lord, 2006). Usikkerheden på ϕ falder således med et stigende antal observationer. Lord (2005) anbefaler, at en uheldsmodel bør være baseret på mindst 1.000 uheld, hvis spredningsparameteren skal bestemmes med et rimeligt lille konfidensinterval. Uheldsmodellens konstanter konvergerer betydeligt tidligere end ϕ . Faktisk synes a-p-værdier at være rimeligt godt bestemt i basismodeller allerede ved 300-500 uheld (Lord, 2005; Agrawal og Lord, 2005; Park og Lord, 2008).

Geedipally og Lord (2008) finder, at en varierende spredningsparameter giver en bedre uheldsmodel for 3-benede kryds i landzone end en fast parameter. En god pointe er, at modellens konfidensinterval er betydeligt mindre med en varierende spredningsparameter set i forhold til en fast. Spredningsparameteren falder, når trafikmængden stiger.

Geedipally et al. (2009) undersøger ti udtryk, hvor spredningsparameteren på forskellig facon varierer med strækningslængde og / eller trafikmængde, samt sammenligner disse med en fast spredningsparameter. Dette testes på tre sæt af data med strækninger i hhv. Texas, Californien og Washington. Modelresultater viser, at k falder, når hhv. trafikmængden og strækningslængden stiger. Det ser ud til, at trafikmængden har større indflydelse på k end strækningslængden. Det er forskellige udtryk for de tre stater, der er bedst, og derfor er det svært at blive klog på, hvordan spredningsparameteren skal udtrykkes.

Geedipally og Lord (2011) finder, at en varierende shape- eller spredningsparameter er bedre end en fast i både Conway-Maxwell-Poisson og NB modeller. De finder også, at k falder med stigende trafikmængder i kryds og med stigende strækningsslængder.

I langt de fleste tilfælde kan en uheldsmodel mest hensigtsmæssigt opstilles som en Negativ binomial (Poisson-Gamma) model. I de tilfælde, hvor der er beskeden variation i vej- og krydsdesign mellem observationer, så kan det være udmærket med en fast spredningsparameter. Ved stor variation mellem observationer er det hensigtsmæssigt at behandle spredningsparameteren som varierende. NB model med fast eller varierende k kan estimeres i f.eks. proc GENMOD, COUNTREG og NLMIXED i SAS. I alle tilfælde kan man få konfidensinterval og teststatistik både for modellens a-p værdier og andre variable samt spredningsparameteren. Man vil også kunne beregne teststatistik for hele modellen. Derved kan usikkerheden på både den forventede og den estimerede uheldstæthed beskrives.

3.2.4 Estimering, optimering og teststatistik

Oftest estimeres konstanterne og spredningsparameteren i en NB model ud fra generalized linear regression (hvor logaritmen af udtrykket er taget på begge sider af lighedstegnet) ved brug af en maximum-likelihood (ML) procedure, der gør brug af en variant af Newton-Raphson metode, hvor iterativ vægtet mindstekvadraters-metode anvendes. Andre estimeringsmetoder f.eks. maximum quasi-likelihood (MQL), method of moments (MM), weighed regression (WR) og Bayesian metoder anvendes. Det er fundet, at ML proceduren, der er standard i statistik software, resulterer i mindre fejl i estimering end MQL, MM og WR, når modellen er rimeligt simpel, og antallet af observationer og den gennemsnitlige uheldstæthed er rimeligt stort (Lord og Mahlawat, 2009). Ved mere komplekse modeller kan Bayesian estimeringsmetoder være en fordel f.eks. Markov Chain Monte Carlo metodikken og andre samplingsteknikker.

Man bør generelt ikke udarbejde uheldsmodeller med færre end 50 observationer (Lord og Mahlawat, 2009), da estimerer for modellens konstanter bliver meget usikre og upålidelige, mens spredningsparameteren bliver skævt estimeret uanset, hvilken estimeringsmetode der anvendes. Jo lavere den gennemsnitlige uheldstæthed er, desto flere observationer behøves. Valget mellem at lave en disaggregeret model (opdelt på år med årsfaktor) og en aggregeret model (årene samlet) bør afhænge af hhv. antal observationer, gennemsnitlig uheldstæthed og spredningsparameteren. Jo flere observationer, jo større gennemsnitlig uheldstæthed og jo nærmere spredningsparameteren er 1, desto oftere bør en disaggregeret model vælges.

En korrelationskoefficient, R^2 , bruges ofte ved modellering til a) at beskrive den overordnede kvalitet af en model, b) at udvælge uafhængige variable, og c) at sammenligne modellers kvalitet. Traditionelt defineres en korrelationskoefficient til at være den andel af den samlede variation, som den pågældende model eller

variabel forklarer. R^2 kan derfor være mellem 0 og 1. En korrelationskoefficient tager således ikke på behørigvis højde for den tilfældige variations andel af en samlet variation. Det gør, at R^2 bliver afhængig af omfanget af overspredning (systematisk variation, hvilket næsten altid er til stede ved modellering af uheld), og det er godt, men bliver desuden afhængig af omfanget af tilfældig variation, hvilket er skidt (Miaou et al., 1996). R^2 stiger jo lavere spredningsparameteren er. Omfanget af tilfældig variation sætter derved en grænse for, hvor stor R^2 kan blive, og denne bliver så afhængig af middelværdien af uheld (Brüde og Larsson, 1993). De almindelige korrelationskoefficienter kan ikke bruges til at vurdere kvaliteten af uheldsmodeller.

Ved opstilling af uheldsmodeller med overspredning, f.eks. NB modeller, kan andelen af systematiske variation, som modellen beskriver, være et udtryk for dens forklaringskraft. Det gøres ved at sammenstille spredningsparameteren for modellen med spredningsparameteren for de oprindelige data:

$$\text{Andel af systematisk variation forklaret} = R_k^2 = 1 - \frac{k_{\text{model}}}{k_{\text{oprindelig}}}$$

Den andel benævnes af nogen som *Elvik's indeks*. Dette indeks blev foreslået af Elvik og indgår bl.a. i en artikel af Fridstrøm et al. (1995). Miaou et al. (1996) foreslår også at anvende dette indeks. Lyon og Persaud (2002) bruger indekset til (ved brug af den inverse spredningsparameter, ϕ , $R_\phi^2 = 1 - \phi_{\text{oprindelig}}/\phi_{\text{model}}$) til at vurdere forbedringen af uheldsmodeller for kryds i byer ved at lade data om gangtrafik indtræde i modeller og finder, at variable om gangtrafik markant forbedrer modellerne. Persaud et al. (2002) bruger samme indeks til vurdering af modeller på tværs af krydstyper og hhv. med alle uheld og personskadeuheld.

Fridstrøm et al. (1995) angiver, at Elvik's indeks giver nogenlunde samme resultat som Freeman-Tukey R^2 og scaled deviance G^2 (baseret på log-likelihood ratio). Dog ser Fridstrøm et al. kun på modeller baseret på observationer med relativt store antal uheld eller personskader. Maher og Summersgill (1996) angiver, at scaled deviance ikke fungerer hensigtsmæssigt, når der er tale om lave middelværdier (få uheld eller personskader i gennemsnit). Mountain et al. (1996) bruger og argumenterer også for Elvik's indeks og Freeman-Tukey R^2 på strækninger og i kryds i både land- og byzone i Storbritannien.

Ved udvikling af modeller er standardafvigelse, konfidensinterval og signifikansniveau for variablen, der skal inkluderes eller ej, relevante at se. Ved NB modeller beregnes de ud fra maximum-likelihood estimatet, spredningsparameteren og Tukey's t-fordeling. I øvrigt vil statistik software også angive et konfidensinterval for spredningsparameteren. Skulle spredningsparameteren ikke være signifikant, så bør det overvejes om modellens fordeling fortsat skal være Poisson-gamma. For at kunne afgøre, hvorvidt det er fornuftigt at inkludere en variabel og hvordan (funktionsudtryk), er man nødt til at afveje hhv. forbedringen og simpliciteten af

modellen. Under forsætning af samme antal observationer i de forskellige modeludtryk kan afvejningen udføres ud fra en af følgende opgørelser:

- Spredningsparameteren, k = se tidligere
- Akaike's Information Criterion (AIC) = $2 \cdot p - 2 \cdot \ln(\text{ML})$
- Bayesian Information Criterion (BIC) = $p \cdot \ln(n) - 2 \cdot \ln(\text{ML})$
- Deviance Information Criterion (DIC) = ...
- Quasi-likelihood Information Criterion (QIC) = ... ,

hvor p er antal parametre, der estimeres i modellen, n er antal observationer, og ML er den maksimerede værdi af likelihood funktionen.

BIC er også kendt som Schwarz criterion. Den model med lavest værdi af k , AIC, BIC, DIC eller QIC er bedst. AIC og BIC vil som regel lede frem til, at den samme model er bedst, dvs., når AIC er lavest, så er BIC også lavest. Men særligt ved få observationer kan de give forskellige resultater. Modeller hvis AIC-værdi er 1-2 højere den laveste værdi, bør overvejes. Er værdien 4-7 højere, bør der være særdeles gode argumenter til at foretrække en model frem for den med lavest værdi, og er værdien mere end 10 højere, så bør modellen forkastes. DIC kan bruges, når modellen er bestemt ud fra en Bayesian metodik som Marko Chain Monte Carlo, eller der er benyttet andre estimeringsteknikker end maximum-likelihood. Det kan også være hensigtsmæssigt at bruge DIC, hvis man benytter hierarkisk modellering eller en varierende spredningsparameter, dog ser det umiddelbart ud til, at AIC også i den situation kan bruges, hvis modellen er maximum-likelihood baseret. Ved GEE modeller bruges QIC, da de modeller er non-likelihood baseret.

Tre andre goodness-of-fit statistikparametre er også interessante: Mean Prediction Bias (MPB), Mean Absolute Deviation (MAD) og Mean Squared Prediction Error (MSPE). Disse er f.eks. interessante, når man vil vurdere, om det er fornuftigt at sammenlægge modeller, f.eks. fra to basismodeller til en faktormodel. Med MPB vurderes, om modellen samlet set over- eller underestimerer og omfanget af dette. Med MAD opgøres den gennemsnitlige estimeringsfejl, og MSPE sætter estimeringsfejlen i anden potens og derved får store fejl langt større betydning. De tre statistikparametre er defineret således:

$$MPB = \frac{\sum_{i=1}^n (y_i - \mu_i)}{n}, \quad MAD = \frac{\sum_{i=1}^n \text{abs}(y_i - \mu_i)}{n}, \quad \text{og} \quad MSPE = \frac{\sum_{i=1}^n (y_i - \mu_i)^2}{n},$$

hvor y_i er den rapporterede uheldstæthed det i 'te sted, μ_i er den forventede uheldstæthed det i 'te sted og n er antal observationer. *abs* betyder, at negative fortegn forsvinder.

Man kan anvende grafer for kumulative residualer (CURE) til at vurdere, hvordan modellen afviger fra observationer for kontinuere variable. CURE metoden blev foreslået af Hauer og Bamfo (1997). En variabel, der er godt beskrevet i modellen, bør kurven for kumulative residualer svinge omkring 0 i dens udstrækning, f.eks.

ÅDT fra 500 til 20.000. CURE grafer viser i hvilke intervaller, at uheldstætheden er over- eller undervurderet, og viser om der findes markante knæk. Derved kan man få en indikation for relevansen af at ændre modelspecifikationen (variable og funktionsudtryk). CURE grafen skal helst være mindre end 2 gange standard afvigelsen i hele variabelens udstrækning. CURE grafer kan også anvendes i vurdering af fornuften ved at lægge modeller sammen.

Som angivet i afsnit 2.3 kan en samlet usikkerhed beregnes ved ”kvadratroden af kvadratsummen”. Ved valg af vej- og krydstype er det relevant at angive samlet usikkerhed for kombinationen af grundmodel og tilhørende sikkerhedsfaktorer (SF’er). I afsnit 2.3 er eksemplificeret en beregning baseret på standardafvigelse. Man kan også udføre beregningen for f.eks. et 95% konfidensinterval.

3.2.5 Alvorlighed af uheld og personskader

Man kan modellere personer involveret i uheld og deres skadesgrad eller uhelds-omkostninger eller uheld efter alvorlighedsgrad på rigtig mange måder. Nogen måder er mere korrekte end andre. Det store problem ved at modellere dette er, at mere end en person kommer til skade i nogen af uheldene, og at alvorligheden af personskader i uheld med mere end én personskade er afhængige. Disse to former for afhængighed – i antal og alvorlighed af skader – øger variansen i fordelingen af personskader. Noget af denne varians er tilfældig og andet systematisk.

I tabel 7 er givet en oversigt af uheld og personskader i Danmark år 2000-2009. Her kan bl.a. ses, at antallet af skader pr. uheld er højest i dødsulykker (1,74) og lavest i uheld med alvorlige skader (1,34). Af tabel 7 ses også, at skadesgraden for én person er afhængig af skadesgraden af andre personer i uheldet. Sandsynligheden for at være dræbt i et uheld, hvor en anden også er dræbt, er 2-3 gange større (14 %), end hvor en anden person er kommet let til skade (5 %).

Uheld med	Antal uheld	Antal personer			Skader pr. uheld	* Procent fordeling dræbt, alvorlig, let
		Dræbte	Alv. skader	Lette skader		
Dræbte	3.569	3.932	1.156	1.108	1,71	14 %, 46 %, 40 %
Alv. skader	30.492	914	34.050	5.881	1,34	9 %, 34 %, 57 %
Lette skader	30.777	758	5.193	38.620	1,45	5 %, 38 %, 57 %
Personskader	59.636	3.932	34.050	38.620	1,28	

Tabel 7. Antal uheld og personskader efter skadesgrad i politirapporterede uheld i Danmark år 2000-2009. * Procentuel fordeling af skader udover den ene dræbte i dødsulykker, den ene alvorlige skade i uheld med alvorlige skader, samt den ene lette skade i uheld med alvorlige skader.

Uheld med fodgængere, cyklister og knallertkørere har kun sjældent mere end én personskade. I tabel 8 er vist en tilsvarende oversigt for motorvejsstrækninger (ej kryds, flettestrækning, rampe, parallelbane og sideanlæg). Her sker kun få uheld

med bløde trafikanter og gennemsnitshastigheden er højere. Uheld på motorveje er således anderledes end uheld i almindelighed.

Uheld med	Antal uheld	Antal personer			Skader pr. uheld	* Procent fordeling dræbt, alvorlig, let
		Dræbte	Alv. skader	Lette skader		
Dræbte	243	292	98	125	2,12	18 %, 36 %, 46 %
Alv. skader	1.238	79	1.532	501	1,71	9 %, 34 %, 57 %
Lette skader	1.471	95	444	2.137	1,82	8 %, 37 %, 55 %
Personskader	2.528	292	1.532	2.137	1,57	

Tabel 8. Antal uheld og personskader efter skadesgrad i politirapporterede uheld på motorvejsstrækninger i Danmark år 2000-2009. * Procentuel fordeling af skader udover den ene dræbte i dødsulykker, den ene alvorlige skade i uheld med alvorlige skader, samt den ene lette skade i uheld med alvorlige skader.

Ragnøy et al. (2002) modellerede antal dræbte, meget alvorligt, alvorligt og let skadede på veje i Norge. De opstillede en model for hver af de fire skadesgrader. De efterviste, at f.eks. fordelingen af antal dræbte i dødsulykker og fordelingen af antal dræbte pr. km vej tilnærmelsesvis var negativ binomial, og derfor opstillede de fire NB modeller. I modellerne indgik ÅDT, antal kørespor, antal kryds pr. km, hastighedsgrænse / vejtype og vejklasse som uafhængige kontinuer eller dummy variable.

Der er flere problemer ved de norske modeller. Dels er sammenhængen mellem forekomsten af personskader og ÅDT (P-værdien) højest sandsynligt ikke ens for alle vejtyper, da sammenhængen mellem uheld og ÅDT er forskellig fra vejtype til vejtype. Det kunne man tage højde for ved at lade et synergiled indgå mellem ÅDT og hastighedsgrænse / vejtype eller opstille en model for hver hastighedsgrænse / vejtype. Dels burde en vejstrækning med f.eks. tre dødsulykker med hver én dræbt have et højere estimeret antal dræbte end en vejstrækning med én dødsulykke med tre dræbte. Det kan en hierarkisk model (uheld – person) tage højde for med én spredningsparameter for uheld og én anden spredningsparameter for antal af personskader. Dels er skadesgraderne modelleret hver for sig, og derfor er der ikke taget højde for øget varians som følge af afhængighed i skadesgrad. Det kan en multivariat model håndtere, og derved vil man i tillæg få en række spredningsparametre – en for hver skadesgrad. Problemet er bare, at der så vidt vides ikke er udviklet en hierarkisk multivariat Poisson-gamma model, hvor alle spredningsparametre kan udledes af, og det kan jo også være, at disse spredningsparametre er varierende. Derfor må man indtil videre nøjes med de ”næstbedste” modeller, som altså ikke er teoretisk korrekte.

I tabel 9 er vist antal skader pr. uheld. I tabel 9 er gennemsnittet er trunkeret (kun værdier fra 1 og opad), og derfor er spredningsparameteren, k , beregnet efter formelen: $k = (\text{Var.} - (\text{Gns.} - 1)) / (\text{Gns.} - 1)^2$. Af tabel 9 ses, at antallet af skader er voldsomt overspredte – k er høj – mest for alvorlige skader og mindst for lette skader. Man kan også erfare, at den tilfældige variation, $(\text{Gns.} - 1) / \text{Var.}$, er størst

blandt dræbte og mindst blandt lette skader. Ud fra det kan siges, at omfanget af systematisk variation er faldende jo mere alvorlig skaden er.

Dræbte		Alvorlige skader		Lette skader		Personskader	
... pr. uheld	Antal uheld	... pr. uheld	Antal uheld	... pr. uheld	Antal uheld	... pr. uheld	Antal uheld
1	3.291	1	27.725	1	25.411	1	48.675
2	219	2	2.207	2	3.793	2	7.423
3	42	3	413	3	1.049	3	2.170
4	12	4	99	4	314	4	785
5	4	5	33	5	125	5	339
6	0	6	10	6	49	6	155
7	0	7	3	7	14	7	43
8	0	8	1	8	5	8	23
9	1	9	0	9	3	9	7
10	0	10	0	10	0	10	2
11	0	11	0	11	1	11	0
12	0	12	0	12	3	12	3
13	0	13	0	13	3	13	3
14	0	14	0	14	2	14	0
15	0	15	0	15	0	15	0
16	0	16	0	16	0	16	1
17	0	17	0	17	0	17	1
18	0	18	0	18	0	18	1
19	0	19	0	19	0	19	1
20	0	20	0	20	0	20	0
21	0	21	0	21	0	21	1
22	0	22	1	22	1	22	1
23	0	23	0	23	0	23	0
24	0	24	0	24	0	24	0
25	0	25	0	25	0	25	2
Gns.	1,1017	Gns.	1,1167	Gns.	1,2550	Gns.	1,2845
Var.	0,1468	Var.	0,1873	Var.	0,4762	Var.	0,5837
k	4,4	k	5,2	k	3,4	k	3,7
Tf var	70,8 %	Tf var	62,3 %	Tf var	53,5 %	Tf var	48,7 %

Tabel 9. Antal skader pr. uheld i Danmark 2000-2009, hhv. dræbte pr. uheld med dræbte, ... personskader pr. personskadeuheld. Note: "Gns." er gennemsnit, "Var." er variansen, "k" er spredningsparameter, og "Tf var" er den tilfældige variations andel af den samlede variation.

Ragnøe et al. (2002) viser også, at andelen af systematisk variation, som model- lerne forklarer, falder jo mere alvorlig skaden er. Spredningsparameteren, k , i de norske modeller bliver derfor ret høj for dræbte og meget alvorligt skadede (2,4),

men lavere for de alvorligt (1,4) og let skadede (1,0). Hvis der rapporteret én dræbt og én let skade, og det forventede antal er 0,4 dræbte og 0,4 lette skader, så vil det estimerede antal dræbte være 0,69 og det estimerede antal lette skader være 0,57. Men da der er flere lette skader end dræbte i Norge, så vil det forventede antal dræbte typisk kun være ca. 5 % af det forventede antal lette skader, så hvis man forventer 0,02 dræbte, så vil det estimerede antal dræbte være 0,065, når der er rapporteret én dræbt. De norske modeller får derved taget godt højde for den tilfældige variation i forekomsten af personskader.

Der findes en række alternativer til den norske metode, f.eks.:

- Model baseret på uheldsomkostninger,
- Modeller opdelt på uheldsalvorlighed – hhv. dødsulykker, uheld med alvorlige skader men uden dræbte, uheld med lette skader men uden dræbte og alvorlige skader, samt materielskadeuheld,
- Modeller opdelt efter uheldsart – hhv. person- og materielskadeuheld,
- Modeller opdelt efter uheldstyper – f.eks. eneuheld, mødeuheld, osv.,
- Bivariat model med estimering af hhv. person- og materielskadeuheld, eller anden opdeling af uheld, og
- Multivariat model med estimering af hhv. dræbte, alvorlige og lette skader, eller opdeling af uheld.

I Danmark opereres med enhedspriser til beregning af uheldsomkostninger. Disse enhedspriser er (2010-priser inklusiv velfærdstab):

- 17.726.781 kr. pr. dræbt i personrelaterede uheldsomkostninger,
- 3.037.046 kr. pr. alvorlig skade i personrelaterede uheldsomkostninger,
- 457.471 kr. pr. let skade i personrelaterede uheldsomkostninger, og
- 664.543 kr. pr. rapporterede uheld i materielskaderelaterede omkostninger.

Hvis man benytter disse omkostninger på uheld i Danmark 2000-2009 findes, at uheldsomkostninger pr. rapporteret uheld varierer mellem 664.543 kr. og 160.205.572 kr. De gennemsnitlige uheldsomkostninger er 1.927.907 kr., mens variansen er 19.176.069 kr. Det giver, at en spredningsparameter på 7,5 og en tilfældig variation på kun 6,6 %. Hvis man opstillede modeller baseret på uheldsomkostninger ville man ende op med en høj spredningsparameter, da en stor del af den systematiske variation ikke vil blive forklaret, fordi trafikmængder og andre uafhængige variable ikke kan forklare prissætningen af hhv. dræbte, alvorlige og lette skader. Man vil derfor få, at den estimerede uheldsomkostning vil ligge alt for tæt på ”den rapporterede uheldsomkostning”. Modeller baseret på uheldsomkostninger (udefra kommende vægte) er derfor ikke anvendelige. Man er nødt til at modellere på det oprindelige antal – altså antal dræbte, dødsulykker, antal eneuheld, osv. – og først derefter gange enhedspriser / vægte på.

Modeller opdelt efter uheldsalvorlighed vil ikke kunne tage højde for variansen i antal personskader pr. uheld. Derved vil f.eks. eneuheldet med en dræbt få for stor

betydning i forhold til mødeuheldet med flere dræbte. Da antal personskader pr. uheld formentligt er afhængig af trafikmængder og andre uafhængige variable, og der samtidig er f.eks. flere dræbte end dødsulykker, vil modeller opdelt efter uheldsalvorlighed give ringere præcision i estimering end modeller opdelt efter skadesgrader, altså dræbte, alvorlige og lette skader.

I Danmark har Vejdirektoratet modelleret hhv. personskadeuheld og rapporterede uheld (person- og materielskadeuheld samlet). Her indgår hverken skadesgrad eller antallet af personskader, hvilket resulterer i en beskeden variation i uheldsomkostninger. Disse modeller kan ikke ”opfange” den markante forskel i skadesgrad, der er mellem f.eks. uheld i hhv. rundkørsler og prioriterede kryds, og uheld på veje med forskellig hastighedsbegrænsning. Men der kan f.eks. stilles krav om et minimumsantal uheld med personskade i forbindelse med sortpletudpegning. Modeller efter skadesgrad vil også være at foretrække frem for denne hidtidige opdeling.

Jonsson et al. (2009) opstillede uheldsmodeller for 3- og 4-benede kryds dels ud fra alle uheld dels ud fra fire uheldstyper; eneuheld, uheld mellem køretøjer i samme retning, mødeuheld, samt uheld mellem køretøjer i krydsende retninger. De fandt, at andelen af alle uheld, som hver af de fire typer udgjorde, afhang af trafikmængden. Der var ikke stor forskel på det samlede forventede antal uheld, om dette blev opgjort ud fra modellen med alle uheld eller som summen fra de fire modeller for uheldstyperne. De angiver, at der kan være for få uheld til at opstille pålidelige uheldsmodeller for uheldstyper. Sådanne modeller kan bruges til at sige, hvad man kan forvente og ved at sammenligne med rapporterede uheld, om der er visse uheldstyper, der er overrepræsenteret. Det er meget relevant, hvis tiltag kan forebygge bestemte uheldstyper.

Alvorligheden opgjort efter uheldsomkostninger er noget varierende for de forskellige uheldstyper. Men man kan hurtigt få for få uheld til at opstille en model, hvis man opdeler i for mange typer. Modeller opdelt efter uheldstype lider derfor af samme problem som modeller opdelt efter uheldsart, når målet er at beskrive alvorligheden af uheld og personskader eller uheldsomkostninger.

Geedipally og Lord (2010) udarbejdede en bivariat negativ binomial model, der samtidigt kunne modellere hhv. ene- og flerpartsuheld på strækninger i det åbne land med 4 eller kørespor men uden midterrabat. For modelopstilling henvises til artiklen. Forfatterne udarbejdede også to univariat negativ binomial modeller hhv. for ene- og flerpartsuheld, og en univariat negativ binomial model for alle uheld. Den bivariate model blev estimeret i proc NLMIXED (maximum-likelihood) i SAS. Den bivariate model har kun én spredningsparameter – altså ikke en for eneuheld og en for flerpartsuheld, selvom de univariate modeller viser, at spredningsparameteren er statistisk signifikant forskellig for de to uheldstyper. Modellen for alle uheld har et mindre konfidensinterval (større præcision) end den bivariate model. Om de to univariate modeller for hhv. ene- og flerpartsuheld er mere præcise end den bivariate model synes at kunne afhænge af antallet af observatio-

ner, men forskellen er ikke stor. Der synes derfor ikke at være en væsentlig grund til at gøre brug af bivariate modeller i hvert ikke, hvis der kun opereres med én spredningsparameter. Mountain et al. (1997) opstiller en række bivariate quasi-Poisson modeller f.eks. med hhv. fatale/alvorlige uheld og uheld med lette skader. De finder bl.a., at alvorligheden afhænger af hastighedsbegrænsning og antal mindre kryds pr. km strækning.

Flere har opstillet multivariate uheldsmodeller (se f.eks. *Ye og Lord, 2011; Ma et al., 2007; Ma og Kockelman, 2006; Miranda-Moreno et al., 2009; Song et al., 2006*). Der anvendes primært Bayesian estimeringsteknik f.eks. Markov Chain Monte Carlo eller brug af Gibbs sampler og Metropolis-Hastings algoritmer. Det kan være meget kompliceret modellering. De multivariate modeller kan opdeles i to typer hhv. nominelle (multinomial, nested og mixed logit, osv.) og ordinale (ordered logit, ordered probit, ordered mixed logit) modeller. Tre af modellerne er beskrevet af Ye og Lord (2011).

En ordinal model opererer med en rangordning f.eks. efter skadesgrad; dræbt, alvorlig skade, let skade og uskadt. Ved ordinale modeller indgår en kumulativ sandsynlighedsfunktion, så en ændring i en variabel vil gøre skaderne generelt enten mere eller mindre alvorlige. Det betyder, at forhold, der har tendens til at trække skadesgraden i retning af en skade midt på skalaen f.eks. autoværn, bliver uhensigtsmæssigt estimeret. For nominelle modeller indgår ikke en kumulativ sandsynlighedsfunktion, hvorfor f.eks. betydningen af autoværn kan blive mere hensigtsmæssigt estimeret. En nominal model er derved mere fleksibel end en ordinal model. Men denne fleksibilitet kræver flere observationer for at opnå samme præcision, faktisk 2-5 gange flere (*Ye og Lord, 2011*). De angiver, at en ordinal probit model som minimum bør være baseret på 1.000 uheld, hvor en multinomial logit model kræver 2.000 og mixed logit 5.000.

Det er meget forskelligt, hvordan modelleringen gribes an, men almindeligt er at opdele enten efter skadesgrad eller efter uheldsalvorlighed (dødsulykker, uheld med alvorlige skader men uden dræbte, osv.). De modeller, der ser ud til at komme tættest på en hierarkisk multivariat Negativ binomial model, synes at være udført af Song et al. (2006) og Miranda-Moreno et al. (2009), men modellerne er meget komplekse.

Uheldsmodeller, der skal anvendes i praksis, skal a) helst skal være lette at forstå, b) helst ikke kræve ekstremt avanceret statistisk indblik for at kunne estimere, og c) være baseret på tilstrækkeligt med data for at være så pålidelige, at praktikere finder dem troværdige. Set i det lys foreslås to tilgange til at lade antal og alvorlighed af uheld og personskader indgå.

Den første tilgang retter sig mod grundmodeller til brug for valg af vej- og krydstype, hvor kun det *forventede* antal uheld og uheldsomkostninger er interessant. Her kunne tre modeller opstilles, hvor der i hovedtræk kun opereres med trafikmængder som uafhængige variable, da øvrige forhold der i større omfang påvirker

uhelds- eller personskadeforekomsten forventes ikke at variere (det kan dog være nødvendigt at tillade en variation ad hensyn til datamængden – f.eks. hastighedsbegrænsning på motorveje – og i så fald må dette forhold indgå med en yderligere uafhængig variabel):

- Model for uheldstæthed: Basismodel type 1-2 (NB model) evt. faktormodel type 3-4
- Model for personskadetæthed: Som model for uheldstæthed, men med antal af personskader i stedet for antal uheld
- Model for fordeling af personskader efter skadesgrad: En multivariat ordinal model (dræbt, alvorlig, og let skade) enten logit eller probit estimeret på baggrund af maximum-likelihood.

Disse tre modeller kan alle estimeres i gængs statistik software såsom SAS, Genstat, SPSS, osv. Det vurderes, at de tre modeller kan estimeres rimeligt pålideligt, når de er baseret på mere end 1.000 rapporterede uheld. Baggrunden for at opdele i et sæt med tre modeller er, at multivariate modeller ikke synes at kunne håndtere overspredningen i antal personskader samtidig med fordelingen efter skadesgrad på en tilstrækkelig præcis måde. Set i lyset af de relativt få dræbte i Danmark vil man ikke kunne estimere en separat model for antal dræbte for en specificeret vej- eller krydstype, som det kræves ved grundmodeller. Her vil det være nødvendigt at bruge information om alle personskader fordelt efter skadesgrad for at få en rimelig pålidelig estimering for dræbte, og her er en ordinal model at foretrække frem for en nominel model. De tre modeller gør det muligt at beregne et forventet antal uheld, dræbte, alvorlige og lette skader, og disse kan efterfølgende omregnes til forventede uheldsomkostninger ved brug af enhedspriser.

Den anden tilgang retter sig mod sortpletmodeller. Her er både et forventet og et estimeret antal uheld, dræbte, alvorlige og lette skader samt uheldsomkostninger interessante. Det vurderes, at en kombination af almindelige uheldsmodeller og modeller for hhv. dræbte, alvorlige og lette skader er at foretrække. Særligt for modeller for dræbte vil det være nødvendigt at gå på tværs af stræknings- eller krydstyper. Fire modeller opstilles:

- Model for uheldstæthed: Modeltype 1-4 (NB model)
- Model for dræbte: Modeltype 3-4 (NB model) med synergiled
- Model for alvorlige skader: Modeltype 3-4 (NB model) evt. med synergiled
- Model for lette skader: Modeltype 3-4 (NB model) evt. med synergiled

Disse modeller kan også estimeres i gængs statistik software. Det vurderes, at disse modeller kan estimeres rimeligt pålideligt, når de er baseret på mere end hhv. 1.000 rapporterede uheld, 500-1.000 dræbte, 500-1.000 alvorlige skader og 500-1.000 lette skader. Modellerne gør det muligt at beregne et forventet og estimeret antal uheld, dræbte, alvorlige og lette skader, og de kan efterfølgende omregnes til uheldsomkostninger ved brug af enhedspriser.

Usikkerheden på uheldsomkostninger vil være en kombination af usikkerheden på alle estimater fra de enkelte modeller, som uheldsomkostninger er beregnet ud fra. Derimod må usikkerheden på enhedspriser sættes til nul, da den ikke kendes, selvom det reelt er den største usikkerhedsfaktor. Der er jo stor økonomisk forskel på, om en alvorlig skade er en brækket lillefinger eller 100 % invaliderende. Usikkerheden kan samlet beregnes ud fra "kvadratrod af kvadratsummen".

Ved sortpletudpegning kan niveau, forskel og usikkerhed på forventede og estimerede uheldsomkostninger anvendes til at angive, hvor statistisk sikker man er på, at stedet er en sort plet. Alternativt kan man på tilsvarende vis udpege sorte pletter efter uheld, og herefter rangordne de sort pletter efter forskel i forventede og estimerede uheldsomkostninger og frasortere de steder, hvor de estimerede uheldsomkostninger ikke er tilstrækkeligt meget større end de forventede.

3.2.6 Forudsætninger

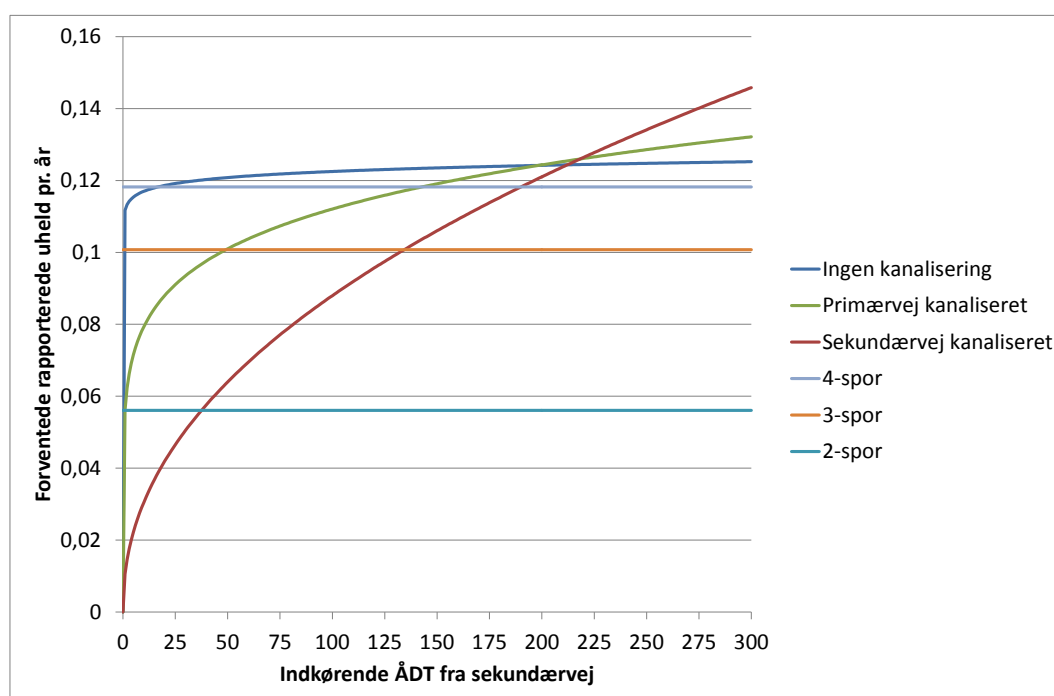
Forud for opstilling af uheldsmodeller er det nødvendigt at sætte en række forudsætninger op. Det gælder f.eks. opsplitning i strækninger og kryds, minimum og maksimum længde af strækninger, håndtering af overgangszoner på strækninger (steder hvor en strækningstype overgår til en anden), indgangsværdier (type og år) for trafik og uheld, håndtering af udeladte værdier, minimum trafik ind i kryds og på strækninger, samt udeladelse, verifikation eller vægtning af ekstreme værdier. Det er også væsentligt at forholde sig til, hvordan uheldsmodellerne bør anvendes og ikke bør anvendes samt deres gyldighedsområde.

Opsplitning i strækninger og kryds samt krav om minimum trafik ind i kryds og på strækninger bør, som angivet i afsnit 3.2.2, sammentænkes, når forudsætninger skal opstilles. I de hidtidige danske uheldsmodeller har – så vidt vides – kryds ikke en fast afgrænset udbredelse. Arealet, hvor krydsende trafik er i konflikt, er typisk, hvad der svarer til længde på 10-60 meter i primærvejens stationering, når der ses bort fra flettestrækninger. Der findes dog rundkørsler i Danmark, der er over 150 meter i diameter. Et gennemsnitligt konfliktareal i prioriterede kryds har måske en bredde i nærheden af 15 meter i det åbne land, mens bredden er større i signalregulerede kryds og rundkørsler.

Politiet rapporterer formentligt en del uheld til at være sket i kryds i et område, der er bredere end konfliktarealet, fordi eksempelvis bagendekollisioner og påkørsler af heller nær konfliktarealet henføres til krydset. Krydsuheld skulle helst registreres med to vejnumre og med krydsets stationering på primærvejen, men det er ikke altid tilfældet, da ikke alle veje har vejnumre eller er stationeret (det har dog næsten alle større veje i det åbne land), samt mangelfuld og fejl i stedfastelse forekommer. Samtidig vil en del uheld, der skyldes tilstedeværelsen af kryds, blive registreret som strækningssuheld. Det er således i praksis vanskeligt at opstille en fast afgrænsning for kryds- og strækningssuheld, der er sagligt velbegrundet og ens for alle kryds- og strækningstyper.

I USA afgrænses ofte og helt generelt i forbindelse udarbejdelse af *Highway Safety Manual* krydsudbredelsen til 250 fod (76 meter) fra krydscentret (AASHTO, 2010; Lord et al., 2008; Harwood et al., 2007). Det vil sige, at strækninger aldrig er tættere på kryds end 75 meter, og et kryds – uanset krydstype – altid har en længde af 150 meter på primærvejen. Derimod stilles der i USA sjældent minimumskrav til trafik på veje, der indgår i uheldsmodeller.

I figur 5 og 6 er Vejdirektoratets seneste uheldsmodeller for kryds- og strækningstyper i det åbne land sammenstillet. Her er det forsøgt at illustrere, hvor mange uheld der forventes at være sket på en vej med en årsdøgnstrafik (ÅDT) på 5.000. Strækninger af en længde på 200 meter, hhv. med 2, 3 og 4 kørespor, er sammenstillet med kryds med en ÅDT på 0-600 på sekundærvejen, hhv. med 3 og 4 ben, samt hhv. ingen, sekundær- og primærkanalisering.

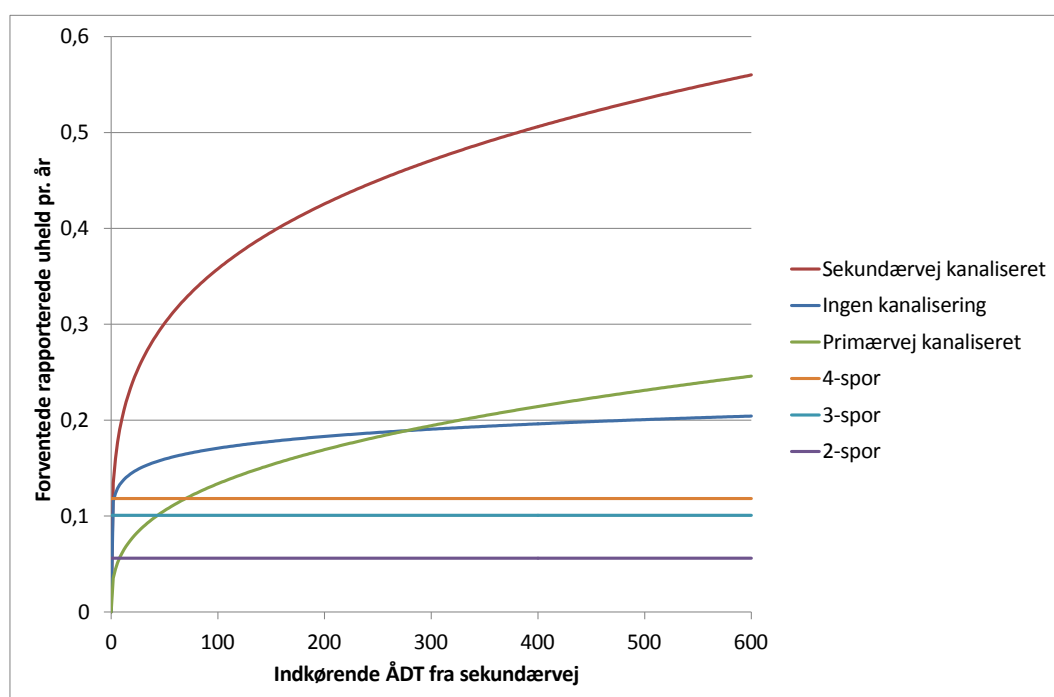


Figur 5. Forventede rapporterede uheld pr. år for 200 meter lange strækninger i åbent land med 5.000 i ÅDT hhv. med 2, 3 og 4 kørespor, og 3-benede kryds i åbent land med 5.000 indkørende ÅDT fra primærvej og varierende indkørende ÅDT fra sekundærvej hhv. med ingen, sekundær- og primærkanalisering. Baseret på Vejdirektoratet uheldsmodeller med data fra 2001-2005.

I figur 5 og 6 er 200 meter lange strækninger sammenlignet med kryds. Det er sandsynligt, at uheld stedfæstet til kryds i åbent land ikke har en udbredelse på 200 meter på primærvejen, hvorfor sammenligningen er kritisabel. Krydset har også en udbredelse på sekundærvejen, som ikke indgår i sammenligningen.

Af figurerne ses, at den forventede uheldstæthed beregnet ud fra uheldsmodeller for kryds er højere end uheldstætheden på 200 meter lange strækninger allerede,

når sekundærvejen bærer en relativ beskeden trafik. Det er selvfølgelig vigtigt at pointere, at modeller for kryds ikke er gyldige ved ÅDT under 500 på sekundærvejen og modeller for strækninger ikke er gyldige ved ÅDT over 500 på sekundærvejen. Men figurerne illustrerer den diskontinuitet, der er mellem modeller for hhv. kryds og strækninger. Når ÅDT på sekundærvejen er 499 og den forventede uheldstæthed beregnes ud fra en strækningsmodel, så forventes få uheld, mens der forventes langt flere uheld beregnet med en krydsmodel med ÅDT på 501 på sekundærvejen.



Figur 6. Forventede rapporterede uheld pr. år for 200 meter lange strækninger i åbent land med 5.000 i ÅDT hhv. med 2, 3 og 4 kørespor, og 4-benede kryds i åbent land med 5.000 indkørende ÅDT fra primærvej og varierende indkørende ÅDT fra sekundærvej hhv. med ingen, sekundær- og primærkanalisering. Baseret på Vejdirektoratets uheldsmodeller med data fra 2001-2005.

Baggrunden for diskontinuiteten mellem modellerne er formentlig flersidig. Dels er krydsmodellerne baseret på få uheld og uheldstætheden er beskeden, så selve estimeringen af modellen er problematisk – og her har brug af Poisson fordeling evt. resulteret i for høje a-værdier og for lave P-værdier. Dels er funktionsudtryk evt. ikke optimale, da mindre kryds og ind-/udkørsler ikke indgår som uafhængige variable i strækningsmodeller, og da brugen af den traditionelle basismodel for kryds kan være problematisk ved lave trafikmængder, som anført i afsnit 3.2.2. Dels er definitionen af krydsenes udbredelse uklar.

For at afhjælpe diskontinuiteten og generelt få mere pålidelige tal for forventet (og estimeret) uheldstæthed vil det være relevant med brug af NB modeller, data om mindre kryds og ind-/udkørsler, brug af andre funktionsudtryk end den traditionel-

le basismodel for kryds, samt en mere klar afgrænsning af kryds. Data om mindre kryds og ind-/udkørsler kan resultere i en stor forbedring af strækningsmodeller, både i relation til sortplet- og grundmodeller.

En klar afgrænsning af kryds kunne være, at kryds sættes til at have en fast udbredelse i meter uafhængig af geometri. Denne udbredelse skal henhøre uheld til kryds, så langt de fleste "krydsuheld" faktisk indgår i udarbejdelse af krydsmodeller. Samtidig skal afgrænsningen af kryds udbredelse muliggøre, at strækningsmodeller fortsat kan opstilles. Hvis krydsene gøres for store, så er der jo ikke meget strækning tilbage. Bibeholdes kravet om en ÅDT på minimum 500 på sekundærvejen for, at stedet er et kryds, så kunne udbredelsen være f.eks. 100 meter fra krydscentret og ud af alle krydsets ben. I tilfælde af, at et andet kryds (ÅDT på minimum 500 på sekundærvejen) er beliggende mindre end 100 meter fra krydscentret, så må de to kryds behandles som forsatte, og må ses som et krydsområde med afstanden mellem krydscentre som uafhængig variabel. Forsatte kryds må derefter opfattes som en særskilt type af kryds. Da minimumskravet til trafik på sekundærvejen for kryds er afgørende for, hvor mange steder, der skal håndteres som værende kryds, så skal minimumskravet ses i sammenhæng med afgrænsningen for udbredelsen af kryds.

En klar afgrænsning af kryds på ovenstående måde vil løse mange problematikker med overgangszoner, da veje i særlig grad ændrer tværsnit nær kryds. Når veje i øvrigt ændrer tværsnit eller andet, f.eks. ny køresporsbredde, cykelsti ophører, ny hastighedsbegrænsning, kurve påbegynder, osv., så opstår en overgangszone. I nogle tilfælde vil der være en anden uheldstæthed i og nær overgangszonen end på strækninger før og efter overgangszonen. En mulighed er at opdele overgangszonen i midten (hvor køresporsbredden er middel eller hastighedstavlen står) og lade hver halvdel af zonen tilgå strækningen før og efter. En anden mulighed er at lade overgangszonen og evt. et stykke før og efter denne udgå eller blive behandlet i en særskilt model. I langt de fleste tilfælde vil det formentligt være smartest at opdele zonen i midten og lade hver halvdel tilgå strækningen før og efter. Visse overgangszoner kan håndteres med særskilte variable, f.eks. overgangen fra en lige strækning til en kurve (tangent og klotoide), ved reduktion af antal kørespor, m.fl. Det vil sige, at klotoide strækningen kan være en særskilt observation.

I USA har man fundet, at en strækning bør minimum være 65-160 meter lang for at optræde som observation og derved få valide modeller (AASHTO, 2010). Denne minimumslængde afhænger af den gennemsnitlige uheldstæthed pr. km. Politiet i USA rapporterer betydeligt flere uheld pr. km vej end i Danmark, altså når vejen har samme trafikmængde. Det skyldes to forhold, dels er trafiksikkerheden ringere i USA, dels er mørketallet (uheld, som Politiet ikke rapporterer) større i Danmark. Derfor er det nok rimeligt at operere med en minimumslængde på ca. 100 meter for veje med mere end 2 kørespor, og på ca. 200 meter for veje med 1-2 kørespor i Danmark. Praksis med at "sammensmelte" korte strækninger til længere for at opnå minimumslængden kan accepteres, når den sammensmeltede strækning er homogen. Det er ikke relevant med maksimumslængder for strækninger, da det er

en fordel at analysere lange ensartede strækninger, både i relation til sorte pletter / grå strækninger og ved valg af strækningstype, i stedet for korte delstrækninger, der er ens. En lang strækning har en mindre usikkerhed på uheldstætheden, end når strækningen betragtes som et sæt af delstrækninger.

Vejdirektoratets seneste modeller er baseret på politirapporterede uheld fra 2001-2005 og trafiktal fra 2005. Politiet i Danmark registrerer også ”ekstrauheld”, men de indgår ikke i uheldsmodeller. Et problem med ekstrauheld er, at stedfæstelsen ofte er ringe. Det samme kan siges om sygehusregistrerede uheld. Selvom mørketallet for politirapporterede uheld er svingende fra egn til egn, så forekommer det vigtigere med en god stedfæstelse. Trafiktal (ÅDT) er beregnet ud fra tællinger af motorkøretøjer i forskellige kategorier og fastsatte personbilsækvivalenter. Det er der næppe grund til at ændre for veje i det åbne land, da fodgængere, cykler og knallerter ikke udgør nævneværdige antal. Det kan dog være, at trafik er bedre at håndtere, som et antal motorkøretøjer og en andel tung trafik, i stedet for ÅDT, da andelen af tung trafik måske har en anden betydning for trafiksikkerheden end de personbilsækvivalenter (udtryk for trafikafvikling) repræsenterer. Trafiktal bør stamme fra hele den undersøgte periode, da et enkelt år ikke nødvendigvis er repræsentativt for perioden. Som angivet i afsnit 3.2.1 kan det være en fordel at opdele uhelds- og trafikdata på år.

Det er nok ikke smart med udeladte værdier, altså uhelds- og trafiktal der mangler for et eller flere år for et kryds eller en strækning. Udeladte værdier opstår typisk som følge af nyanlæg og ombygninger, og efter disse vil der være en kortere eller længere tilvænningsperiode med usædvanlige uhelds- og trafiktal, som desværre indgår, hvis udeladte værdier accepteres. Det anbefales at undgå udeladte værdier, også fordi det muliggør uheldsmodeller byggende på paneldata.

Vægtning af ekstreme værdier er vigtig, når modeller opstilles med baggrund i en normal- eller Poisson fordeling. Vægtning af ekstreme værdier er derimod ikke relevant, når der opstilles NB modeller. I stedet er det hensigtsmæssigt med en validering af ekstreme værdier, altså en manuel kvalitetssikring (stedfæstelse af uheld, trafiktal samt vej- og krydstype) af steder med særligt høje værdier for uheld og / eller trafik.

Sortpletudpegning kan foretages på forskellige måder. I afsnit 2.1 var fokus rettet mod valg af modelspecifikation (der peges på NB modeller) og strategiske valg af, hvad den forventede uheldstæthed skal repræsentere (der peges på, at erfaring med forskellige strategier er væsentlig for at foretage valg). I sortpletudpegning er fire angivelser interessante, nemlig antallet af hhv. forventede og estimerede uheld (alternativt uheldsomkostninger), samt usikkerheden på disse to antal. Disse kan sammenfattes til i hvilken udstrækning det estimerede antal uheld ud fra en statistisk betragtning (signifikansniveau) kan siges at være højere end det forventede antal uheld. Vistisen (2002) har nøje beskrevet, hvordan signifikansniveauet kan beregnes. Det er lidt usikkert, om hendes metodik kan anvendes i tilfældet af en varierende spredningsparameter.

Angivelse af en uheldsmodels gyldighedsområde er sjældent noget, der levnes større opmærksomhed. Almindeligvis angives intervaller for indgangsværdier blandt observationer til modellen, f.eks. 2-sporede strækningers ÅDT var mellem 400 og 22.600. Men det er sjældent, at der opstilles et egentligt gyldighedsområde, da man ikke opgør modellens forklarende egenskaber i yderkanterne af (og de prediktive egenskaber uden for ...) intervallerne for indgangsværdier. Eksempelvis kunne man for 2-sporede strækninger sige, at en model med gyldighedsområde på 400-22.600 ÅDT bør have gode forklarende egenskaber, både i en situation med lidt trafik, hvor uheld oftest er alvorlige eneuheld, og i en situation nær kapacitetsgrænsen, hvor uheld oftest er materielskadeuheld med flere parter, samt i situationer mellem disse.

Der er flere måder at angive modellens forklarende egenskaber. De tre goodness-of-fit parametre, Mean Prediction Bias (MPB), Mean Absolute Deviation (MAD) og Mean Squared Prediction Error (MSPE), kan ved at opdele indgangsværdierne i passende delintervaller bruges til at opgøre, hvor god eller ringe modellen er i delintervallerne. Ved at stille passende krav kan et gyldighedsområde fastsættes.

Modellens prediktive egenskaber – altså mulighed for at forudsige uheldstæthed uden for indgangsværdiernes intervaller – kan også belyses. Det kan f.eks. være, at man ønsker at kunne sige noget om, hvad der sikkerhedsmæssigt kan ske, hvis køresporsbredden på en motorvej reduceres til under eksisterende mindstemål eller hvis en motorvej med en trafikmængde under laveste niveau åbner. En almindelig måde at teste modellers prediktive egenskaber er ved at udarbejde en ny model ud fra en del af observationerne, f.eks. halvdelen eller to tredjedele i en ende af intervallet, og så opgøre f.eks. de tre førnævnte goodness-of-fit parametre for de resterende observationer i den anden ende af intervallet. Samtidig kan man sammenligne den ny model med den oprindelige, hvor alle observationer indgår. På den måde får man et indtryk af, i hvilken udstrækning variabelnes forklarende egenskaber afhænger af indgangsværdiernes intervaller, og på denne måde kan rimeligheden i at udvide modellens gyldighedsområde til noget udenfor indgangsværdiernes intervaller belyses.

3.3 Opsummering

De nuværende danske uheldsmodeller er af ældre dato og trænger til at blive opdateret både med hensyn til modelspecifikation og data. De nuværende danske uheldsmodeller kan ikke anvendes til valg af vej- og krydstype, og kan ikke beskrive antallet og alvorligheden af personskader. Det kan være, at definitionen af kryds med fordel kan ændres.

Gennemgangen af state-of-the-art uheldsmodeller tyder på følgende:

- Det er hensigtsmæssigt at undersøge fordelingen af uheld eller personskader før en model opstilles, hvilket i hovedtræk vil sige middelværdi og varians.

- Modeller for tætheden af uheld og personskader kan i næsten alle tilfælde opstilles med baggrund i en Negativ binomial (Poisson-gamma) fordeling. Det specielle ved en Negativ binomial model (NB model) er, at modellens middelværdi (det forventede antal uheld) er lavere end variansen. Det bør dog altid undersøges, hvilken fordeling der er rimelig at antage. Det kan eksempelvis gøres for en foreløbig basismodel og en senere færdig model.
- NB modeller kan opstilles med en fast eller varierende spredningsparameter. Spredningsparameteren beregnes ud fra modellens middelværdi og varians. Med varierende spredningsparameter menes, at den afhænger af en eller flere af modellens uafhængige variable. En varierende spredningsparameter kan være en stor fordel. Det er derfor vigtigt at undersøge, hvordan spredningsparameteren bedst kan beskrives.
- Estimeringen af relativt simple NB modeller kan fint gøres med teknikken med maximum-likelihood.
- Udvalget af potentielle uafhængige variable er oftest den forhåndenværende datatilgængelighed. Det er meget sjældent, at særlige registreringer udføres for at opstille uheldsmodeller. Blandt potentielle uafhængige variable udvælges modellers uafhængige variable oftest med baggrund i statistiske test, dog er det ikke ualmindeligt, at trafikmængder indgår selvom disse ikke opfylder kravene i statistiske test. Det anbefales at benytte signifikansniveauet for variabelens sammenhæng med uheldstætheden samt AIC (ved modeller uden årsfaktorer eller HGLM modeller med årsfaktorer) eller QIC (ved GEE modeller med årsfaktorer) for modellen til at vælge, om den uafhængige variabel skal indgå i modellen. Det anbefales at optage variable trinvis i modellen efter signifikansniveau. (AIC er Akaike's Information Criterion, HGLM er Hierarchical Generalized Linear Models, QIC er Quasi-likelihood Information Criterion, og GEE er Generalized Estimating Equation)
- Brug af årsfaktorer kan være en fordel, hvis der er tilstrækkeligt med data for hvert år (eller anden periodelængde f.eks. to eller tre år). Områdefaktorer kan også være en fordel, dog er brug af områdefaktorer ikke velundersøgt.
- Variable for trafik (eksponering) kan med fordel opstilles så detaljerede som muligt f.eks. opdelt på trafikstrømme og trafikarter, i stedet for overordnede tal for årsdøgntrafik.
- Det anbefales at undersøge flere funktionsudtryk, og valget af funktionsudtryk bør være understøttet af statistiske tests. Simple udtryk, f.eks. uden synergi- og additive led, bør foretrækkes frem for komplekse udtryk, hvis forskelle i AIC eller QIC er beskedne.
- Modeller for personskader opstilles bedst for hver skadesgrad (NB modeller), hvis der er tilstrækkeligt med data til at gøre det. Alternativt må hhv. en model for personskadetæthed (NB model) samt en model for fordeling af skader efter skadesgrad (ordinal probit eller logit) opstilles.
- Samvariation mellem modellens uafhængige variable bør beskrives, og der bør tages stilling til, om det er nødvendigt at håndtere stærk samvariation.
- Det anbefales at gøre brug af empirical Bayes til opgørelse af estimerede uheld og personskader.

- Det anbefales, at modellens forklaringskraft udtrykkes ved Elvik's indeks eller Freeman-Tukey R^2 .
- Konfidensinterval og standardafvigelse for estimerede konstanter i modellen bør angives.
- Modellens gyldighedsområde bør beskrives.
- Det forekommer ikke at være velundersøgt, hvordan kryds' udbredelse bedst defineres. Det forekommer samtidigt, at den oftest benyttede basismodel for kryds ($UHT = a \times N_{pri}^{P_1} \times N_{sek}^{P_2}$) er diskutabel.

4. Datagrundlag

For at kunne opstille uheldsmodeller må data om uheld (og personskader), trafik og veje (og kryds geometri og regulering) foreligge. For at belyse tilgængelige data er der foretaget et udtræk af alle uheld i VISopslag år 2000-2009. Desuden er der oplysninger om trafik, vejtype samt signalregulering udtrukket for uheld i Vejman.dk for samme periode. I store træk synes der at foreligge oplysninger om veje og kryds udformning i Vejman.dk, når der foreligger oplysninger om trafik. Derfor bruges alene oplysningen om trafik som indikator for, at relevante oplysninger om trafik og veje foreligger i et nødvendigt omfang til, at det pågældende uheld kan indgå i opbygning af uheldsmodeller. På grund af Kommunalreformen ved indgangen til 2007, så er det lidt usikkert, hvor opdateret oplysningerne er i Vejman.dk for kommuneveje.

Det følgende gennemgår mængden af tilgængelige data om uheld og personskader i tiårsperioden 2000-2009, der kan anvendes til opbygning af uheldsmodeller for strækninger og kryds i det åbne land.

Data, 2000-2009	Administrativ vejtype samt zone				
	Motorveje	Motortrafikveje	Øvrige veje, landzone	Øvrige veje, byzone	Uoplyst
Dræbte	326	108	2.296	1.202	0
Alvorlige skader	1.914	403	12.908	18.825	0
Lette skader	2.683	469	15.401	20.067	0
Personskader	4.923	980	30.605	40.094	0
Personskadeuheld	3.244	609	21.074	34.709	0
Materielskadeuheld	5.406	786	19.751	65.427	4
Ekstrauheld	9.040	576	13.539	34.760	304
Uheld	17.690	1.971	54.364	134.896	308
Uheld m. trafikdata	17.427	1.927	34.399	27.414	1

Tabel 10. Fordeling af politiregistrerede uheld og personskader 2000-2009 på vejtype og zone.

Tabel 10 er opbygget ved at knytte parameteren adm_vejtype_vis fra Vejman.dk til uheld fra VISopslag. Øvrige veje er herefter opdelt på by- og landzone ud fra uheldsoplysninger. Nogle uheld på motorveje og motortrafikveje indgår ikke i Vejman.dk og har ikke oplysninger om trafik, samt er derfor manuelt tildelt en vejtype. 308 primært ekstrauheld kan hverken opgøres efter vejtype eller zone.

Af tabel 10 ses, at antallet af dræbte på motorveje samt antallet af hhv. dræbte, alvorlige og lette skader på motortrafikveje er under 500. Der vil ad den grund ikke kunne opstilles pålidelige særskilte modeller for de skader på de vejtyper, altså eksempelvis en model for dræbte på motorveje. En model for dræbte, der

skal kunne opgøre estimerede dræbte, må derfor bero på flere vejtyper i det åbne land. Det skal samtidig pointeres, at tabel 10 inkluderer uheld både på strækninger og i kryds.

I det følgende ses alene på uheld og personskader på motorveje, motortrafikveje og øvrige veje i landzone. Uheld på øvrige veje i byzone samt uheld med uoplyst vejtype frasorteres således. Ekstrauheld er også frasorteret, da en betydelig andel af disse uheld har stedfæstelsesoplysninger af ringe kvalitet.

Motorveje	Strækning	Flettestrækning	Rampe	Sideanlæg	Kryds
Dræbte	287	10	10	0	19
Alvorlige skader	1.513	68	87	9	237
Lette skader	2.120	110	93	1	359
Personskader	3.920	188	190	10	615
Personskadeuheld	2.502	132	150	10	450
Materielskadeuheld	3.783	321	240	29	1.033
Uheld	6.285	453	390	39	1.483
Uheld m. trafikdata	6.281	444	374	19	1.355

Tabel 11. Politirapporterede uheld og personskader 2000-2009 på motorveje opdelt efter "vejudformning". Uheld på flettestrækning er indtruffet ved fletninger mellem fra- og tilkørselsramper og motorvejsstrækning/parallelbane. Uheld på rampe er indtruffet på rampe eller parallelbane. Uheld i kryds er indtruffet enten i almindelige rampekryds, for enden af motorvej eller ved vigepligtsregulerede kryds på sideanlæg.

Opdelingen af uheld efter "vejudformning" er udført ved at gøre brug af oplysninger fra uheldsoplysninger om vejnumre, vejdele, vejudformning samt vejtyper (angivet for elementer/partier i uheld).

Af tabel 11 ses, at der er rapporteret godt 6.000 uheld og knap 4.000 personskader på strækninger. Det vil være muligt at opstille pålidelige uheldsmodeller for disse strækninger, og modeller for antal personskader og fordelingen af skader efter skadesgrad. Man vil også kunne opstilles særskilte modeller for hhv. lette skader og dræbte/alvorlige skader. Det vil højst sandsynligt være muligt at opstille en grundmodel for 4-sporede motorveje med 110/130 km/t hastighedsbegrænsning, med nødspor, uden vejbelysning, og med køresporbredder på ca. 3,5 meter. En sådan samlet grundmodel vil inkludere beskrivelse af uheldstæthed, personskadetæthed og fordeling af skader efter skadesgrad.

Antallet af uheld og personskader på både flettestrækninger og ramper er for få til at opstille modeller for disse. Ved at "udvide" flettestrækninger til at omfatte en længere del af motorvejen vil der muligvis være tilstrækkeligt med uheld til at opstille en model. En anden mulighed kunne være at betragte flettestrækning og rampe som én observation. Ved en tilkørselsrampe kunne man eksempelvis lade uheld sket på rampen (ej rampekryds), flettestrækning og 150 meter af motorvejen

(en vejside) opstrøms ift. start af flettestrækning samt 400 meter nedstrøms ift. slutning af flettestrækning indgå. Data om f.eks. længde af rampe og flettestrækning, rampedesign samt trafik på rampe og motorvejsstrækning kan indgå i modeller. Det vil være hensigtsmæssigt at foretage analyser af uheld ved og nær flettestrækninger for at fastsætte længder op- og nedstrøms. Med en sådan ”udvidet” rampe-flette-opstilling vil det måske være muligt at opstille pålidelige uheldsmodeller samt modeller for antal personskader og fordeling af skader efter skadegrad. Det er tvivlsomt, om der kan opstilles en grundmodel for rampe-flettesteder, da rampedesignet varierer betydeligt.

Uheldsmodeller for sideanlæg (rastepladser og serviceanlæg) vil ikke kunne opstilles. Det vil være nødvendigt at se mere bredt på anlæg, hvor der parkeres og / eller tankes brændstof. Da sideanlæg er meget forskellige i udformning vil det være en vanskelig opgave at relatere udformning og trafik for disse anlæg til en uheldstæthed.

Som det ses af tabel 11 indgår mange uheld i kryds i vejtypen motorveje. De 1.483 uheld er fordelt med 451 uheld i T-kryds, 718 uheld i F-kryds, 51 uheld i rundkørsler, og 263 uheld i andre kryds (øvrige kryds, stikryds samt krydstype uoplyst). Af de 1.483 uheld skete mindst 741 uheld i signalregulerede kryds. Af de 1.483 uheld skete 286 uheld i byzone. Der er ikke tilstrækkeligt med uheld til at opstille pålidelige uheldsmodeller særskilt for motorvejsrampekryds. Det vil derfor være nødvendigt at lægge motorvejsrampekryds sammen med andre kryds i det åbne land for at opstille uheldsmodeller, og evt. med en variabel angive, at krydset er et rampekryds.

Motortrafikveje	Strækning	Flettestrækning	Rampe	Sideanlæg	Sti	Kryds
Dræbte	84	1	0	0	0	16
Alvorlige skader	273	11	1	0	1	107
Lette skader	250	15	4	0	0	174
Personskader	607	27	5	0	1	297
Personskadeuheld	357	23	5	0	1	223
Materielskadeuheld	400	29	14	2	1	340
Uheld	757	52	19	2	2	563
Uheld m. trafikdata	749	52	19	2	1	536

Tabel 12. Politirapporterede uheld og personskader 2000-2009 på motortrafikveje opdelt efter ”vejudformning”. Uheld på flettestrækning er indtruffet ved fletninger mellem fra- og tilkørselsramper og motortrafikvejsstrækning/parallelbane. Uheld på rampe er indtruffet på rampe eller parallelbane. Uheld i kryds er indtruffet enten i almindelige rampekryds, for enden af motortrafikvej eller ved vigepligtsregulerede kryds på sideanlæg.

Af tabel 12 ses, at antallet af uheld og personskader på motortrafikveje er få. Det vil dog være muligt at opstille en rimelig pålidelig model for strækninger, men en sådan uheldsmodel skal baseres på mange år – måske 10 eller flere år. Da antallet

af uheld i det enkelte år er for få til at opstille hensigtsmæssige årsfaktorer vil det være nødvendigt at sammenlægge 2-3 år og knytte faktorer til sådanne perioder. Alternativt må motortrafikveje lægges sammen med øvrige veje i det åbne land, og så opstille en model, hvor eksempelvis vejtype eller hastighedsbegrænsning indgår som variabel.

Der vil ikke kunne opstilles modeller for flettestrækninger, ramper, sideanlæg og stier i relation alene til motortrafikveje. Et spørgsmål er, om det giver mening at lægge flettestrækninger og ramper på motortrafikveje sammen med flettestrækninger og ramper på motorveje. Forskellen i trafikafviklingen er formentlig for stor disse steder til, at det er fornuftigt.

Som det ses af tabel 12 indgår tillige mange uheld i kryds i vejtypen motortrafikveje. De 563 uheld er fordelt med 112 uheld i T-kryds, 356 uheld i F-kryds, 68 uheld i rundkørsler, og 27 uheld i andre kryds (øvrige kryds og krydstype uoplyst). Af de 563 uheld skete mindst 242 uheld i signalregulerede kryds. Af de 563 uheld skete 116 uheld i byzone. Det vil derfor også være nødvendigt at lægge kryds ad motortrafikveje sammen med andre kryds i det åbne land for at opstille uheldsmodeller for kryds.

Øvrige veje	Strækning	Flettestrækning	Rampe	Sideanlæg	Sti	Kryds
Dræbte	1.684	0	0	7	10	595
Alvorlige skader	8.378	15	10	45	142	4.318
Lette skader	9.072	20	3	37	111	6.158
Personskader	19.134	35	13	89	263	11.071
Personskadeuheld	13.351	25	11	78	218	7.391
Materielskadeuheld	10.459	49	30	214	82	8.917
Uheld	23.810	74	41	292	300	16.308
Uheld m. trafikdata	13.994	64	24	58	207	12.006

Tabel 13. Politirapporterede uheld og personskader 2000-2009 på øvrige veje i det åbne land opdelt efter "vejudformning".

Af tabel 13 kan erfares, at antallet af uheld og personskader på flettestrækninger og ramper også er beskedent på øvrige veje i det åbne land. Sammenlægges disse med uheld på flettestrækninger og ramper på motortrafikveje, vil der fortsat være for få uheld til at opstille modeller. Det anses for mest hensigtsmæssigt alene at opstille modeller for flettestrækninger og ramper ved motorveje, og ikke at lade uheld sådanne steder ved motortrafikveje og øvrige veje i det åbne land indgå.

Af tabel 13 ses, at antallet af uheld på sideanlæg også er beskedent. Selv ved sammenlægning af sideanlæg ved motorveje, motortrafikveje og øvrige veje i det åbne land vil det ikke være muligt at opstille modeller for sideanlæg. Ligeså kan det siges om uheld på stier.

Antallet af uheld på strækninger er meget stort på øvrige veje i det åbne land, og for over halvdelen af disse kan trafikdata tilknyttes. Antallet af dræbte, alvorlige og lette skader er også så stort, at separate modeller for disse tre skadesgrader kan opstilles. Uheldsoplysninger om vejtype under elementer/parter kunne tyde på, at når motortrafikveje og øvrige veje i det åbne land lægges sammen, så er der tilstrækkeligt med uheld til at opstille tre modeller for hhv. 2 sporede veje, 2 sporede veje med midterrabat samt 3-4-sporede veje med/uden midterrabat. Det er dog usikkert, hvor stor en andel af uheldene på veje med midterrabat og på veje med 3-4 kørspor er indtruffet nær kryds, hvor ”midterrabatten” blot er en lang helle.

Der vil uden tvivl kunne opstilles uheldsmodeller for 2-sporede veje uden midterrabat, idet mindst $\frac{3}{4}$ af alle strækningsskader med trafikdata på motortrafikveje og øvrige veje i det åbne land er sket på en sådan strækningstype. På den strækningstype er der indtruffet ca. 16.000 uheld, 1.000 dræbte, 4.500 alvorlige og 5.000 lette skader, hvor der samtidig forefindes trafikdata. Der er således gode muligheder for at opstille sortpletmodeller, grundmodeller og separate modeller for de tre skadesgrader for denne strækningstype.

For krydsene på øvrige veje i det åbne land, motorveje og motortrafikveje er det ikke opgjort i hvilket omfang, der er trafikdata for alle vejben. I tabel 14 nedenfor er det opgjort, hvor mange uheld og personskader der er rapporteret i de kryds, hvor der findes trafikdata for mindst 1 vejben.

Kryds	Alle kryds				Signalregulerede kryds			
	Uheld	Dræbt	Alvorlig	Let	Uheld	Dræbt	Alvorlig	Let
T-kryds	7.249	213	1.660	2.397	668	8	83	138
F-kryds	6.136	209	1.396	2.091	2.408	39	338	560
Rundkørsler	1.484	6	135	183	0	0	0	0
Øvrige kryds	420	7	75	126	125	1	15	29
Jernbane	42	5	1	3	4	0	0	0
Stikryds	22	0	7	7	4	0	3	1
Uoplyst	1.480	39	223	369	275	3	14	37

Tabel 14. Politirapporterede uheld og personskader 2000-2009 i kryds ad motorveje, motortrafikveje og øvrige veje i det åbne land, hvor der forefindes trafikdata for mindst ét vejben.

Mange af uheldene i øvrige kryds og kryds med uoplyst krydstype er formentligt sket i T-kryds, F-kryds og rundkørsler.

Tabel 14 kunne pege i retning af, at det vil være muligt at opstille pålidelige uheldsmodeller for T-kryds og F-kryds, både vigepligts- og signalregulerede, samt for rundkørsler i det åbne land. Det vil ikke være muligt at opstille modeller for jernbaneoverskæringer og stikryds. For vigepligtsregulerede T-kryds og F-kryds samt signalregulerede F-kryds vil det formentligt også være muligt at opstille modeller for antallet af personskader samt fordeling af skader efter skadesgrad. For

vigepligtsregulerede T-kryds og F-kryds vil tillige formentligt være muligt at opstille separate modeller for dræbte/alvorlige skader og lette skader.

Hvis kryds "udvides" til også at omfatte strækninger op til f.eks. 100-200 meter væk fra krydsstationeringen, så vil antallet af uheld og personskader, der kan indgå i modeller for kryds stige. Der er ikke foretaget analyse af stigningens størrelse. Det er her vigtigt at påpege, at en "udvidelse" af kryds vil reducere antallet af uheld og personskader, der kan indgå i modeller for strækninger og ramper.

Som angivet førhen er det uklart, hvor opdateret vej- og trafikdata er i Vejman.dk særligt for kommuneveje. I årene fremover er det sandsynligt, at flere data vil blive tilgængelige til brug for opstilling af modeller for uheld og personskader. Data kan tilgå fra f.eks. projekter om landsdækkende trafikmodel, digitalt vejnet, samt at flere kommuner indgår i Vejman.dk. Man kan også forestille sig en bedre udnyttelse af højopløselige luftfoto og avanceret fotogrammetri, så data om flere veje – også med relativ beskedne trafikmængder – bliver tilgængelige.

5. Forslag til nye uheldsmodeller

I nærværende kapitel findes et forslag til fremtidige danske uheldsmodeller for veje i det åbne land. Kapitlet indledes med at pege på nogle få analyser, der kan være ønskværdige at gennemføre for at få forudsætningerne for fremtidige uheldsmodeller på plads. Der opstilles en initierende vej- og krydstypeinddeling. Slutteligt angives en procedure for frembringelse af endelige modeller.

5.1 Ønskværdige indledende analyser

Som påpeget tidligere kan det evt. være hensigtsmæssigt at afgrænse kryds, ramper og flettestrækninger anderledes end ved hidtidige uheldsmodeller i Danmark. Det handler om udbredelsen af kryds og flettestrækninger, minimumstrafikmængder på sekundærveje i kryds, samt om ramper kan modelleres sammen med flettestrækninger.

En anden afgrænsning af udbredelsen af kryds kan være hensigtsmæssig. Dels er nogle uheld nær kryds stedefæstet til krydset, mens andre uheld ikke er. Dels har strækninger nær kryds ofte et andet tværprofil end strækninger længere væk fra kryds. En anden afgrænsning af krydsudbredelsen vil have til formål, at en større andel af uheldene, der skyldes krydssets tilstedeværelse, faktisk indgår i modeller for kryds, samt at strækninger, der indgår i uheldsmodeller for strækninger, har mere homogene tværprofiler, og korte strækninger i større omfang undgås.

Det foreslås at gennemføre analyser for hhv. T-kryds, F-kryds og rundkørsler. Analyserne kan have til formål at beskrive tre forhold, dels krydsenes fysiske (svingbaner, heller, belysning, osv.) udbredelse fra krydscenter, vige- eller stoplinjer, dels forekomsten af nærliggende andre kryds (f.eks. forsatte T-kryds), og dels uheld på strækninger op mod kryds. Analyserne kan udføres med strækningsintervaller af 25-50 meter. Ud fra beskrivelserne vil man bedre kunne fastlægge en god afgrænsning for udbredelsen af kryds, og illustrere om det er muligt for at opstille særlige modeller for forsatte kryds.

En anden analyse omhandler minimumstrafikmængder på sekundærveje. Her er det relevant at opføre i hvilket omfang, der findes oplysninger om trafikmængde på sekundærveje med årsdøgnstrafik under 500. For disse kryds bør der opgøres uheldstætheder. Denne analyse kan evt. udføres sammen med analysen om krydsudbredelse.

En anden afgrænsning af flettestrækninger på motorveje kan også være hensigtsmæssig. En analyse kunne her have til formål at beskrive uheld på strækninger på motorvejen hhv. før, ved og efter flettestrækninger, opdelt hhv. på fra- og tilkørsler. Også her kunne analyserne udføres med strækningsintervaller af 25-50 meter.

Analysens formål er at beskrive, hvornår tilstedeværelsen af en fra- eller tilkørsel får udpræget indflydelse på uheldsforekomsten. Med beskrivelsen vil man bedre kunne fastlægge en afgrænsning for flettestrækninger.

Analysen af flettestrækninger kan udføres sammen med en analyse af ramper for at belyse, om det er muligt at opstille uheldsmodeller, der betragter rampe og flettestrækning som én samlet observation.

5.2 Initierende vej- og krydstypeinddeling

Omfanget af tilgængelige data blev belyst i kapitel 4. Ud fra dette kapitel er det muligt at opstille et bud på en initierende vej- og krydstypeinddeling. Følgende inddeling foreslås:

Sortpletmodeller til beregning af forventede og estimerede uheldstætheder:

Strækninger:	Kryds:
Motorveje	Kryds med 3 ben/grene
Flettestrækninger-ramper ved motorveje	Kryds med 4 ben/grene
2-sporede veje med/uden midterrabat	Prioriterede T-kryds
3-4-sporede veje med/uden midterrabat	Forsatte prioriterede T-kryds
	Prioriterede F-kryds
	Rundkørsler
	Signalregulerede T-kryds
	Signalregulerede F-kryds

For strækninger kan potentielle uafhængige variable være f.eks. trafikmængde (ÅDT), tværprofil (kørespør, midterrabat, kantbane, osv.), tracé (kurveradius, kurvelængde, osv.), hastighedsbegrænsning, vejbelysning og rampelængde. For 2-sporede veje kan det evt. vise sig, at flere forskellige modeller kan udarbejdes.

For kryds kan potentielle uafhængige variable være f.eks. trafikmængder (ÅDT, hhv. primær og sekundær), tværprofil (køre-, sving- og cirkulationsspor, osv.), hastighedsbegrænsning, signalstyring, længde mellem forsatte kryds, osv. For modeller for kryds med hhv. 3 og 4 ben/grene (hvor både prioriterede og signalregulerede kryds samt rundkørsler indgår) kan krydstype indgå. For disse modeller kan spredningsparameteren vise sig at blive så stor, at modellerne må siges at være u hensigtsmæssige. Det kan evt. vise sig at være en fordel at sammenlægge rundkørsler i åbent land med rundkørsler i byområder, og så operere med en uafhængig variabel for zone. En sådan løsning kan evt. også vise sig som en fordel med signalregulerede T-kryds.

Sortpletmodeller til beregning af forventede og estimerede personskadetætheder:

Strækninger:	Kryds:
Dræbte – alle strækninger	Dræbte – alle kryds
Alvorlige skader – motorveje	Alvorlige skader – alle kryds
Lette skader – motorveje	Lette skader – alle kryds
Alvorlige skader – andre vejstrækninger	
Lette skader – andre vejstrækninger	

Det må komme an på en prøve, hvorvidt det er hensigtsmæssigt at lade ramper og flettestrækninger indgå i kategorien motorveje. Med andre vejstrækninger menes øvrige veje i det åbne land og motortrafikveje. Modeller for personskadetætheder kan bygge på de samme uafhængige variable, som modeller for uheldstætheder, dog kan det være relevant at ”sammenfatte” nogle variable til egentlige vej- eller krydstyper.

Der foreslås tre sæt grundmodeller for strækninger i det åbne land. Med ét sæt grundmodeller menes én model for uheldstæthed, én model for personskadetæthed samt én model for fordeling af skader efter skadesgrad. De tre sæt er:

- *Motorvejsstrækninger* med 4 kørespor og nødspor, 110/130 km/t hastighedsbegrænsning, ingen vejbelysning, ingen afstandsmærker, kørespor ca. 3,5 meter brede, ...
- *Højklasset 2-sporede veje* uden midterrabat, med midtlinje og kantpæle, med minimum x,x meter brede kantbaner, uden ind- og udkørsler, horisontalkurver skal minimum have radius af xxxx meter, 80/90 km/t hastighedsbegrænsning, ingen vejbelysning, ingen overhalingsforbud, kørespor 3,2-4,0 meter brede, ...
- *Lavklasset 2-sporede veje* uden midterrabat, uden kantbaner, mellem 0 og x ind- og udkørsler pr. km, ingen vejbelysning, horisontalkurver med minimum radius af xxx meter, 60/70/80 km/t hastighedsbegrænsning, ingen vejbelysning, kørespor 2,5-3,5 meter brede, ingen overhalingsforbud, ...

Disse grundmodeller har trafikmængde (ÅDT) som uafhængig variabel. Det kan dog være relevant også at benytte f.eks. hastighedsbegrænsning, bredde af kørespor og kantbaner, forekomst af midtlinje samt antal ind-/udkørsler som andre uafhængige variable.

Der er højst sandsynligt ikke tilstrækkeligt med data til at opstille et sæt grundmodeller for 6-sporede motorveje. De 6-sporede motorveje adskiller fra 4-sporede ved ofte at have korte afstande mellem til- og frakørsler, vejbelysning, samt korte afstande mellem kørespor og autoværn. Derfor vil det næppe give mening at lade 6-sporede motorveje indgå i den foreslåede grundmodel for motorvej og så operere med en variabel for antal kørespor – men det er relevant at prøve. Et problem er nemlig, at der kun findes beskeden viden om de sikkerhedsmæssige konsekvenser af at udvide motorveje fra 4 til 6 spor.

Der er højest sandsynligt heller ikke tilstrækkeligt med data til at opstille sæt af grundmodeller for 2-sporede veje med midterrabat, 3-, 4- og 6-sporede veje (ej motorveje) med og uden midterrabat, herunder 2+1 veje. Også her er der mangel på viden. Det er særligt 2-sporede veje med midterrabat samt 6-sporede veje i det åbne land, der er stærkt begrænset viden om.

En særlig problemstilling er, at udvidelser af motorveje fra 4 til 6 spor og af andre veje fra 2 til 4 eller 6 spor medfører en øget kapacitet, og derved et øget interval i mulig årsdøgntrafik. Gyldighedsområdet for foreslåede grundmodeller vil derfor ikke kunne dække dette øgede interval for ÅDT. Dette er vigtigt at pointere, hvis det alligevel findes nødvendigt at lade sikkerhedsfaktorer (SF'er) for flere køre-spør indgå.

Det er uvist, om det er muligt at opstille grundmodeller for kryds. Problemet er, at antallet af uheld og personskader for velspecificerede varianter af kryds er ukendt. Det kan dog være relevant at analysere følgende krydstyper, hvoriblandt mere velspecificere varianter kan identificeres:

- Prioriteret T-kryds uden kanalisering
- Prioriteret T-kryds med kanalisering på både primær- og sekundærvej
- Prioriteret F-kryds med kanalisering på både primær- og sekundærvej
- Signalreguleret T-kryds
- Signalreguleret F-kryds
- Rundkørsler med 1 cirkulationsspor og 4 vejgrene

Egentlige krydstællinger af trafikstrømme i 4-benede kryds anses for nødvendige til opstilling af relevante grundmodeller for F-kryds og rundkørsler. I T-kryds kan trafikstrømme beregnes ud fra tællinger på de tre strækninger før krydset. For alle fem nævnte krydstyper kan det være relevant at sammenlægge kryds i åbent land med kryds i byområder, og benytte en variabel for zone.

Der er ikke tilstrækkeligt med data til at opstille sæt af grundmodeller for forsatte T-kryds, rundkørsler med 2 cirkulationsspor samt 2-planskryds. Effekten af at omdanne F-kryds til forsatte T-kryds er dog rimeligt velbelyst, og derfor vil en modellen for F-kryds kunne anvendes sammen med SF'er til at beskrive uheldstæthed mv. i forsatte T-kryds. Viden om 2-planskryds er også rimelig stor, så her vil de foreslåede modeller sammen med SF'er også kunne anvendes til at beskrive uheldstætheden i 2-planskryds.

Der findes betydeligt mindre viden om forskelle i sikkerhedsniveau på rundkørsler med hhv. 1 og 2 cirkulationsspor, samt rundkørsler med forskellige antal vejgrene. Det kan evt. være hensigtsmæssigt at medtage rundkørsler med 2 cirkulationsspor og rundkørsler med et andet antal end 4 vejgrene i en grundmodel, og så benytte variable for antal cirkulationsspor og vejgrene.

5.3 Procedure for modeludarbejdelse

Processen med at frembringe uheldsmodeller involverer en række trin, der hver især er væsentlige. Disse trin er i det følgende beskrevet.

1. Formål

Det første trin er at beskrive formålet med modelarbejdet. Her er det væsentligt at besvare flere spørgsmål:

- *Hvad skal modellerne anvendes til?* Som angivet tidligere er det væsentligt i hovedtræk at opstille modeller til hhv. sortpletudpegning samt valg af vej- og krydstype på forskellig vis.
- *Hvad skal modellerne repræsentere?* Det er vigtigt at angive, hvilke vej- og krydstyper der ønskes at indgå i modelarbejdet.
- *Hvordan skal den trafikale risiko beskrives?* Forekomsten af uheld og personskader kan opgøres på en række måder. Det bør angives, hvordan tætheden af uheld (f.eks. i alt, efter uheldsart, efter uheldstype) og personskader (f.eks. i alt, efter skadesgrad) ønskes opgjort.

2. Analyseenheder

En klar definition af analyseenheder (observationer) er vigtig før datafangsten iværksættes. Denne definition kan opdeles i fire dele:

- *Fysisk definition:* Her defineres så præcist som muligt, hvad der menes med f.eks. motorvejsstrækning eller 4-benet signalreguleret kryds. Heri indgår bl.a. udbredelsen af kryds og derved også udbredelsen af strækning.
- *Trafikal definition:* Her defineres, om modellerne skal betragte bestemte intervaller af trafikmængde og i givet fald hvilke.
- *Geografisk definition:* Her defineres, hvor analyseenheder må stamme fra. Det kan være, at visse områder vælges f.eks. Danmark, landzone, eller bestemte vejbestyrer, f.eks. statsveje.
- *Tidsmæssig definition:* Her defineres, hvilken tidsperiode som modellerne skal baseres på. Det er her vigtigt at angive den samlede tidsperiode f.eks. 2004-2009, samt om tidsperioden skal kunne opdeles, f.eks. i enkelte år.

3. Variable og datafangst

Forud for datafangsten bør en liste af eksisterende forhold, der kan formodes at kunne have en større betydning for trafiksikkerheden (tæthed af uheld, dræbte, osv.) for de definerede analyseenheder, udarbejdes. Denne liste håndteres som potentielle variable. Efter datafangst bør der for variable på denne liste, hvor det

ikke har været muligt at indsamle data, udføres en diskussion af disse udeladte variable.

Data indsamles fra de forskellige kilder. Data kan eksempelvis organiseres på følgende vis:

- *Sted-ID*: Hver analyseenhed tildeles en identifikation
- *Uheld og personskader*: Antal uheld, personskader, dræbte, alvorlige og lette skader for analyseenheden i det aktuelle tidsrum evt. opdelt på uheldstyper
- *Tidsperiode*: Periode eller enkelt år
- *Geografisk område*: Region, politidistrikt, kommune eller lign.
- *Trafik*: Årsdøgntrafik, evt. fordelt efter trafikstrømme, trafikarter, mv.
- *Fysisk udformning*: Beskrivelse af analyseenhedens regulering, geometri, afmærkning, omgivelser og evt. signalstyring. Her indgår tillige længde, udbredelse og antal/tæthed af ind-/udkørsler og mindre kryds.

Ofte er det nødvendigt at organisere data i databaser, fordi data skal kunne sættes op som paneldata med forskellig periodelængde. Det er vigtigt at kunne opgøre den gennemsnitlige tæthed af uheld og personskader, og andelen af steder med en tæthed på nul, for de enkelte år samt for perioder med flere år.

En databasestruktur er også særlig vigtig for strækninger, da en strækning fra et kryds til et andet i mange tilfælde kan opdeles i delstrækninger især med hensyn til variation i fysisk udformning, men evt. også med hensyn til trafik og geografisk område.

4. Indledende analyser

Før en model udvikles, er det vigtigt at foretage valideringer, at få en god forståelse af data, samt at foretage en række valg af betydning for modelspecifikationen.

For steder med ekstreme værdier (meget høje antal uheld eller personskader og meget høje eller lave mængder af trafik) bør der udføres valideringer. Det indbefatter en manuel gennemgang af stedfæstelsen af uheld og en vurdering af trafikmængder i forhold til nærliggende trafiktællinger.

Herefter udføres en analyse, der skal lede frem til et valg af organisering af data med hensyn til tidsperiode. Fire tal opgøres for hver potentiel organisering af data (f.eks. opdelt efter perioder på 1 år, 2 år, 3 år, ... 10 år): Gennemsnitlig tæthed af uheld og personskader, andelen af steder med en tæthed på nul, varians i tæthed samt spredningsparameteren. Særlig ud fra andelen af steder med en tæthed på nul og spredningsparameteren kan det vurderes, hvordan en model bedst kan organiseres med hensyn til tidsperioder, dvs. relevansen af årsfaktorer.

Man kan efterfølgende udføre en analyse, der skal lede frem til en organisering af data med hensyn til geografisk område. Det kan være væsentligt at opgøre f.eks. antallet af steder, gennemsnitlig uheldstæthed og -frekvens, samt variationen i den fysiske udformning i de enkelte geografiske områder. Ud fra sådanne opgørelser kan man tage stilling til, hvorvidt variationen mellem områderne gør det relevant med områdefaktorer, og om omfanget uheld og personskader gør brug af områdefaktorer muligt.

I den sidste indledende analyse opstilles en korrelationsmatrix for de potentielle uafhængige variable. Det er særligt stærk korrelation mellem variable for trafik og variable for fysisk udformning, som man skal være opmærksom på. Almindeligvis vil variable for trafik blive valgt i modeludviklingen, mens evt. flere variable for den fysiske udformning fravælges (ej signifikante), fordi de er stærkt korrelerede med variable for trafik. Stærk korrelation blandt variable for fysisk udformning forekommer ofte, så her vil man typisk være nødsaget til at fravælge variable i den efterfølgende modeludvikling. Denne indledende analyse er vigtig, da den kan medvirke til at undgå potentielle skævheder og misfortolkninger som følge af samvarierende eller endogene uafhængige variable.

5. Modeludvikling

Forud for udvikling af modellen må man vælge, hvordan modellens variable skal udvælges ud fra statistiske tests. Det kan være, at visse variable, især for trafik, vælges at skulle indgå uanset udkommet af statistiske tests. Det er også vigtigt at angive et minimum for, hvor god hele modellen skal være for at blive accepteret.

Herefter vælges estimeringsteknikken. Det anbefales at gøre brug af maximum-likelihood metodikken, da de mere sofistikerede Bayesian metodikker næppe er besværet værd.

Næste step vil være at vælge fordeling for residualer. Umiddelbart vil oplysninger om spredningsparameteren fra forrige trin, samt illustration af fordeling kunne lede til et fornuftigt valg. I praktisk taget alle tilfælde vil en form for Poisson fordeling – typisk Negativ binomial fordeling – være fornuftig. Statistik software, der kan håndtere den valgte estimeringsteknik og fordeling for residualer, kan så vælges.

Sidste step er opstilling af modellen. Det kan vælges at udføre regressionsanalyse for variable for trafik før andre variable betragtes. Derved opstilles basismodeller. I denne proces vil det være hensigtsmæssigt at afprøve en række funktionsudtryk, hvor alene uafhængige variable for trafik indgår. I denne proces er det også smart at afprøve, om og hvordan spredningsparameteren varierer.

Efterfølgende testes om andre variable er statistisk signifikante. En mulighed er først at teste eventuelle års- og områdefaktorer. Dernæst testes variable for den fysiske udformning.

I tilfælde af, at års- og områdefaktorer ikke indgår, anbefales det at bruge kombinationen af statistisk signifikans og AIC som vurderingsmetode for, om en variabel bør indtræde i modellen. Ellers anvendes kombinationen statistisk signifikans og QIC/DIC som metode.

Hvis variable for den fysiske udformning indgår i den indtil videre udarbejdede model, kan det anbefales at undersøge forskellige funktionsudtryk for disse. Det kan indebære synergi- og additive led. Hvis der er signifikante synergiled mellem variable for fysisk udformning og trafik, bør det overvejes, om og hvordan modellen kan opdeles i to eller flere modeller. Efter dette arbejde kan analyser for spredningsparameteren atter udføres, hvis de forrige analyser for denne viste, at den tydeligvis varierede.

Ved opstilling af multivariate ordinale modeller for skadesgrader gennemgås i vid udstrækning samme processer, dog er visse valg anderledes. Det er væsentligt at påpege, at en model for tæthed af personskader kan indeholde andre uafhængige variable end en model for tæthed af uheld. Ligeledes kan en multivariat model for skadesgrader indeholde andre variable end en model for tæthed af personskader.

Det anbefales, at den færdige model analyseres ud fra et CURE plot, samt at der foretages en undersøgelse og fastsættelse af gyldighedsområde. Disse analyser og undersøgelser kan evt. være en iterativ proces ved opstilling af funktionsudtryk.

6. Modelbeskrivelse

Den frembragte model beskrives ved funktionsudtryk, tests statistik, usikkerhed, indgangsværdier, gyldighedsområde, samt udeladte og ej signifikante variable.

6. Analyser af designelementer

Gennem tiden er der udført mange med-uden studier, der søger at opgøre sikkerhedseffekter af diverse vejtekniske geometriske eller signaltekniske elementer, både i Danmark og udlandet. I Danmark har man for veje i åbent land bl.a. søgt at belyse effekter af fysisk midteradskillelse, cykelstier, kantlinjer, køresporsbredde, kantbanebredde, yderrabatbredde, trug versus grøfter, hældning på vejskråninger, faste genstande, kryds' reguleringsform, hastighedsgrænser ved lyskryds, design af rundkørsler, osv.

Med-uden studier opdeles normalt i to typer af studier. I tværsnitstudier sammenlignes typisk uheldstætheder for to eller flere forskellige vej- eller krydsdesign ved samme trafikmængder eller der sammenlignes uheldsfrekvenser (uheld pr. køretøjsenhed). Med tværsnit- i ordet "tværsnitstudie" henvises ikke til vejens tværsnit men til, at en population, f.eks. alle veje i det åbne land, opdeles på tværs af en variabel i nogle snit (kategorier). Den anden type af studie kaldes for case-control studier. Case-control studier er anderledes ved, at udvalgte "cases" indgår i studiet, og en case altid har en "control". Case og control skal være "matchende", dvs. at der skal kontrolleres for skævheder (confounding factors). Visse forhold kan dog godt variere, nemlig den eller de risikofaktorer, som undersøges. En variant af tværsnitstudier er bi- eller multivariate modeller. I disse modeller betragtes flere resultater (f.eks. skader efter skadesgrad og uheld efter uheldstyper) samtidigt. Derved kan de forskellige kategorier, f.eks. prioriterede kryds versus rundkørsler, sammenlignes for f.eks. alvorlighed af personskader.

Op til 1990'erne, hvor trafiksikkerhedsforskningen var en relativ ny disciplin, troede mange, at man kunne opgøre en sikkerhedsfaktor (SF) af f.eks. kantbaner ved blot at sammenligne uheldsfrekvenser for veje med kantbaner med veje uden kantbaner – altså tværsnitstudie. Det kan man ikke, bl.a. fordi uheldsfrekvensen afhænger af eksponeringen. De sidste 20-25 år har man derfor næsten fuldstændigt gået bort fra disse tværsnitstudier, da resultaterne ofte er meget upålidelige. Denne slags tværsnitstudier anvendes dog stadig til at beskrive forskelle i trafiksikkerhed f.eks. mellem køn, vejtyper osv. – men altså kun til at beskrive forskelle ikke til at forklare dem. I stedet er tværsnitstudier, hvor man sammenligner uheldstæthed og hvor eksponeringen indgår blandt flere uafhængige variable i en faktormodel, taget mere i anvendelse. Disse faktormodeller adskiller stadig mere fra de oprindelige tværsnitstudier ved, at stadig flere kontroller med efterfølgende datatilpasning udføres for at sikre et pålideligt resultat.

Det er velkendt, at trafiksikkerhed er et resultat af utrolig mange forhold. Matchningen i et case-control studie kan tage højde for nogle af disse forhold men ikke alle. Selvom case-control studier er langt bedre til at opgøre en SF end tværsnitstudiers sammenligning af uheldsfrekvenser, så kan resultater være upålidelige. Det er forhold som udvælgelse af cases, matchning af controls, ej korrekte modi-

fiktorer og problemer med endogene uafhængige variable, der er afgørende for, om case-control studier er pålidelige.

I de senere år bi- og multivariate modeller blevet afprøvet. Disse modeller adskiller sig fra almindelige modeller ved, at man får to eller flere resultater af samme model ved en beregning. Man har forsøgt sig med modeller, der giver en tæthed af personer opdelt på skadesgrad, af uheld opdelt på alvorlighed, samt af uheld opdelt på situationer. Fordelen ved sådanne modeller er, at der findes korrelationer mellem f.eks. skadesgrader, og korrelationerne indgår i modellen. Bi- og multivariate modeller opstilles ofte på tilsvarende vis som faktormodeller, altså et tværsnitsstudie med kontroller og tilpasninger.

Både faktor-, bi- og multivariate uheldsmodeller kan lide af de samme problemer som andre med-uden studier, nemlig udeladte og endogene variable. Men uheldsmodellerne kan også lide af et andet problem – samvariation mellem uafhængige variable. Hvis der ikke tages hånd om dette, så kan samvariationen medføre, at der frembringes ulogiske og utroværdige resultater.

Metodeudviklingen af før-efter uheldsevalueringer har ført til meget pålidelige studier de to seneste årtier. Resultater af pålidelige før-efter uheldsevalueringer har flere gange været meget forskellige fra resultater af især tværsnitsstudier. Et eksempel er, at implementering af en lavere hastighedsgrænse som regel udviser en forbedret trafiksikkerhed i før-efter uheldsevalueringer, mens tværsnitsstudier kan give vidt forskellige resultater. Metodesvage tværsnitsstudier viser som regel, at jo lavere hastighedsgrænsen er, desto højere er uheldsfrekvensen.

Før-efter uheldsevalueringer kan ikke nødvendigvis beskrive samtlige relevante sikkerhedseffekter. Det kan skyldes, at visse designelementer sjældent bygges om, og før-efter uheldsevalueringer derfor ikke kan omfatte tilstrækkeligt med uheldsdata til at beskrive en pålidelig sikkerhedseffekt, f.eks. af kørebanebredde. Det kan også skyldes, at visse designelementer oftest bygges om samtidig med andre, f.eks. vejbelysning.

Der er således fortsat behov for andre undersøgelsesmetoder end før-efter uheldsevalueringer. Af de førnævnte alternativer ses på:

- Tværsnitsstudier med uheldsmodeller
- Case-control studier
- Bi- og multivariate modeller

Disse metoder kan eventuelt kombineres med andre undersøgelsesmetoder f.eks. tidsserie-analyse. De oprindelige tværsnitsstudier, hvor sikkerhedseffekter blev vurderet ud fra uheldsfrekvenser, bør man se helt bort fra.

I de følgende afsnit ses på, hvilke designelementer på strækninger og i kryds i det åbne land, der kunne være relevante at undersøge. Der ses også på om og hvordan

sikkerhedseffekter for disse designelementer kan undersøges, dvs. hvordan en undersøgelse i hovedtræk kan udføres, samt om der er tilstrækkeligt datagrundlag for at udføre en undersøgelse.

6.1 Relevante analyser

En sikkerhedseffekt (SF) angiver en forskel i sikkerhedsniveau mellem to design af strækninger eller kryds. Det vil være relevant at få SF'er for ombygninger (fra et design til et andet), som vejbestyrelser udfører eller ønsker at udføre, samt for alternative design, der indgår i valg af vej- og krydstype.

Ved ombygninger og alternativer, der påvirker kapacitet, hastighedsniveau og / eller oplevet serviceniveau væsentligt, kan der forekomme migration af uheld og induceret vækst i et sådant omfang, at det vil være nødvendigt at se på et større influensvejnet og tage højde for ændringer i trafikmængder – for at få et samlet billede af konsekvenser for trafiksikkerheden. Et eksempel kunne være at udvide en motorvej fra 4 til 6 kørespor, og her kunne flere forskellige ændringer have konsekvenser for trafiksikkerheden:

- Effekt af design: Trafiksikkerheden på motorvejen påvirkes af ændret design – alt andet lige, dvs. trafikmængder, trafiksammensætning, osv. er uændret.
- Effekt af flyttet biltrafik: Udvidelsen vil sandsynligvis flytte trafik fra andre veje til motorvejen, hvilket kan påvirke sikkerheden samlet set i området.
- Effekt af længere bilture: Udvidelsen vil sandsynligvis øge rejsehastigheden i bil i området, hvilket vil øge rejselængden. Det vil øge mængden af biltrafik.
- Effekt af overflytning fra andre transportformer: Den øgede rejsehastighed vil sandsynligvis gøre bilen som transportform mere attraktiv, og flytte personer og gods fra andre transportformer til bilen. Det vil øge mængden af biltrafik og mindske mængden af anden trafik i området.
- Effekt af tilstrømning af jobs og beboere: Samlet vil området med motorvejen sandsynligvis få en tilstrømning af jobs og beboere, som følge af en hurtigere infrastruktur. Det vil øge mængden af trafik i området med motorveje, men samtidig mindske trafikken i andre områder.

Når vi i nærværende kapitel ønsker at kunne angive SF'er, er det alene følgen på trafiksikkerheden af det ændrede design, altså en alt andet lige betragtning. Det forudsættes, at øvrige effekter, som eksempelvis de fire sidstnævnte ovenfor, kan håndteres på anden vis.

På de følgende sider er forsøgt at vurdere, hvilke designelementer i kryds og på strækninger det kunne være relevant og muligt at undersøge ved at udføre med-uden studier, såsom case-control studier og faktoruheldsmodeller. I den sammenhæng er det vurderet, om den eksisterende viden allerede er tilstrækkelig, eller ny viden er påkrævet. Det er vurderet, hvilken metodik vil være mest givtig med hensyn til pålidelige SF'er. Før-efter uheldsevalueringer er i mange tilfælde mere

pålidelige end med-uden studier. Og endelig er det vurderet, om en undersøgelse i Danmark kan føre frem til resultater, der kan forbedre grundlaget til at kunne opstille pålidelige SF'er.

Følgende relevante analyser vurderes mulige:

- 1) Bredde og forekomst af kørespor, kantbane, nødspor og cykelstier samt evt. midter-, skille- og yderrabatter
Der er udført masser af undersøgelser om bredder og forekomst af tværsnitselementerne i det åbne land, dog er de fleste af undersøgelseerne over 20 år gamle. Der er lille til stor grad af usikkerhed om SF'er af bredder og forekomst af tværsnitselementer. Få gode nye undersøgelser tyder på, at SF'er af bredder af tværsnitselementer afhænger af trafikmængde og hastighedsniveau.
- 2) Forekomst og intensitet af vejbelysning
Der er stor grad af usikkerhed om SF'er af vejbelysning på motorveje og veje i øvrigt i det åbne land. Det er uklart, om der er taget højde for ulykker med vejbelysningsmaster i dagslys og endogene variable.
- 3) 2-planskryds i forhold til andre krydstyper
Der er stor grad af usikkerhed om SF'er af 2-planskryds i forhold til andre krydstyper. Der findes ikke sammenligninger af 2-planskryds og rundkørsler. Et problem i sådanne undersøgelser er, at sammenligninger ikke nødvendigvis er særligt udspecificerede, hvorved der er mindre eller større variation i design både blandt 2-planskryds og den eller de alternative krydstyper.
- 4) Design af 2-planskryds – ramper og flettestrækninger
Der er nogen grad af usikkerhed om SF'er af design af flettestrækninger og ramper. Næsten alle undersøgelser på området er amerikanske.
- 5) Længde af og andre designforhold for svingbaner
Der er stor grad af usikkerhed om SF'er af længde af og andre designforhold for svingbaner. Der findes kun få undersøgelser på området.
- 6) Design af rundkørsler
Der er stor grad af usikkerhed om SF'er for de mange forskellige varianter af og elementer i rundkørsler. Her tænkes på forhold såsom midterøens diameter, bredde og antal af cirkulationsspor, forekomst og typer af heller, shunts, osv.
- 7) Omgivelserne
Der er stor grad af usikkerhed om SF'er af forskellige omgivelser. Der tænkes her på elementer som støjværn, beplantning, kløft-skråning-åbent, bygværker, ind- og udkørsler, reklamer, rastepådsler, serviceanlæg, mv.

8) Type af vejbelægning

Der er nogen grad af usikkerhed om SF'er af typer af belægning på motorveje og veje i øvrigt samt i og ved kryds og rundkørsler i det åbne land. Typen af belægning indbefatter forhold som f.eks. farve, genskin, støjforhold, friktion, afvanding og vandsprøjt, osv.

9) Vedligeholdelsestilstand af vejbelægning, afmærkning og tavler

Der er kun en lille grad af usikkerhed om SF'er af genasfaltering. Til gengæld forekommer der er at være stor grad af usikkerhed om SF'er af ujævnheder, huller i vejen, sporkøring, osv., samt lapninger og udbedringer heraf. Desuden er viden om SF'er om vedligehold af afmærkning og tavler beskeden. Viden om vedligeholdelsestilstandens betydning for trafiksikkerheden er i øvrigt ikke specificeret for de enkelte vej- og krydstyper og hastighedsklasser.

Det kan i dag vanskeligere at frembringe pålidelige resultater, da antallet af uheld og personskader – heldigvis – er lavere end for 10, 20, 30 og 40 år siden. Et andet problem er tilgængeligheden til vej- og trafikdata. Til gengæld er det i dag langt nemmere at verificere eller detaljere vejdata og data for omgivelserne, da luft- og vejfotos er tilgængelige via internettet.

6.2 Potentielle analysemetoder

Som nævnt ses på tværsnitstudier, case-control studier samt bi- og multivariate modeller evt. i samspil med andre metodikker. I det følgende ses på metoder, og der indgår konkrete eksempler med tidligere undersøgelser.

6.2.1 Tværsnitstudier

Faktoruheldsmodellen, også kaldet kovariatmodellen, blev også omtalt i kapitel 3, men nedenfor er lige vist formeludtrykkene igen:

$$\text{Strækning: } UHT = a \times N^P \times e^{\sum_{i=1}^n b_i x_i} \quad (3)$$

$$\text{Kryds: } UHT = a \times N_{\text{pri}}^{P_1} \times N_{\text{sek}}^{P_2} \times e^{\sum_{i=1}^n b_i x_i} \quad (4)$$

I faktormodeller indgår konstanterne b_i og variablene x_i , der beskriver stedets udformning (f.eks. kanalisering, belysning, osv.).

Estimering af sikkerhedsfaktorerne (SF'er) opgøres ud fra estimater for variablene x_i . Hvis f.eks. vejbelysning indgår som dummy-variabel ($b=0$ uden vejbelysning, og $b=x$, x med belysning), så kan uheldstætheden opgøres for de to situationer hhv. med og uden vejbelysning. Derved fås en SF.

Baggrunden for, at én faktormodel er bedre at benytte til opgørelse af SF'er end to basismodeller (alternativt to faktormodeller), er, at usikkerheden på en SF bliver

mindre. Bruges atter eksemplet med vejbelysning, så vil to basismodeller, hhv. en for veje med vejbelysning og en for veje uden vejbelysning, have to a-værdier og to P-værdier. Derved bliver SF'en afhængig af hele fire værdier, hvor den ved én faktormodel kun var afhængig af én værdi. Fordelen ved at bruge én faktormodel er derfor indlysende, men der er også ulemper.

Det må anses for uhensigtsmæssigt at operere med vej- eller krydstyper med vidt forskellige P-værdier i samme model, da man i så fald får beskrevet betydningen af de andre uafhængige variable (f.eks. vejbelysning) forkert. Hvis undersøgelsen går ud på at vurdere forskellen på to vejtyper eller krydstyper med forskellige P-værdier, f.eks. hhv. signalregulerede kryds og rundkørsler, bliver det altså særligt problematisk. Her må det frarådes at benytte en faktormodel. Det må også i vid udstrækning frarådes at sammenligne to basis- eller faktormodeller, da det er vanskeligt at få styr på samvarierende og endogene uafhængige variable på tværs af de to modeller. Et case-control studie og / eller en multivariat model vil her være at foretrække.

En anden ulempe ved faktormodeller kan være, at man overser eller misfortolker samvariation mellem uafhængige variable. Hvis f.eks. veje med vejbelysning i gennemsnit har væsentlig mere trafik end veje uden vejbelysning kan konstanter i modellen blive usikre. En mulighed er at udelade observationer (ved stratificeret tilfældig frasortering), så de to gennemsnit nærmer sig hinanden. En anden mulighed er at modellere forekomsten af vejbelysning i forhold til både trafikmængde og uheldsforekomst samtidig med modellering uheldsforekomst – og derved mere betragte vejbelysning som en endogen variabel.

En tredje ulempe (der også eksisterer, når to modeller sammenlignes) er endogene uafhængige variable. Man kunne forestille sig, at f.eks. vejbelysning er etableret, hvor der førhen er sket mange uheld i mørke eller hvor en stor andel af trafikken kører i mørke. Her vil brug af instrumentvariable såsom forekomst af sprituheld samt fordeling af trafik i dagslys og mørke evt. kunne nedbringe vejbelysningens endogene egenskaber.

For at få et udbytterigt tværnsnitsstudie med en velunderbygget dokumentation er det således vigtigt at have styr på P-værdier, samvarierende og endogene variable. Det er således essentielt at kende forskelle i trafikforhold, vejtekniske forhold og uheldsforekomst i detaljer.

Der er gennem årene gennemført mange tværnsnitsstudier. I det følgende er fem studier refereret.

Hadi et al. (1995) konkluderede med baggrund i 36 NB faktoruheldsmodeller for 2-, 4- og 6-sporede motorveje og almindelige veje, hhv. med og uden midterrabat, at visse veje var sikrede end andre, samt bredere kørespor, nødspor, kantbaner og midterrabatter resulterer i bedre sikkerhed. Selvom modellerne er udarbejdet på fornuftig vis, analyseres ikke mulige problemer med udeladte, samvarierende og

endogene uafhængige variable. Variable er antaget enten at være kontinuere uden maksima eller minima (bredder, trafikmængde, antal kryds, hastighedsgrænser, længder) eller dummies (forekomst af midterrabat, kantsten). Der er ingen synergivvariable. De påpeger, at disse antagelser muligvis ikke er korrekte, men undersøger det ikke. Usikkerheder for estimer angives ikke. I realiteten må man konkludere, at deres undersøgelser ikke indeholder velunderbygget dokumentation for årsag-virkningssammenhænge, altså kausale sammenhænge. Deres konklusioner virker dog forventelige, og det er besnærende og de accepteres derfor af mange.

Council og Stewart (1999) bruger nogenlunde samme set-up som Hadi et al. til at sammenligne sikkerheden på 2-sporede veje uden midterrabat med 4-sporede veje med eller uden midterrabat. Alle veje er i åbent land. De konkluderer, at ombygning af 2-sporede veje til 4-sporede veje med midterrabat vil reducere antallet af uheld med 40-60 procent, mens en ombygning til 4-sporede veje uden midterrabat kun vil reducere antallet med 0-20 procent. Forfatterne angiver usikkerheder for deres estimer. Forfatterne er her langt mere usikre på, om deres konklusioner holder. Således angiver de, at indflydelsen fra ind-/udkørsler og mindre kryds ikke er undersøgt (udeladt variabel), og der savnes flere observationer af 2-sporede veje med særligt høje trafikmængder, og der mangler estimer for alvorligheden af personskader. Endelig angiver de, at før-efter uheldsevalueringer er nødvendige for at verificere deres konklusioner.

Stamatiadis et al. (2009) prøver at finde sikkerhedseffekten af bredder for hhv. kantbaner og midterrabat. De opstiller NB modeller for 4-sporede veje i åbent land hhv. med og uden midterrabat. De finder eksempelvis, at forekomsten af autoværn i midterrabat øger antallet af uheld med 118 procent, mens forekomsten af venstresvingsbaner på veje med midterrabat reducerer antallet af eneuheld med 28 procent, men øger antallet af flerpartsuheld med 57 procent. De finder samtidig, at sikkerheden forbedres, jo bredere (kontinuære variable) kantbaner og midterrabat er (forbedringen er ca. 5-6 procent for yderligere 30 cm kantbane og ca. 8 procent for yderligere 3 meter midterrabat), men at smalle kantbaner og midterrabatter er mere usikre end ingen kantbaner og midterrabatter. De angiver, at udeladte variable såsom hastighed, venstresvingende trafik, mv., hvor de ikke har data, kan være baggrunde for de fundne resultater. Resultaterne for bredder af kantbaner og midterrabat er næppe logiske. Igen må man konkludere, at studiet ikke indeholder velunderbygget dokumentation for årsag-virkningssammenhænge, samtidig er deres resultater ikke helt forventelige, og måske derfor angiver, at megen yderligere forskning er nødvendig – de tror ikke på egne resultater. Man kan tilføje, at man burde ikke håndtere bredder som kontinuære variable, og man burde ikke have sammenlagt data fra tre forskellige delstater med vidt forskellig sikkerhedsniveau uden brug af områdefaktorer.

Donnell et al. (2011) undersøger sikkerhedseffekten af belysning ved kryds i to delstater i USA. De diskuterer de potentielle problemer med endogene variable, og påpeger vigtigheden af brug af instrumentvariable. De angiver, at et problem er at håndtere flere endogene variable samtidigt. De vælger at opstille NB uhelds-

modeller for hver delstat og hhv. for trafik i dagslys og trafik i mørke. I modellerne indgår forskellige krydstyper og -varianter, der beskrives med en lang række uafhængige variable. De finder, at belysning ved kryds er forbundet med en reduktion i uheld på 23 procent i Californien og 12 procent i Minnesota. Et problem ved modellerne er ringe forklaringskraft. Studiet her glimrer ved at opdele uheld og trafik i dagslys og mørke, og derved søger at håndtere endogenitetsproblemet med belysning. Derimod sammenblandes vidt forskellige krydstyper og -varianter, hvilket er uheldigt, da det introducerer megen samvariation og endogenitet. Dette kan være forklaringen på forskellen i effekten af belysning mellem de to delstater, men må samtidig øge usikkerheden på effekten betydeligt.

Kim og Washington (2006) studerede endogene variable og brugte venstresvingsspor som eksempel. De modellerede tværkollisioner i 155 4-benede kryds. Kryds var defineret med udbredelse af 250 fod (76 meter) ad vejene fra krydscentret. Først estimerede de en almindelig NB model, som så ud på følgende måde:

$$UHT_{\text{tværkollision}} = 0,01494 \times N_{\text{pri}}^{0,3227} \times N_{\text{sek}}^{0,2868} \times e^{\sum 0,6897 \times V + 0,1142 \times I - 0,5925 \times L},$$

hvor $UHT_{\text{tværkollision}}$ er antallet af tværkollisioner, N_{pri} er ÅDT på primærvejen, N_{sek} er ÅDT på sekundærvejen, V er forekomsten af venstresvingsspor (1 hvis venstresvingsspor ellers 0), I er antallet af ind-/udkørsler ad vejene mindre end 76 meter fra krydscentret, og L er forekomsten af vejbelystning (1 hvis vejbelystning ellers 0).

Ud fra ovenstående model beregnes ($e^{0,6897}=1,99$), at forekomsten af venstresvingsspor øger antallet af tværkollisioner med 99 procent. Men Kim og Washington opstillede herefter to nye modeller, hvor hhv. tætheden af tværkollisioner og forekomsten af venstresvingsspor estimeres samtidigt ved hjælp af en speciel teknik. De tager derved højde for, at forekomsten af venstresvingsspor afhænger af: N_{pri} , $UHT_{\text{tværkollision}}$, antallet af ind-/udkørsler samt om krydset er signalreguleret eller ej. Den nye model for tætheden af tværkollisioner ser således ud:

$$UHT_{\text{tværkollision}} = 0,1334 \times N_{\text{pri}}^{0,1289} \times N_{\text{sek}}^{0,2756} \times e^{\sum -0,1653 \times V + 0,1583 \times I - 0,6520 \times L}$$

Ud fra denne model beregnes ($e^{-0,1653}=0,85$), at forekomsten af venstresvingsspor reducerer antallet af tværkollisioner med 15 procent. Et meget anderledes resultat. Ved at sammenligne de to modeller kan også erfares, at estimaterne for de andre variable også ændrer sig. Donnell et al. (2011) antyder, at flere uafhængige variable kan være endogene. I Kim og Washingtons modeller kan man forestille sig, at også variable for antallet af ind-/udkørsler og forekomsten af vejbelystning er endogene. Eksempelvis er næsten alle signalregulerede kryds i USA vejbelyste, mens ikke-signalregulerede kryds ofte ikke er vejbelyste.

Gennemgangen af de fire studier illustrerer vanskeligheder ved at producere velunderbygget dokumentation for sikkerhedseffekter på basis af tværsnitsstudier.

6.2.2 Case-control studier

I case-control studier opereres med kategoriske værdier. Det undersøges om en konsekvens (uheld eller personskade) afhænger af en risikofaktor. Hvis man f.eks. vil undersøge betydningen af forekomsten af vejbelysning for uheldsforekomsten, så kan man bruge følgende opstilling:

		Vejbelysning	
		Ja	Nej
Uheld	Ja	A	C
	Nej	B	D

Effekten af vejbelysning vil efterfølgende beregnes til odds ratio værdien (relativ risiko): $(A/B)/(C/D)$. I eksemplet ovenfor er "cases" vejstrækninger med mindst ét uheld, mens "controls" er vejstrækninger uden uheld. Risikofaktoren (tiltaget) der undersøges er vejbelysning. Forestil dig, at der er 1.000 vejstrækninger, heraf 500 med uheld og 500 uden, så kunne de fordele sig således:

		Vejbelysning	
		Ja	Nej
Uheld	Ja	100	400
	Nej	120	380

Odds ratio værdien for vejbelysning vil så være: $(100/120)/(400/380) = 0,79$. Det kan oversættes til, at tilstedeværelsen af vejbelysning reducerer uheldsforekomsten med 21 procent.

Et "trick" i case-control studier kan være at udtrække uheld fra nogle strækninger eller kryds i mange år f.eks. 10 år for derefter at opdele efter år, så hver strækning eller kryds repræsenteres 10 gange. Man lader dog selvfølgelig altid de matchede par af cases og controls være fra samme år. Og undersøger om strækninger eller kryds har ændret sig i det undersøgte tidsrum.

Det er simpelt at beregne, hvor mange strækninger eller kryds, der er nødvendige i en undersøgelse, hvis en sikkerhedseffekt skal angives med et konfidensinterval af en bestemt maksimal størrelse. Metoden herfor gives dog ikke.

I case-control studier er det tilladt at operere med flere cases end controls og vice versa, men det komplicerer de efterfølgende statistiske beregninger.

Case-control studier har oftest været anvendt til at opgøre effekten af tiltag, der primært påvirker alvorligheden af personskader, f.eks. sikkerhedsseler, cykelhjelme, osv. Men der er også studier, der har set på effekten af tiltag, der primært påvirker forekomsten af uheld.

Et eksempel er Gross et al. (2009) studie af bredde af kørespor og kantbaner på tosporede veje i åbent land uden midterrabat. I dette studie indgår strækninger fra to delstater, Pennsylvania og Washington, men hver stat behandles hver for sig. Kørebane er mellem 24 og 36 fod bred (7,32 – 10,97 m), og bredden af hhv. kørespor og kantbane inddeles i kategorier. Efter nøje overvejelser og vurderinger udvælger man to matchningsvariable; ÅDT og strækningslængde. Det vil sige, at case og control altid har nogenlunde samme trafik og længde. Der opstilles så en model, der tager højde for skævheder (covariates): Hastighedsbegrænsning, bredde af yderrabat, tracé, og regionale forskelle indenfor en delstat. Med tricket med opdeling i år er det muligt at sammenligne 87.666 ud af 362.311 strækninger. Resultatet er relativ risiko (odds ratio værdi) mellem en baseline opstilling (3,66 m kørespor og 1,83 m kantbane) og den aktuelle bredde af kørespor og kantbane. Undersøgelsens resultater ser rimeligt fornuftige ud set i forhold til andre undersøgelser om kørespors- og kantbanebredde. Man kan dog stadig have indvendinger mod den konkrete metodik. Eksempelvis kunne bredde af yderrabat være indgået som matchningsvariabel i stedet for covariat pga. en betydelig samvariation med bredde af kørespor og kantbane. Man kunne også lade yderligere forhold indgå som matchningsvariabel eller covariat, f.eks. sprituheld, trafiksammenstilling, antal ind- og udkørsler samt type af længde- og kantafmærkning.

Et af de måske bedste case-control studier indenfor trafiksikkerhed er nok Evans (1986) studie af bl.a. sikkerhedssellers effekt. Han benytter en såkaldt "Double pair comparison method", hvor case og control er personer fra samme køretøj, oftest personbiler. Eksempelvis sammenstilles dræbte førere med sele med passagerer uden sele og omvendt dræbte passagerer uden sele med førere med sele. For at tage højde for siddepositionen i køretøjet opereres med et yderligere sæt af data, hvor dræbte førere uden sele sammenstilles med passagerer med sele og omvendt. Data stratificeres også efter alder og køn. Med den metode mener Evans, at der i praksis tages højde for nærmest alle skævheder. Det synes også at være tilfældet, men man skal huske på, at sikkerhedseffekten af selen er baseret på situationer med mere end én person i bilen. Da personer uden sele har det med at skade andre personer i bilen (Nukene og Daniel, 2011), når de kastes rundt i kabinen, så er sikkerhedseffekten af selen i tilfælde med kun én person i bilen faktisk bedre.

Det er sværere at udføre et studie af vej- eller signaltekniske elementer med brug af Evans metodik. Men det kan lade sig gøre i f.eks. en før-efter uheldsevaluering. Eksempelvis kunne vejkanthak evalueres ved, at man ser på de steder, hvor plakaterne kun er opstillet i én kørselsretning, selvom de kunne være interessante også for den anden kørselsretning. Her bruges uheld i retningen med plakater som case og sammenstilles med den anden retning, og omvendt. For at tage højde for eventuelle forskelle i risiko i de to retninger bruges så et yderligere sæt data, nemlig uheld opdelt på retninger fra før plakaterne sættes op. Derved tages der højde for mange skævheder, dog kan regressionseffekten være forskellig for de to retninger.

Det er vanskeligere at tage højde for skævheder, når der ikke haves en før-efter situation. Effekten af vejbelysning er typisk blevet studeret ved at se på odds ratio

værdien bestemt af uheld i hhv. mørke og dagslys på veje hhv. med og uden belysning. Her vil der være flere skævheder, f.eks. eksponering i dagslys og mørke, uheld med belysningsmaster, sprituheld, natteblindhed, osv. Vejbelysningen i det åbne land er måske også sat op, hvor der i forvejen var de største uheldsproblemer i mørke. Således kunne endogene variable være på tale.

Case-control metoden er at foretrække frem for tværsnitsstudier, fordi man nærmest tvinges til at undgå samvariation mellem uafhængige variable, og til at få en forståelse for detaljer om trafik og vejtekniske forhold.

6.2.3 Bi- og multivariate modeller

Med bi- og multivariate modeller kan flere f.eks. skadesgrader og uheldstyper modelleres samtidigt i en og samme model. Ud fra en gennemgang af studier, hvor sådanne modeller er opstillet, forekommer det ikke hensigtsmæssigt, med det udviklingsstadium disse modeller er på, at benytte disse modeller til at opgøre tætheder (antal) af uheld og personskader opdelt efter skadesgrader, typer, osv., men alene at benytte disse modeller til at angive fordelingen af uheld og personskader efter skadesgrader, typer, mv. Det skyldes, at modellerne ikke er udviklet til at kunne håndtere flere spredningsparametre samtidigt, f.eks. en for hver skadesgrad.

Disse modeller kan opdeles i ordinale og nominelle modeller. I en ordinal model er respons (udkom) rangordnet, f.eks. dræbt, alvorlig skade, let skade og uskadte, mens det ikke er tilfældet i en nominel model, f.eks. eneuheld, bagendekollisioner, mødeuheld og tværkollisioner. I en ordinal model indgår én nyttefunktion med de uafhængige variable, samt en kumulativ sandsynlighedsfunktion, der fordeler, f.eks. efter skadesgrad, med baggrund i den resulterende værdi af nyttefunktionen. Både residualer for nyttefunktion og sandsynlighedsfunktionen er oftest antaget at være normeltfordelt. I en nominel model indgår en nyttefunktion for hver kategori, f.eks. hver uheldstype, med evt. forskellige uafhængige variable i de enkelte nyttefunktioner. Herefter fordeles, f.eks. efter uheldstype, med baggrund i den resulterende værdi af nyttefunktionen for den enkelte kategori samt summen af de resulterende værdier fra alle nyttefunktioner. En nominel model er derved mere fleksibel end en ordinal model, da flere nyttefunktioner indgår. En nominel model kræver derfor også flere observationer end en ordinal model for at opnå samme lave statistiske usikkerhed.

Der er – så vidt vides – ikke arbejdet med bi- og multivariate modeller inden for trafiksikkerhed i Danmark. Det er tvivlsomt, om der i multivariate modeller kan indgå uskadte personer, da uskadte passagerer ikke registreres i Danmark. Det er kun uskadte fører af køretøjer og uskadte fodgængere, der registreres. Således bliver antallet af uskadte personer på en måde trunke.

Multivariate modeller kan f.eks. belyse forskelle i fordelingen af skadesgrader eller uheldstyper mellem signalregulerede kryds og rundkørsler, eller opgøre ind-

virksomheden af hastighedsgrænser på skadesgrader. Modellerne kan også bruges til at belyse f.eks. sikkerhedssellers og personers alder indvirkning på skadesgrader.

6.3 Opsummering

Der er fortsat behov for at udføre med-uden studier, da nogle sikkerhedseffekter (SF'er) ikke kan undersøges på baggrund af de mere pålidelige før-efter studier. Baggrunden herfor er, at nogle typer af ombygninger forekommer sjældent, mens andre typer af ombygninger oftest forekommer samtidigt med andre ændringer. Et eksempel er, at vejbelysning i det åbne land oftest etableres i forbindelse med anlæg af vejudvidelser, signalreguleringer og rundkørsler. Det er sjældent, at vejbelysning er det eneste element, der etableres.

En liste med designelementer i kryds og på strækninger, der kunne være relevante at undersøge i med-uden studier, er opstillet. Listen anses ikke for at være udtømmende. Der kan være flere designelementer, der er relevante at studere.

Tre metodisk forskellige måder at udføre med-uden studier er gennemgået. Det er hhv. tværsnitsstudier, case-control studier samt bi- og multivariate modeller.

Den "letteste" metode, tværsnitsstudier, indebærer store risici for, at de estimerede SF'er er ukorrekte. Det er vanskeligt med denne metode at opstille en pålidelig dokumentation for den årsag-virkningssammenhæng, som en SF'er er udtryk for. Problemet er samvarierende og endogene uafhængige variable. For at undgå dette problem kan man enten udføre en vel-specificeret datafangst, hvilket kan resultere i få observationer, eller at tage højde for samvariation og endogenitet på baggrund af komplicerede simultane modelestimeringer.

Case-control metoden sætter store krav til datafangsten. Denne metodik gør det nemmere at undgå problemer med samvariation og endogenitet. Med den rette tilgang vil case-control studier ofte udgøre en rimelig pålidelig dokumentation. Problemet med case-control metoden er ofte få observationer.

Bi- og multivariate modeller kan anvendes til at opgøre fordelingen af uheld og personskader, f.eks. efter skadesgrader og uheldstyper. Det synes ikke rimeligt at benytte disse modeller til at opgøre tætheden af uheld og personskader.

Set i lyset af listen af designelementer, der kunne være relevante at undersøge, forekommer det, at tværsnitsstudier og case-control metoden er de mest relevante metoder at anvende. Det anbefales at gøre brug af en datafangst med mange krav til variable for trafik og fysisk udformning, frem for en metodik med komplekse korrektioner for skævheder, herunder samvariation og endogenitet.

Litteraturliste

AASHTO (2010): *Highway Safety Manual, 1st Edition, Volume 1-3*. American Association of State Highway and Transportation Officials, Washington DC, USA.

Abbess, C., Jarrett, D. og C. C. Wright (1981): Accidents at black spots: Estimating the effectiveness of remedial treatment with special reference to the 'regression-to-mean' effect. *Traffic Engineering + Control*, vol. 22, pp. 535-542.

Agrawal, R. og D. Lord (2006): Effect of Sample Size on Goodness-of-Fit Statistic and Confidence Intervals of Crash Prediction Models Subjected to Low Sample Mean Values. *Transportation Research Record*, no. 1950, pp. 35-43.

Ashton, W. D. (1966): *The theory of road traffic flow*. King's College, London, Storbritannien.

Bélanger, C. (1994): An Estimation of the Safety of 4 legged Unsignalized Intersections. *Proceedings of Transportation Research Board 73th Annual Meeting*, Washington DC, USA.

Council, F. M. og J. R. Stewart (1999): Safety Effects of the Conversion of Rural Two-Lane to Four-Lane Roadways Based on Cross-Sectional Models. *Transportation Research Record*, no. 1665, pp. 35-43.

Donnell, E. T., Porter, R. J., Shankar, V. N., Gross, F., Hu, W., Gemar, M. og J. S. Oh (2011): Estimating the safety effects of roadway lighting at intersections: A cross-sectional study approach. *Proceedings of Transportation Research Board 90th Annual Meeting*, Washington DC, USA.

Elvik, R. (2007): *State-of-the-art approaches to road accident black spot management and safety analysis of road networks*. TØI report 883/2007, Transport Økonomisk Institutt, Oslo, Norge.

Evans, L. (1986): Double pair comparison – a new method to determine how occupant characteristics affect fatality risk in traffic crashes. *Accident Analysis & Prevention*, vol. 18, pp. 217-227.

Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R. og L. K. Thomsen (1995): Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. *Accident Analysis & Prevention*, vol. 27, pp. 1-20.

Geedipally, S. R., Lord, D. og B-J. Park (2009): Analyzing Different Parametrizations of the Varying Dispersion Parameter as a Function of Segment Length. *Proceedings of Transportation Research Board 88th Annual Meeting*, Washington DC, USA.

Geedipally, S. R. og D. Lord (2011): Examining the Crash Variances Estimated by the Poisson-Gamma and Conway-Maxwell-Poisson Models. *Proceedings of Transportation Research Board 90th Annual Meeting*, Washington DC, USA.

Geedipally, S. R. og D. Lord (2010): Investigating the effect of modeling single-vehicle and multi-vehicle crashes separately on confidence intervals of Poisson-gamma models. *Accident Analysis & Prevention*, vol. 42, pp. 1273-1282.

Greibe, P. og S. Hemdorff (1995): *Uheldsmodel for bygader – Del 1: Model for 3- og 4-benede kryds*. Vejdirektoratet, notat 22, København, Danmark.

Greibe, P. og S. Hemdorff (1998): *Uheldsmodel for bygader – Del 2: Model for strækninger*. Vejdirektoratet, notat 59, København, Danmark.

Greibe, P. og S. Hemdorff (2001): *Håndbog i trafiksikkerhedsberegninger – Brug af uheldsmodeller og andre vurderinger*. Rapport 220, Vejdirektoratet, København, Danmark.

Gross, F, Jovanis, P. P., Eccles, K. og K-Y. Chen (2009): *Safety Evaluation of Lane and Shoulder Width Combinations on Rural, Two-Lane, Undivided Roads*. Federal Highway Administration, FHWA-HRT-09-031, Washington DC, USA.

Hadi, M. A., Aruldas, J., Chow, L-F. og J. A. Wattleworth (1993): Estimating Safety Effects of Cross-Section Design for Various Highway Types Using Negative Binomial Regression. *Transportation Research Record*, no. 1500, pp. 169-177.

Harwood, D. W., Bauer, K. M., Richard, K. R., Gilmore, D. K., Graham, J. L., Potts, I. B., Torbic, D. J. og E. Hauer (2007): *Methodology to Predict the Safety Performance of Urban and Suburban Arterials*. National Cooperative Highway Research Program, NCHRP web-only document 129, Transportation Research Board, Washington DC, USA.

Hauer, E. (1997): *Observational before-after studies in road safety*. Pergamon, Elsevier Science Ltd., Oxford, Storbritannien.

Hauer, E. (2001): Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. *Accident Analysis & Prevention*, vol. 33, pp. 799-808.

Hauer, E. (2004): Statistical Road Safety Modeling. *Transportation Research Record*, no. 1897, pp. 81-87.

Hauer, E. og J. Bamfo (1997): Two tools for finding what function links the dependent variable to the explanatory variables. *Proceedings of International Cooperation in Theories and Concepts in Traffic Safety (ICTCT) conference*, Lund, Sverige.

Hauer, E., Council, F. M., og Y. Mohammedshah (2004): Safety Models for Urban Four-Lane Undivided Road Segments. *Transportation Research Record*, no. 1897, pp. 96-105.

Hemdorff, S. (2006): *AP-parametre til uheldsmodeller – Baseret på data for 2001-2005*. Elektronisk notat, Vejdirektoratet, København, Danmark.

Heydecker, B. G. og J. Wu (2001): Identification of Sites for Road Accident Remedial Work by Bayesian Statistical Methods: An Example of Uncertain Inference. *Advances in Engineering Software*, vol. 32, pp. 859-869.

Jonsson, T., Lyon, C., Ivan, J. N., Washington, S., van Schalkwyk, I. og D. Lord (2009): Investigating Differences in the Performance of Safety Performance Functions Estimated for Total Crash Count and for Crash Count by Crash Type. *Proceedings of Transportation Research Board 88th Annual Meeting*, Washington DC, USA.

Joshua, S. C. og N. J. Garber (1990): Estimating truck accident rate and involvements using linear and Poisson regression models. *Transportation Planning and Technology*, vol. 15, pp. 41-58.

Jovanis, P. P. og H. L. Chang (1986): Modeling the relationship of accidents to miles travelled. *Transportation Research Record 1068*, pp. 42-51, Washington DC, USA.

Kim, D-G. og S. Washington (2006): The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. *Accident Analysis & Prevention*, vol. 38, pp. 1094-1100.

Lee, Y., Nelder, J. og Y. Pawitan (2006): *Generalized Linear Models with Random Effects: Unified Analysis via H-likelihood*. Chapman & Hall/CRC.

Lord, D., Washington, S. P. og J. N. Ivan (2005): Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis & Prevention*, vol. 37, pp. 35-46.

Lord, D., Geedipally, S. R., Persaud, B. N., Washington, S. P., van Schalkwyk, I., Ivan, J. N., Lyon, C. og T. Jonsson (2008): *Methodology to Predict the Safety*

performance of Rural Multilane Highways. National Cooperative Highway Research Program, NCHRP web-only document 126, Transportation Research Board, Washington DC, USA.

Lord, D. og B. N. Persaud (2000): Accident Prediction Models With and Without Trend. *Transportation Research Record*, no. 1717, pp. 102-108.

Lord, D. og M. Mahlawat (2009): Examining the Application of Aggregated and Disaggregated Poisson-gamma Models to Low Sample Mean Bias. *Proceedings of Transportation Research Board 88th Annual Meeting*, Washington DC, USA.

Lord, D., Kuo P-F. og S. R. Geedipally (2010): Comparing the Application of the Product of Baseline Models and Accident Modification Factors and Models with Covariates: Predicted Mean Values and Variance. *Proceedings of Transportation Research Board 89th Annual Meeting*, Washington DC, USA.

Lord, D. og F. Mannering (2010): The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A*, vol. 44, pp. 291-305.

Lyon, C. og B. Persaud (2002): Pedestrian Collision Prediction Models for Urban Intersections. *Transportation Research Record*, no. 1818, pp. 102-107.

Ma, J. og K. M. Kockelman (2006): Bayesian Multivariate Poisson Regression for Models of Injury Count, by Severity. *Proceedings of Transportation Research Board 85th Annual Meeting*, Washington DC, USA.

Ma, J., Kockelman, K. M., og P. Damien (2008): A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis & Prevention*, vol. 40, pp. 964-975.

Madsen, J. C. O. (2005): *Skadesgradsbaseret sortpletudpegning – fra crash prevention til loss reduction i de danske vejbestyrelses sortpletarbejde*. Ph.d. afhandling, Trafikforskningsgruppen, Institut for samfundsudvikling og planlægning, Aalborg Universitet, Aalborg, Danmark.

Maher, M. J. og I. Summersgill (1996): A comprehensive methodology for the fitting of predictive accident models. *Accident Analysis & Prevention*, vol. 28, pp. 281-296.

Maycock, G. og R. D. Hall (1984): *Accidents at 4-arm roundabouts*. TRL, report LR 1120, Crowthorne, Storbritannien.

McCullagh, P. og J. A. Nelder (1989): *Generalized linear models*. Chapman & Hall, Storbritannien.

- Miaou, S-P. og D. Lord (2003): Modeling Traffic Crash-Flow Relationships for Intersections: Dispersion Parameter, Functional Form, and Bayes versus Empirical Bayes. *Proceedings of Transportation Research Board 82th Annual Meeting*, Washington DC, USA.
- Miaou, S-P. og H. Lum (1993): Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis & Prevention*, vol. 25, pp. 689-793.
- Miaou, S-P., Lu, A. og H. S. Lum (1996): Pitfalls of Using R^2 To Evaluate Goodness of Fit of Accident Prediction Models. *Transportation Research Record 1542*, pp. 6-13, Washington DC, USA.
- Miranda-Moreno, L. F., Fu, L., Ukkusuri, S. og D. Lord (2009): How to incorporate accident severity and vehicle occupancy into the hotspot identification process? *Proceedings of Transportation Research Board 88th Annual Meeting*, Washington DC, USA.
- Mitra, S. og S. Washington (2007): On the nature of over-dispersion in motor vehicle crash prediction models. *Accident Analysis & Prevention*, vol. 39, pp. 459-468.
- Mountain, L., Fawaz, B. og D. Jarrett (1996): Accident prediction models for roads with minor junctions. *Accident Analysis & Prevention*, vol. 28, pp. 695-707.
- Mountain, L., Maher, M. og B. Fawaz (1998): The influence of trend on estimates of accidents at junctions. *Accident Analysis & Prevention*, vol. 30, pp. 641-649.
- Mountain, L., Maher, M. og B. Fawaz (1997): The effects of trend over time on accident model predictions. *Proceedings of European Transport Conference*, Cambridge, Storbritannien.
- Nukenine, S. K. L. og J. Daniel (2011): Impact of Rear-seat Occupant Belt Use on the Injury Severity of Belted Drivers. *Proceedings of Transportation Research Board 90th Annual Meeting*, Washington DC, USA.
- Persaud, B. og C. Lyon (2007): Empirical Bayes before-after safety studies: Lessons learned from two decades of experience and future directions. *Accident Analysis & Prevention*, vol. 39, pp. 546-555.
- Persaud, B., Lord, D. og J. Palmisano (2002): Calibration and Transferability of Accident Prediction Models for Urban Intersections. *Transportation Research Record*, no. 1784, pp. 57-64.
- Ragnøy, A., Christensen, P. og R. Elvik (2002): *Skadegradstetthet – SGT*. TØI rapport 618/2002, Transportøkonomisk Institutt, Oslo, Norge.

Satterthwaite, S. P. (1976): *Day to day variations in road accident frequency in Britain*. Traffic Studies Group, University College London, Storbritannien.

Song, J. J., Ghosh, M., Miaou, S. og B. Mallick (2006): Bayesian multivariate spatial models for roadway traffic crash mapping. *Journal of Multivariate Analysis*, vol. 97, pp. 246-273.

Stamatiadis, N., Pigman, J., Sacksteder, J., Ruff, W. og D. Lord (2009): *Impact of Shoulder Width and Median Width on Safety*. National Cooperative Highway Research Program, NCHRP report 633, Transportation Research Board, Washington DC, USA.

Sørensen, M. (2005): Amternes sortpletudpegning – Metode og erfaring. *Dansk Vejtidskrift*, vol. 82, no. 1, pp. 13-16.

Turner, S. og A. Nicholson (1998): Intersection accident estimation: The role of intersection location and non-collision flows. *Accident Analysis & Prevention*, vol. 30, pp. 505-517.

Vistisen, D. (2002): *Models and methods for hot spot safety work*. Ph.d. afhandling, Institut for informatik og matematisk modellering, Danmarks Tekniske Universitet, Lyngby, Danmark.