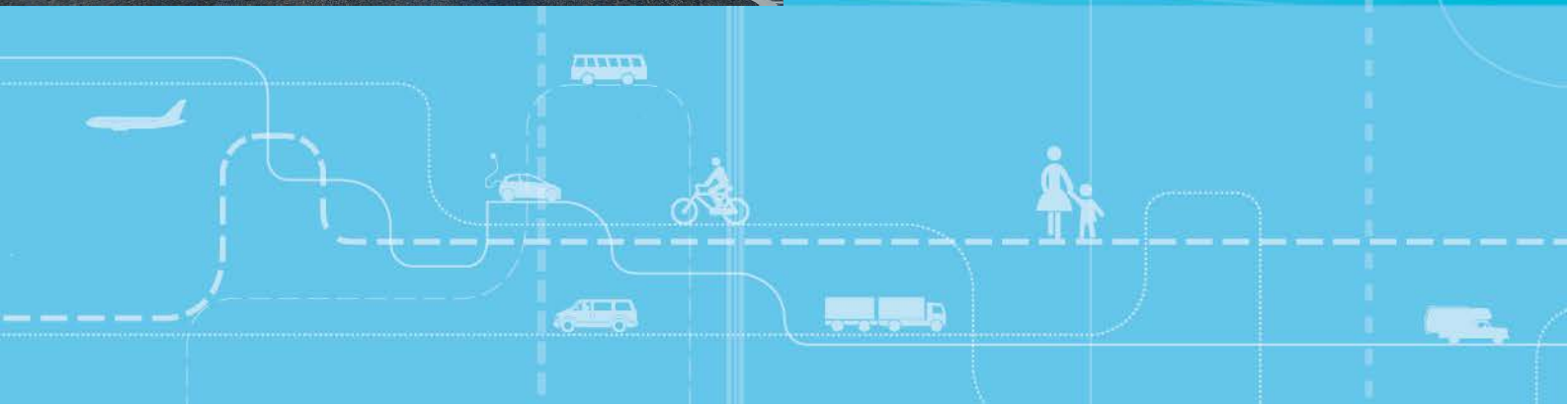


# Utvikling av ulykkesmodeller for ulykker på riks- og fylkesvegnettet i Norge (2010-2015)





# Utvikling av ulykkesmodeller for ulykker på riks- og fylkesvegnettet i Norge (2010-2015)

Alena Høye

Forsidebilde: Statens vegvesen

Transportøkonomisk institutt (TØI) har opphavsrett til hele rapporten og dens enkelte deler. Innholdet kan brukes som underlagsmateriale. Når rapporten siteres eller omtales, skal TØI oppgis som kilde med navn og rapportnummer. Rapporten kan ikke endres. Ved eventuell annen bruk må forhåndssamtykke fra TØI innhentes. For øvrig gjelder [åndsverklovens](#) bestemmelser.

ISSN 0808-1190

ISBN 978-82-480-2031-8 Papirversjon

ISBN 978-82-480-1759-2 Elektronisk versjon

Oslo, august 2016

**Tittel:** Utvikling av ulykkesmodeller for ulykker på riks- og fylkesvegnettet i Norge (2010-2015)

**Forfatter:** Alena Høye  
**Dato:** 08/2016  
**TØI rapport** 1522/2016  
**Sider:** 49  
**ISBN papir:** 978-82-480-2031-8  
**ISBN elektronisk:** 978-82-480-1759-2  
**ISSN:** 0808-1190  
**Finansieringskilde:** Statens vegvesen, Vegdirektoratet

**Prosjekt:** 4023J – Ulykkesmodeller  
**Prosjektleder:** Alena Høye  
**Kvalitetsansvarlig:** Michael W.J. Sørensen  
**Fagfelt:** 21 - Sikkerhet og tiltak

**Emneord:** Trafikksikkerhet;  
Ulykkesmodell;  
Negativ binomial modell;  
Variabel  
overspredningsparameter

#### Sammendrag:

Ulykkesmodeller er utviklet for riks- og fylkesvegnettet i Norge for å beregne normale antall personskaueulykker, lett skadde, hardt skadde, drepte og samlet antall drepte eller hardt skadde som en funksjon av trafikkmengde, strekningslengde, vegtype, fartsgrense, antall kjørefelt, antall kryss, automatisk fartskontroll, vegbelysning og andre vegegenskaper. Den valgte modellformen (negative binomialmodeller med variabel overspredningsparameter) gjør det mulig å benytte resultatene bl.a. i evalueringer av trafikksikkerhetstiltak med den empiriske Bayes metoden som i dag anses som den beste for å redusere risikoen for å systematisk overestimere virkninger av trafikksikkerhetstiltak i før-etter studier. Resultatene kan også brukes i ulike verktøy som brukes av Statens vegvesen for å gjøre analyser av ulykkessituasjonen i vegnettet og for å beregne forventede virkninger av trafikksikkerhetstiltak. Modellene er oppdatert fra TØI-rapport 1323/2014 med et nytt datagrunnlag fra 2010-2015.

*Transportøkonomisk Institutt  
Gaustadalleen 21, 0349 Oslo  
Telefon 22 57 38 00 - [www.toi.no](http://www.toi.no)*

**Title:** Development of crash prediction models for national and county roads in Norway (2010-2015)

**Author:** Alena Høye  
**Date:** 08/2016  
**TØI Report:** 1522/2016  
**Pages:** 49  
**ISBN Paper:** 978-82-480-2031-8  
**ISBN Electronic:** 978-82-480-1759-2  
**ISSN:** 0808-1190  
**Financed by:** The Norwegian Public Roads Administration

**Project:** 4023J - Crash prediction models  
**Project Manager:** Alena Høye  
**Quality Manager:** Michael W.J. Sørensen  
**Research Area:** 21 - Safety and crash countermeasures  
**Keywords:** Road safety  
Crash prediction model  
Generalized negative binomial model  
Overdispersion

#### Summary:

Crash prediction models for national and county roads in Norway were developed in order to calculate predicted numbers of injury crashes, slightly injured, seriously injured, fatalities, and the total number of killed or seriously injured, as a function of traffic volume, segment length, road category, speed limit, number of lanes, number of intersections, speed cameras, road lighting, and other road characteristics. Models were calculated as generalized negative binomial models (negative binomial models with a variable overdispersion parameter). Results can be used in evaluations of road safety measures with the empirical Bayes method which is regarded as state of the art with respect to control for regression to the mean. The models can also be applied in conjunction with several tools of the Norwegian Public Roads Administration for road safety analyses and estimation of expected effects of road safety measures. The models are updated from TØI Report 1323/2014 based on crash data from 2010-2015.

**Language of report:** Norwegian

*Institute of Transport Economics  
Gaustadalleen 21, 0349 Oslo, Norway  
Telefon 22 57 38 00 - [www.toi.no](http://www.toi.no)*

# Forord

På oppdrag av Vegdirektoratet er det utviklet ulykkesmodeller for å beregne normale antall ulykker, skadde og drepte på riks- og fylkesveger i Norge. Normale ulykkes- og skadetall som beregnes med modellene kan benyttes eksempelvis i programmer som TS-Effekt og S-kost. Den valgte modellformen gjør det også mulig å benytte resultater fra modellberegninger i evalueringer av trafikksikkerhetstiltak med den empiriske Bayes metoden.

TØIs prosjektleder har vært Alena Høye som også har skrevet rapporten. Oppdragsgivers kontaktperson har vært Arild Ragnøy. Arild Ragnøy og Henning Fransplass, begge fra Vegdirektoratet, har bidratt med nyttige kommentarer til tidligere versjoner av rapporten. Datagrunnlaget for modellberegningene er i hovedsak skaffet til veie av Håvard Lerstad fra Triona, basert på data i den nasjonale vegdatabanken (NVDB). Statens vegvesen har i tillegg skaffet informasjon om forsterket midtoppmerking og vegbelysning.

Michael W.J. Sørensen har stått for kvalitetssikring av rapporten. Trude Rømming har tilrettelagt rapporten for utgivelse i papirform og elektronisk.

Oslo, august 2016  
Transportøkonomisk institutt

*Gunnar Lindberg*  
direktør

*Michael W.J. Sørensen*  
avdelingsleder



# Innhold

## Sammendrag

### Summary

<b>1</b>	<b>Bakgrunn</b> .....	<b>1</b>
<b>2</b>	<b>Modellform og prediktorvariabler</b> .....	<b>4</b>
2.1	Modellform.....	4
2.2	Ulykker og skadde .....	5
2.3	Segmentlengde og antall år.....	6
2.4	Trafikkmengde.....	7
2.5	Fartsgrense.....	12
2.6	Antall kjørefelt .....	13
2.7	Kryss, rundkjøringer og ramper .....	14
2.8	Type veg.....	18
2.9	Middeler og midtrekkverk.....	20
2.10	Forsterket midtoppmerking.....	21
2.11	Automatisk trafikkontroll (ATK).....	22
2.12	Vegbelysning .....	26
2.13	Broer.....	28
2.14	Fylke .....	30
2.15	Konstantterm .....	31
2.16	Overspredningsparameter .....	31
<b>3</b>	<b>Modellene</b> .....	<b>32</b>
3.1	Koeffisientene og beregningseksempel.....	32
3.2	Hvor gode er modellene? .....	34
<b>4</b>	<b>Trendfaktor</b> .....	<b>37</b>
4.1	Trendjustering av de forrige modellene .....	37
4.2	Trendjustering av predikerte (normale) ulykkestall .....	38
4.3	Trendjustering av overspredningsparameteren.....	40
4.4	Trendjustering av trafikkmengden.....	42
<b>5</b>	<b>Bruerveiledning for Ulykkesmodeller.xlsx</b> .....	<b>44</b>
<b>6</b>	<b>Referanser</b> .....	<b>46</b>





## Sammendrag

# Utvikling av ulykkesmodeller for ulykker på riks- og fylkesvegnettet i Norge (2010-2015)

TØI rapport 1522/2016

Forfatter: Alena Høye

Oslo, 2016, 5 sider

---

*Ulykkesmodeller er utviklet for riks- og fylkesvegnettet i Norge for å beregne normale antall personskadeulykker, lett skadde, hardt skadde, drepte og samlet antall drepte eller hardt skadde som en funksjon av trafikkmengde, strekningslengde, vegtype, fartsgrense, antall kjørefelt, antall kryss, ATK, vegbelysning og andre vegegenskaper. Den valgte modellformen (negative binomialmodeller med variabel overspredningsparameter) gjør det mulig å benytte resultatene bl.a. i evalueringer av trafikksikkerhetstiltak med den empiriske Bayes metoden som i dag anses som den beste for å redusere risikoen for å systematisk overestimere virkninger av trafikksikkerhetstiltak i før-etter studier. Resultatene kan også brukes i ulike verktøy som brukes av Statens vegvesen for å gjøre analyser av ulykkes situasjonen i vegnettet og for å beregne forventede virkninger av trafikksikkerhetstiltak.*

Ulykkesmodellene som er utviklet på oppdrag av Vegdirektoratet er beregnet med ulykkesdata fra den nasjonale vegdatabanken (NVDB) fra 2010-2015. Rapporten bygger på tidligere rapporter om utvikling av ulykkesmodeller for det norske vegnettet (Ragnøy, Christensen & Elvik, 2002; Ragnøy & Elvik, 2003; Høye, 2014). Ragnøy et al. (2002) har utviklet ulykkesmodeller og en metode for å vekte sammen registrerte og normale ulykkestall til forventede ulykkestall. Basert på denne metoden har Ragnøy & Elvik (2003) gjort en trafikksikkerhetsanalyse av stamvegnettet. Høye (2014) har utviklet modeller som bygger på ulykkesdata fra 2006-2011 og en videreutvikling av modellene og metodene som er brukt av Ragnøy et al. (2002). Den aktuelle rapporten presenterer en oppdatering av modellene som er utviklet av Høye (2014).

Modellene kan brukes for å beregne normale antall personskadeulykker (Psu), lett skadde (LS), hardt skadde (HS), drepte og det samlede antall drepte eller hardt skadde (D/HS) på riks- og fylkesvegnettet. Normale ulykkestall beregnes som en funksjon av trafikkmengden og ulike vegegenskaper. Et normalt ulykkestall er antall ulykker som skjer i gjennomsnitt på en veg av en gitt lengde med en gitt trafikkmengde og gitte vegegenskaper i ett gitt tidsrom. Modellformen som er benyttet (negative binomialmodeller med variabel overspredningsparameter) gjør det også mulig å beregne forventede ulykkestall som en funksjon av normale ulykkestall, registrerte ulykkestall og overspredningsparameteren. Forventede ulykkestall er antall ulykker som man på lang sikt forventer at vil skje på en vegstrekning, ut fra generelle vegegenskaper og ulykkene som har skjedd på strekningen. Overspredningsparameteren brukes for å vekte sammen normale og registrerte ulykkestall. Overspredningsparameteren varierer som en funksjon av trafikkmengde og strekningslengde, slik at registrerte ulykkestall får en større vekt på vegstrekninger med store normale ulykkestall.

Når man beregner det forventede antall ulykker på en strekning med unormalt mange ulykker antar man at resultatet viser hvor mange ulykker man kan forvente på strekningen på lang sikt. Man tar dermed hensyn til at antall ulykker kan ha vært tilfeldig høyt og at antall ulykker på langt sikt mest sannsynlig vil ligge nærmere det normale ulykkestallet. Ved å benytte forventede ulykkestall, istedenfor registrerte ulykkestall, kan man i en før-etter studie unngå å overestimere virkningen av trafikksikkerhetstiltak. Tiltak settes ofte inn på vegstrekninger hvor det har vært spesielt mange ulykker og derfor vil antall ulykker i mange tilfeller gå ned (komme nærmere gjennomsnittet) uansett om man setter inn et tiltak eller ikke.

## Modell og prediktorvariabler

**Modellformen** som er brukt for alle skadegrader er en negativ binomialmodell (NB modeller) med variabel overspredningsparameter. Modellformen er følgende:

$$E(n) = e^{\sum_i \text{Prediktor}_i * \text{Koeff}_i}$$

$E(n)$  er det predikerte (normale) ulykkestall (dvs. antall Psu, LS, HS, drepte eller D/HS), prediktorene er trafikkmengden og ulike vegegenskaper og  $i$  er subskript for prediktorene. I tillegg inneholder modellen prediktorer og koeffisienter for overspredningsparameteren.

Fordelen med denne modellformen er at den tar hensyn til overspredningen i ulykkesdata og at overspredningen varierer som en funksjon av trafikkmengde, segmentlengde og antall år. Dermed er det mulig å benytte resultatene bl.a. i evalueringer av trafikksikkerhetstiltak med den empiriske Bayes (EB) metoden. Denne anses i dag som den beste for å redusere risikoen for å systematisk overestimere virkninger av trafikksikkerhetstiltak i før-etter studier. I en EB studie sammenligner man det registrerte ulykkestallet etter at et tiltak er satt inn på en veg med det forventede ulykkestallet. Det forventede ulykkestallet beregnes ved å vekte sammen det registrerte ulykkestall og det normale ulykkestall. Det normale ulykkestall er antall ulykker som er predikert av modellen, dvs. det gjennomsnittlige antall ulykker på en vegstrekning med de gitte vegegenskapene. Vekten er en funksjon av overspredningsparameteren og det normale antall ulykker. Det forventede antall ulykker ligger alltid mellom det normale og det registrerte antall.

**Prediktorvariablene** i modellene for alle skadegrader er følgende:

- **Segmentlengde og antall år:** Segmentlengde og antall år inngår i alle modellene som eksponeringsvariabler, slik at det normale antall ulykker alltid øker proporsjonalt med segmentlengde og antall år. Antall år er prediktor fordi det ikke foreligger data fra alle år for alle segmentene; segmenter med vesentlige endringer (som for eksempel endringer av fartsgrensen) inngår kun med ulykkesdata fra etter endringen i datafilen som ligger til grunn for modellberegningene.
- **Trafikkmengde (ÅDT):** Trafikkmengden inngår i modellene som  $\ln(\hat{ADT})$ . Dermed er den prosentvise økningen av normale ulykkestall uavhengig av hvor høy trafikkmengden er i utgangspunktet. Hver økning av trafikkmengden med  $X$  prosent medfører alltid den samme prosentvise økningen for hver skadegrad.
- **Fartsgrense:** For hver fartsgrense er det definert en dummyvariabel som er én ved den respektive fartsgrensen og null ellers. Dummyvariabler er valgt framfor en tallvariabel for å ta hensyn til generelle forskjeller mellom veger med ulike fartsgrenser som gjør at antall ulykker kan gå opp og ned mellom ulike fartsgrenser.

- **Antall kjørefelt:** For antall kjørefelt er det definert dummyvariabler på samme måte og av de samme grunnene som beskrevet for fartsgrense. Segmenter med kun ett kjørefelt er utelatt fra datasettet fordi dette er en liten og svært heterogen gruppe av helt ulike typer veger og dermed ikke noen meningsfylt kategori.
- **Plankryss, rundkjøringer og ramper (planskilte kryss):** For X-kryss, T-kryss, rundkjøringer og ramper er det beregnet variabler som er den naturlige logaritmen av antallet per kilometer pluss én (pluss én for å unngå å ta logaritmen av null).
- **Type veg:** Det er definert fem dummyvariabler for ulike typer veg: 1 = motorveg; 2 = tofeltsveg med planskilte kryss (tidligere betegnet som motortrafikkveg); 3 = TEN-T-veg (ikke motorveg / tofeltsveg med planskilte kryss); 4 = øvrig europa-/riksveg (ikke motorveg / tofeltsveg med planskilte kryss / TEN-T-veg); 5 = fylkesveg. Disse vegklassene er valgt fordi det finnes generelle forskjeller i vegstandarden (bl.a. kjørefelt- og skulderbredde) mellom klassene; f.eks. har alle motorveger midtrekkverk og TEN-T-veger har generelt en noe høyere standard enn øvrige europa- og riksveger).
- **ATK:** Det er definert tre dummyvariabler for automatisk trafikkontroll (ATK), punkt-ATK (PATK), streknings-ATK (SATK) i én retning og SATK i begge retninger. For PATK er det ikke skilt mellom strekninger med PATK i en vs. begge retninger. Strekninger som ligger inntil 100 meter oppstrøms og 3 km nedstrøms for et ATK-punkt regnes som PATK-strekninger. For SATK regnes alle strekninger mellom de to fotoboksene som SATK-strekninger.
- **Vegbelysning:** Det er definert én dummyvariabel for vegbelysning.
- **Midtdeler / -rekkverk:** Alle segmentene er delt inn i segmenter med både midtdeler og midtrekkverk, segmenter med midtdeler og uten midtrekkverk, segmenter med midtrekkverk og uten midtdeler, og segmenter uten verken midtdeler eller midtrekkverk. For hver av de fire gruppene er det definert én dummyvariabel. Det er ikke definert to separate dummyvariabler (én for med/uten midtdeler og én for med/uten midtrekkverk) fordi det kan være interaksjonseffekter mellom midtdeler og midtrekkverk, dvs. at virkningen av midtrekkverk kan være forskjellig på veger med og uten midtdeler og at virkningen av midtdeler kan være forskjellig på veger med og uten midtrekkverk.
- **Forsterket midtoppmerking:** Forsterket midtoppmerking er kombinasjonen av vanlig oppmerking og rumleriller. Denne variabelen omfatter kun forsterket midtoppmerking på veger som ikke er motorveger og som ikke har midtdeler eller midtrekkverk. Det skilles ikke mellom ulike typer forsterket midtoppmerking.
- **Fylke:** For hvert fylke er det definert én dummyvariabel. Disse skal fange opp generelle forskjeller mellom fylkene (f.eks. forskjeller i topografi, vær og befolkningstetthet).
- **Konstantterm:** Alle modellene inneholder en konstantterm.

For hver av modellene er det i tillegg estimert koeffisienter for å beregne **overspredningsparameteren** som en funksjon av segmentlengde, antall år og trafikkmengde. Overspredningen synker med økende segmentlengde, antall år og trafikkmengde.

## Hvor gode er modellene?

Det er beregnet flere ulike goodness-of-fit (GOF) indikatorer: Pseudo-R<sup>2</sup>, mean square prediction error (MSPE) og Elvik-index. Når man sammenligner ulike modeller (med ulike prediktorer) viser GOF indikatorene kun svært små forskjeller mellom ulike modellvarianter for hver skadegrad. Sammenlignet med modellene av Høye (2014) er de aktuelle modellene omtrent like «gode» eller bedre. I motsetning til modellene til Høye (2014) er avvikene mellom de faktiske og predikerte totale ulykkestallene forholdsvis små for alle skadegrader. Sorterer man avvikene etter trafikkmengden er det likevel en del avvik mellom predikerte og registrerte ulykkestall i enkelte ÅDT-grupper.

## Regneark Ulykkesmodeller

I vedlagte regneark Ulykkesmodeller 2016.xlsx kan man beregne:

- Predikerte (normale) antall Psu, LS, HS, drepte og D/HS som en funksjon av de prediktorvariablene som er beskrevet ovenfor
- En overspredningsparameter for hver skadegrad
- En vekt for hver skadegrad som kan brukes til å beregne forventede ulykkes/-skadetall
- Hvis man i tillegg oppgir registrerte ulykkes- og skadetall, beregnes også forventede antall Psu, LS, HS, drepte og D/HS.

Det er også mulig å omregne resultater til år mellom 1997 og 2030.

## Summary

# Development of crash prediction models for national and county roads in Norway (2010-2015)

TOI Report 1522/2016

Author: Alena Høye

Oslo, 2016 49 pages, Norwegian language

---

*Crash models for national and county roads in Norway were developed in order to calculate predicted numbers of injury crashes, slightly injured, seriously injured, fatalities and the total number of killed or seriously injured, as a function of traffic volume, segment length, road category, speed limit, number of lanes, number of intersections, speed cameras, road lighting and other road characteristics. Models were calculated as generalized negative binomial models (negative binomial models with a variable overdispersion parameter). Results can be used in evaluations of road safety measures with the empirical Bayes method which is regarded as state of the art with respect to control for regression to the mean. The models can also be applied in conjunction with several tools of the Norwegian Public Roads Administration for road safety analyses and estimation of expected effects of road safety measures.*

Crash models were developed on behalf of the Norwegian Public Roads Administration based on crash data in the national roads data base from the years 2010-2015. The models can be used to calculate predicted numbers of injury crashes, slightly injured, seriously injured, fatalities and the total number of killed or seriously injured (KSI) on national and county roads as a function of traffic volume, segment length, road category, speed limit, number of lanes, number of intersections, speed cameras, road lighting, and other road characteristics. Models were calculated as generalized negative binomial models (negative binomial models with a variable overdispersion parameter) which allows the calculation of expected crash numbers as a function of the model predictions, crash counts and an overdispersion parameter. Expected crash numbers refer to the number of crashes that can be expected on a road segment in the long run, based on general road characteristics and the specific crash history of the road segment in question. The statistical weights for model predictions and crash counts are calculated as functions of the overdispersion parameter. The overdispersion parameter varies as a function of traffic volume, segment length, and number of years. The statistical weight for the crash counts increases with increasing segment length and traffic volume.

The expected numbers of crashes will always be between the model prediction and the observed numbers of crashes. In before-after studies of the safety effects of measures at high-crash locations, regression to the mean is less likely to affect the results when the analysis is based on expected crash numbers in the before period instead of observed crash counts. Regression to the mean occurs when the observed number of crashes in the before period was exceptionally high. One would then expect the number of crashes to decrease in the after period even without any (effective) safety measures.

## Model and predictors

The models were developed as generalized negative binomial (NB) models in which the overdispersion parameter is estimated as a function of traffic volume, segment length and number of years. The model form is as follows:

$$E(n) = e^{\sum_i \text{Predictor}_i * \text{Coeff}_i}$$

$E(n)$  is the predicted number of crashes (i.e. number of injury crashes / injuries / fatalities), predictors are traffic volume and a number of road characteristics, and  $i$  is the subscript for the predictors.

A generalized NB model was chosen because it takes into account that crash counts usually are overdispersed, and that overdispersion is not a constant but depending on traffic volumes, segment length and the number of years. The model results can be used in before-after evaluations of road safety measures with the empirical Bayes (EB) method. The EB method controls for regression to the mean by comparing the observed number of crashes in the after period with the expected number of crashes. The expected number of crashes is a function of the actual crash count and the model prediction of the number of crashes for the same road section and period of time. Crash counts and model predictions are weighted with a function of the overdispersion parameter. The expected number of crashes is always between the crash count and the model prediction.

The crash models for all dependent variables (injury crashes and numbers of injuries / fatalities) are based on the following predictors:

- **Segment length and number of years:** The coefficients for the natural logarithm of segment length and number of years are set equal to one, thus normal crash numbers increase proportionally with segment length and number of years. The number of years is a predictor because segments with substantial changes (e.g. speed limit reductions) are represented with crash data only from after the change was made in the data the models are based on.
- **Traffic volume:** The predictors for traffic volume is the natural logarithm of AADT (annual average daily traffic).
- **Speed limit:** For each speed limit a dummy variable is defined in order to take into account that crash numbers not necessarily are a monotonous (or other) function of speed limit because of general differences between roads with different speed limits (e.g. many roads with a 70 km/h speed limit had previously an 80 km/h speed limit that was reduced because of exceptionally high crash numbers).
- **Number of lanes:** For each number of lanes a dummy variable is defined for the same reasons as for speed limit. Segments with one lane are a highly heterogeneous group of different kinds of roads and only a small proportion of all data. They are therefore omitted from the data.
- **At-grade intersections, roundabouts and ramps (grade separated intersections):** For four-armed intersections, three-armed intersections, roundabouts, and ramps the natural logarithm of the number per kilometer plus one is calculated (one is added in order to avoid taking the logarithm of zero on segments without any of these).
- **Road category:** A dummy variable is defined for each of the following road categories: Motorway; two-lane road with grade separated intersections; TEN-T road (other than motorway or two-lane road with grade separated intersections); European or state highway (other than those previously mentioned); county road.

- **Speed cameras and section control:** Three dummy variables are defined for speed cameras, speed camera, section control in one direction and section control in both directions. Roads “with speed camera” include 100 meter upstream and 3 km downstream of speed cameras. Road “with section control” include the whole length of road between the two section control speed cameras.
- **Road lighting:** One dummy variable is defined for road lighting.
- **Median and median guardrail:** A dummy variable is defined for each of the following: Median with guardrail; median without guardrail; guardrail that separates opposing directions of traffic without median; neither guardrail nor median. Dummy variables for the presence vs. absence of either median or guardrail would not allow to detect interaction effects between guardrail and median presence, which is why four dummy variables for each possible combination were defined.
- **Center line rumble strips:** This variable comprises only centerline rumble strips on roads other than motorways with no median or median guardrail.
- **County:** For each county a dummy variable is defined. These variables are meant to represent general differences between counties such as topography, weather and population density.
- **Constant:** A constant term is included in all models.

Additionally, all models contain coefficients for calculating the **overdispersion parameter** as a function of the natural logarithms of segment length, number of years and traffic volume. The overdispersion parameter decreases with increasing segment length, number of years and traffic volume.

## Crash model spreadsheet

The attached spreadsheet Ulykkesmodeller 2016.xlsx can be used to calculate:

- Model predictions for number of injury crashes, slightly injured, seriously injured, fatalities and KSI
- An overdispersion parameter for each dependent variable
- A statistical weight for each dependent variable that can be used to calculate the expected numbers of injury crashes, slightly injured, seriously injured, fatalities and KSI
- If observed numbers of crashes / injuries are entered, expected numbers of crashes / injuries are calculated as well.

It is also possible to convert the results to years between 1997 and 2030.





# 1 Bakgrunn

Rapporten beskriver ulykkesmodeller som er utviklet på oppdrag av Vegdirektoratet med ulykkesdata fra den nasjonale vegdatabanken (NVDB) fra årene 2010-2015. Modellene kan brukes for å beregne det normale antall personskadeulykker (Psu), lett skadde (LS), hardt skadde (HS), drepte (D), samt hardt skadde eller drepte (D/HS) på riks- og fylkesveger i Norge. I denne rapporten brukes «ulykkestall» for lesbarhetens skyld synonymt med antall personskadeulykker, skadde eller drepte.

Rapporten bygger på tidligere rapporter om utvikling av ulykkesmodeller for det norske vegnettet (Ragnøy, Christensen & Elvik, 2002; Ragnøy & Elvik, 2003; Høye, 2014). Ragnøy et al. (2002) har utviklet ulykkesmodeller og en metode for å vekte sammen registrerte og normale ulykkestall til forventede ulykkestall. Basert på denne metoden har Ragnøy & Elvik (2003) gjort en trafikksikkerhetsanalyse av stamvegnettet. Høye (2014) har utviklet modeller som bygger på ulykkesdata fra 2006-2011 og en videreutvikling av modellene og metodene som er brukt av Ragnøy et al. (2002). Den aktuelle rapporten presenterer en oppdatering av modellene som er utviklet av Høye (2014). Forskjeller til modellene av Høye (2014) er følgende:

- Trafikkmengde inngår kun med  $\ln(\text{ÅDT})$  som prediktor ( $\ln(\text{ÅDT})^2$  som også var med i de forrige modellene ble droppet).
- Kurve- og stigningsvariablene er droppet fra modellene da disse variablene har vist seg å være lite hensiktsmessige i den praktiske anvendelsen av modellene.
- Nye variabler i modellene er vegbelysning, samt punkt- og streknings-ATK.

For øvrig inngår de samme prediktorvariablene i modellene, det er de samme vegene som inngår i datagrunnlaget, og modellformen er også den samme.

For flere faktorer er det gjort testberegninger med ulike konstellasjoner av prediktorvariabler, dels fordi variablene er nye i modellene og dels fordi variabeldefinisjonene i datamaterialet er endret noe. Dette gjelder:

- Trafikkmengde
- Ramper
- ATK
- Vegbelysning
- Broer (som er tilgjengelig i datamateriale men som ikke er med som prediktor i de endelige modellene)

Modellene er beregnet basert på data for riks- og fylkesvegnettet (unntatt ettfeltsveger, rundkjøringer og ramper). De fullstendige endelige modellene er vist i kapittel 3, samt i vedlagte regneark *Ulykkesmodeller 2016.xlsx*. I det vedlagte regnearket kan man beregne normale og forventede antall Psu, LS, HS, drepte og D/HS, samt overspredningsparametere og vektorer for empiriske Bayes (EB)-beregninger. Det er også mulig å gjøre trendjusteringer for å omregne normale ulykkes- og skadetall, samt overspredningsparametere, til år mellom 1997 og 2030.

**Normale ulykkestall:** Ulykkesmodellene er statistiske modeller som kan brukes til å beregne normale ulykkestall som en funksjon av en rekke ulike vegegenskaper, bl.a. strekningslengde, trafikkmengde, fartsgrense og antall kjørefelt. For hver vegegenskap er det estimert en **koeffisient** for hver skadegrad. Koeffisienten er et tall som ganges med vegegenskapen (eller en funksjon av vegegenskapen) før denne inngår i formelen for modellberegningen.

Et normalt ulykkestall er antall ulykker som skjer i gjennomsnitt på en veg av en gitt lengde og med de vegegenskapene som inngår i modellberegningen. Normale ulykkestall er som regel ikke nøyaktig identiske med registrerte ulykkestall på hver eneste vegstrekning, men i et større datamateriale kan man forvente et det kun er et lite avvik mellom det totale antall registrerte og normale ulykkestall. Ulykkesmodellene er, forenklet sagt, beregnet slik at avvikene mellom registrerte og normale ulykkestall er så små som mulig.

**Forventede ulykkestall:** Forventede ulykkestall er antall ulykker som man på lang sikt forventer at vil skje på en konkret vegstrekning. Antakelsen bak beregningen av forventede ulykkestall er at registrerte antall ulykker er påvirket av:

- Vegegenskapene som inngår i modellberegningen og dermed i beregningen av normale ulykkestall.
- Variasjon rundt det normale ulykkestall som skyldes delvis tilfeldigheter og delvis systematiske faktorer som ikke er fanget opp av ulykkesmodellen (Hauer et al., 2002).

Forventede ulykkestall beregnes som en funksjon av:

- Normale ulykkestall: Dermed er det tatt hensyn til den generelle sammenhengen mellom en rekke vegegenskaper og antall ulykker.
- Registrerte ulykkestall: Dermed er det tatt hensyn til at normale ulykkestall ikke har en perfekt sammenheng med faktiske ulykkestall, og at det faktiske ulykkestall kan være påvirket av faktorer som ikke er fanget opp av ulykkesmodellen som brukes for å beregne det normale ulykkestall.
- Overspredningsparameter: Denne inngår i vekten for å beregne et vektet gjennomsnitt av normale og registrerte ulykkestall:

$$\text{Forv. ul.} = a * \text{Norm. ul.} + (1 - a) * \text{Reg. ul.}$$

Vekten  $a$  i denne formelen beregnes som en funksjon av det predikerte ulykkestall og overspredningsparameteren:

$$a = \frac{1}{1 + \frac{\text{Predikert ulykkestall}}{\text{Overspredningsparameter}}}$$

Overspredningsparameteren, og dermed også vekten, er jo mindre desto større det normale ulykkestall er. Dermed får registrerte ulykkestall en større vekt på vegstrekninger med store normale ulykkestall.

***Anvendelsesmulighet for forventede ulykkestall, empiriske Bayes (EB) studier:***

Forventede ulykkestall kan bl.a. brukes i EB analyser for å kontrollere for regresjonseffekter. I en EB studie sammenligner man det registrerte ulykkestallet etter at et tiltak er satt inn på en veg med det forventede ulykkestallet. Det forventede ulykkestallet etter at tiltaket er satt inn beregnes som en funksjon av det forventede ulykkestallet før tiltaket ble satt inn (som beskrevet ovenfor) og den gjennomsnittlige endringen av antall ulykker på lignende veger. Ved å beregne det forventede ulykkestallet før tiltaket er satt inn tar man hensyn til at det registrerte antall ulykker før tiltaket ble satt inn kan ha vært tilfeldig høyt. Dermed kontrollerer man for regresjonseffekten, dvs. at antall ulykker mest sannsynlig går ned (er nærmere gjennomsnittet) etter å ha vært unormalt høyt i en periode.

## 2 Modellform og prediktorvariabler

Ulykkesmodellene er beregnet basert på data for riks- og fylkesvegnettet i årene 2010-2015. Kun vanlige strekninger inngår i datagrunnlaget. Ettfeltsveger, rundkjøringer, ramper, tunneler og kommunale veger inngår ikke i datagrunnlaget. Segmenter hvor trafikkmengden (årsdøgntrafikk, ÅDT), antall år, lengde fartsgrense eller antall kjørefelt er uoppgitt inngår heller ikke i datagrunnlaget. Modellformen og de enkelte prediktorvariablene er beskrevet i de følgende avsnittene.

### 2.1 Modellform

Modellformen som er brukt for alle skadegrader er en negativ binomialmodell (NB modeller) med variabel overspredningsparameter. Modellformen er følgende:

$$E(n) = e^{\sum_i \text{Prediktor}_i * \text{Koeff}_i}$$

$E(n)$  er det predikerte (normale) ulykkestall (dvs. antall Psu, LS, HS, DR eller D/HS), prediktorene er trafikkmengden og ulike vegsegenskaper,  $i$  er subskript for prediktorene.

**Negative binomialmodeller** tar hensyn til overspredning i dataene, dvs. at spredningen er større enn man ville forvente hvis fordelingen av antall ulykker på enkelte vegsegmenter hadde fulgt en poissonfordeling. En negativ binomialfordeling er bestemt av to parametere, gjennomsnittet og overspredningsparameteren. En poissonfordeling derimot er bestemt av kun én parameter, gjennomsnittet. Legger man en poissonfordeling til grunn kan både koeffisientene og konfidensintervallene for koeffisientene være påvirket av overspredningen i dataene og dermed ikke være forventningsrette (skjeve).

En vanlig NB modell beregner en **fast overspredningsparameter**.

Overspredningsparameteren indikerer hvor mye de registrerte ulykkestallene varierer rundt de predikerte ulykkestallet (Hauer, 2001). Overspredningsparameteren kan brukes for å beregne vekten i EB analyser (se kapittel 1). Et problem med en fast overspredningsparameter er at predikerte ulykkestall er avhengige av i hvor lange delstrekninger man deler opp en strekning. Et annet problem er at man i EB analyser ikke har mulighet for å ta hensyn til at overspredningen varierer som en funksjon av strekningslengden, trafikkmengden mv. (Hauer, 2001).

Ved å beregne en **variabel overspredningsparameter** (istedenfor en fast overspredningsparameter som i vanlig NB regresjon) tar man hensyn til at overspredningen varierer som en funksjon av strekningslengde, antall år og trafikkmengde (Geedipally, Lord & Park, 2009). Ved bruk av en variabel overspredningsparameter vil problemene som er knyttet til en fast overspredningsparameter i mindre grad oppstå (Hauer, 2001; Lord & Park, 2008; Miaou & Lord, 2003).

De fleste prediktorvariablene i modellene er **dummyvariabler** (dvs. at variabelen har en verdi på enten null eller én). For fartsgrenser er det for eksempel definert syv dummyvariabler, én for hver fartsgrense, unntatt 80 km/t (fartsgrense 80 km/t er referansekategori og har ingen koeffisient; fartsgrensene 100 og 110 km/t er slått sammen til én variabel). Grunnen til at mange variabler inngår i modellene som et sett med dummyvariabler er at man på denne måten bedre kan fange opp ikke-lineære sammenhenger mellom variabelen og antall ulykker eller skadde enn om man hadde definert én tallvariabel (for eksempel fartsgrense som én variabel som kan ha verdiene 30, 40, 50 mv.).

For dummyvariabler kan koeffisientene direkte brukes til å beregne endringen av den avhengige variabelen som en funksjon av prediktorvariabelen. Det relative antall ulykker på veger med fartsgrense X i forhold til veger med referansefartsgrensen (80 km/t) beregnes som:

$$\text{Rel. ant. ulykker} = e^{\text{Koeff.}}$$

Koeff. er koeffisienten for fartsgrense X. Eksempelvis har fartsgrense 50 km/t en koeffisient på 0,128 for Psu. Veger med fartsgrense 50 km/t har følgelig  $e^{0,128} = 1,138$  ganger så mange Psu som veger med fartsgrense 80 km/t (hvis alt annet er likt).

**Trafikkmengden** inngår med  $\text{Ln}(\text{ÅDT})$  som prediktor i modellen. Det relative antall ulykker på en veg med trafikkmengde X i forhold til en veg med trafikkmengde Y kan man beregne som en funksjon av trafikkmengdene X og Y, samt koeffisienten for  $\text{Ln}(\text{ÅDT})$ :

$$\text{Rel. ant. ulykker} = \frac{e^{\text{Ln}(X) * \text{Koeff}}}{e^{\text{Ln}(Y) * \text{Koeff}}}$$

De fleste øvrige variablene er definert som **tallvariabler** og inngår som  $\text{Ln}(\text{antall}/\text{km} + 1)$  i modellene. Dette gjelder variablene kryss, rundkjøringer og ramper. Den naturlige logaritmen brukes for å unngå at det predikerte antall ulykker stiger eksponentielt som en funksjon av prediktorvariabelen. Antall/km +1 brukes for å unngå å ta logaritmen av null. Det relative antall ulykker på en veg med X enheter (for eksempel kryss) i forhold til en veg med Y enheter kan beregnes som en funksjon av X, Y og koeffisienten:

$$\text{Rel. ant. ulykker} = \frac{e^{\text{Ln}(X/\text{km} + 1) * \text{Koeff.}}}{e^{\text{Ln}(Y/\text{km} + 1) * \text{Koeff.}}}$$

## 2.2 Ulykker og skadde

Følgende variabler brukes som avhengige variabler i modellberegningene: Antall Psu, antall LS, antall HS (meget alvorlig eller alvorlig skadde), antall drepte og antall D/HS. Det beregnes egne modeller for hver av disse variablene. Tabell 2.2.1 viser deskriptiv statistikk for alle fem variablene. De fleste segmentene har ingen Psu, skadde eller drepte.

Tabell 2.2.1: Deskriptiv statistikk for antall Psu, LS, HS, drepte og D/HS (N=76046).

	Psu	LS	HS	Drepte	D/HS
<b>Alle år (2010-2015)</b>					
<b>Sum (totalt antall)</b>	18314	22399	2552	658	3210
<b>Gjennomsnitt</b>	0,241	0,295	0,034	0,009	0,042
<b>Standardavvik</b>	0,745	1,034	0,221	0,104	0,259
<b>Min.</b>	0	0	0	0	0
<b>Maks.</b>	27	33	6	5	9
<b>Antall segmenter med null</b>	64134	65181	73958	75460	73528
<b>Andel segmenter med null</b>	84 %	86 %	97 %	99 %	97 %
<b>Per år</b>					
<b>Sum (antall per år)</b>	3347	4092	456	118	574
<b>Gjennomsnitt</b>	0,044	0,054	0,006	0,002	0,008
<b>Standardavvik</b>	0,175	0,222	0,042	0,021	0,049
<b>Min.</b>	0,0	0,0	0,0	0,0	0,0
<b>Maks.</b>	28,1	28,1	3,0	2,4	3,0
<b>Antall per mill. kjøretøy-km</b>	0,106	0,129	0,014	0,004	0,018
<b>Endring per mill. kjøretøy-km<sup>1</sup></b>	-29 %	-34 %	-24 %	-25 %	-25 %

<sup>1</sup> Endring i forhold til antall per mill. kjøretøykilometer i Høye (2014).

## 2.3 Segmentlengde og antall år

Produktet av segmentlengde og antall år inngår i alle modellene som eksponeringsvariabel. Prediktorvariabelen i modellene er den naturlige logaritmen av produktet av segmentlengde og antall år og koeffisienten for denne variabelen settes i alle modellene lik 1 (i statistikkprogrammet som ble brukt til modellberegningene, Stata, er det kun mulig å definere én eksponeringsvariabel, derfor er produktet brukt istedenfor hver variabel for seg).

Hadde man estimert koeffisienter for segmentlengde og antall år sammen med koeffisientene for de øvrige prediktorvariablene, hadde antall predikert ulykker / skader ikke nødvendigvis variert proporsjonalt med segmentlengde og antall år. Dette forutsetter at segmentlengde og antall år i seg selv er relevante prediktorvariabler.

Segmentlengden kan ifølge Hauer (2001) være korrelert med bl.a. fart og avkjørselstetthet. Hvis man estimerer koeffisienter for segmentlengde (uten at disse settes lik 1), mister man imidlertid muligheten for å generalisere resultatene fra modellberegningene på vegstrekninger som er lengre eller kortere enn segmentene i ulykkesdatafilen. Koeffisientene vil også være avhengige av hvordan vegnettet er delt opp i segmenter og vil følgelig kunne endre seg ved en annen oppdeling. Det er liten grunn til å tro at segmentlengden i seg selv er en relevant prediktor (Elvik, 2012).

Antall år kan i det foreliggende datamaterialet ha sammenheng med vegegenskaper og med andre faktorer som påvirker ulykkesrisikoen. Ved å bruke antall år som en prediktor og ved å ikke sette koeffisienten lik 1, vil man i en viss grad kunne kontrollere for en slik sammenheng mellom periodelengde (antall år) og antall ulykker / skader. Det har imidlertid vist seg at man trolig «overkontrollerer» trenden, dvs. at virkningen av en del variabler som har endret seg over tid fanges opp av prediktoren antall år istedenfor av de aktuelle variablene; dette viser seg bl.a. i usannsynlig store koeffisienter for Ln(antall år) hvis disse ikke er satt lik 1.

**Deskriptiv statistikk** for segmentlengde og antall år er vist i tabell 2.3.1. Datafilen som brukes til utviklingen av ulykkesmodellene inneholder 76 046 segmenter (rader) og data for til sammen 50 149 km veg. Den opprinnelige datafilen (med armer, ramper, rundkjøringer, tunneler, segmenter med  $\text{ÅDT} = 1$  og lignende) inneholder data for 53 056 km veg. Dette betyr at 2 907 km veg, eller 5,5% av veglengden i den opprinnelige datafilen, er blitt «renset bort» (i Høye, 2014, var andelen som ble renset bort 6,3% av veglengden). Det som er blitt renset bort er segmenter som armer, ramper, rundkjøringer, tunneler og segmenter med  $\text{ÅDT}=0$  eller antall år=0 eller manglende verdier på antall år,  $\text{ÅDT}$ , lengde, fartsgrense eller antall kjørefelt.

Tabell 2.3.1: Deskriptiv statistikk for segmentlengde og periodelengde (uvektet).

	Segmentlengde (m)	Periodelengde (antall år)
Antall segmenter	76 046	76 046
Sum	50 149 498	
Gjennomsnitt	659	5,693
Standardavvik	369	1,036
Min.	1	0,044
Maks.	1 199	6

Den vektete gjennomsnittlige periodelengden er 5,55 år. Dette betyr at segmentene i gjennomsnitt omfatter trafikkmengde fra årene 2010 til juli 2015. Midten av denne perioden er i 2012 og resultatene gjelder derfor i gjennomsnitt året 2012.

## 2.4 Trafikkmengde

I de forrige modellene var både  $\text{Ln}(\text{ÅDT})$  og  $\text{Ln}(\text{ÅDT})^2$  prediktorvariabler. Med  $\text{Ln}(\text{ÅDT})^2$  i tillegg til  $\text{Ln}(\text{ÅDT})$ , som i alle modellene hadde en negativ koeffisient i modellene til Høye (2014), ble modellene noe bedre og predikerte antall ulykker/skader bedre enn modellene med kun  $\text{Ln}(\text{ÅDT})$ . En ulempe (i den praktiske anvendelsen av modellresultatene) med to prediktorvariabler for  $\text{ÅDT}$  er at den prosentvise økningen av det normale antall ulykker per prosent økning av trafikkmengden endrer seg med trafikkmengden, noe den ikke gjør hvis bare  $\text{Ln}(\text{ÅDT})$  er prediktor. I de aktuelle modellene er kun  $\text{Ln}(\text{ÅDT})$  benyttet som prediktor da  $\text{Ln}(\text{ÅDT})^2$  ikke har vist seg å forbedre modellene vesentlig (se nedenfor i dette avsnittet).

I forbindelse med utviklingen av ulykkesmodellene av Høye (2014) er det gjort mange forsøk med alternative trafikkmengde-variabler, hvorav ingen har ført til en vesentlig forbedring av modellene i (bl.a. dummyvariabler for høy/middels/lav  $\text{ÅDT}$ , flere  $\text{ÅDT}$ -variabler for høy, middels og lav  $\text{ÅDT}$ ,  $1/\text{ÅDT}$  og interaksjonseffekter mellom  $\text{Ln}(\text{ÅDT})$  og type veg). Av de alternative trafikkmengde-prediktorene ga interaksjonsvariablene  $\text{Ln}(\text{ÅDT}) * \text{type veg}$  de mest lovende resultater. Her var det beregnet én variabel for hver type veg (vegene var delt inn i enten fem eller 11 ulike typer) som er lik  $\text{Ln}(\text{ÅDT})$  på segmenter på de respektive typene veg, og null ellers. Det viste seg at modellene predikerte antall  $\text{Psu}$  og  $\text{LS}$  noe bedre enn modellene med  $\text{Ln}(\text{ÅDT})$  og  $\text{Ln}(\text{ÅDT})^2$ . Forskjellene var imidlertid verken store eller konsistente mellom ulike skadegrader og modellvarianter. Ulempene med interaksjonsvariablene er at koeffisientene for trafikkmengde blir tilsynelatende ulogiske og vanskelige å tolke og at modellene trolig er mindre robuste ved endringer av trafikkmengden (resultatene gjelder kun ved den fordelingen av trafikkmengden på hver av vegtypene / -klassene som i det datasettet som ble brukt til å utvikle modellene). Det er derfor ikke gjort nye forsøk med slike interaksjonsvariabler.

**Deskriptiv statistikk** for trafikkmengden er vist i tabell 2.4.1. Tabellen viser både «snitt» ÅDT og «siste» ÅDT. Snitt-ÅDT er den som brukes i modellberegningene. Dette er for hvert segment den gjennomsnittlige trafikkmengden i den perioden som inngår i datafilen. Siste ÅDT er for hvert segment den mest aktuelle ÅDT-verdien som foreligger (dvs. som regel fra 2015). Gjennomsnittlig uvektet ÅDT er 2466, mens den er 1787 når man tar hensyn til segmentlengden. Forskjellen skyldes at de lengste segmentene har lavest ÅDT.

Tabell 2.4.1: Deskriptiv statistikk for ÅDT.

	ÅDT (siste)	ÅDT (snitt)	ÅDT (snitt; Høye, 2014)
<b>Antall</b>	76 046	76 046	73 170
<b>Gjennomsnitt (uvektet)</b>	2 466	2 393	2 690
<b>Standardavvik</b>	5 705	5 555	5 726
<b>Gjennomsnitt (vektet med segmentlengde)</b>	1 787	1 733	1 643
<b>Min</b>	5	5	1
<b>Maks</b>	99 894	98 043	92 957

Tabell 2.4.2 viser hyppighetsfordelingen for ÅDT. Tabellen viser at over halvparten (59,0%) av segmentene har en ÅDT på under 1000 og at 94,6% av segmentene har en ÅDT på under 10000. Det samlede årlige antall kjøretøykilometer er 31714 millioner (det var 30799 i Høye, 2014). Dette gjelder datamaterialet som er brukt i modellberegningene. Det totale antall kjøretøykilometer på alle vegene i Norge (inkludert kommunale vegger, rundkjøringer, ramper mv.) er anslagsvis 45,417 millioner i 2016.

Tabell 2.4.2: Hyppighetsfordeling av ÅDT med samlet segmentlengde.

ÅDT	Antall	Andel	Kum. andel	Gj.sn. lengde (m)	Samlet lengde (km)	Andel av veglengden	Mill. kjt-km per år	Andel
< 100	6 168	8,1 %	8,1 %	792	4 884	9,7 %	136	0,4 %
100-499	25 785	33,9 %	42,0 %	749	19 318	38,5 %	1 906	6,0 %
500-999	12 904	17,0 %	59,0 %	682	8 804	17,6 %	2 326	7,3 %
1 000-1 999	11 137	14,6 %	73,6 %	634	7 057	14,1 %	3 664	11,6 %
2 000-2 999	6 133	8,1 %	81,7 %	601	3 686	7,3 %	3 288	10,4 %
3 000-3 999	3 084	4,1 %	85,8 %	535	1 651	3,3 %	2 099	6,6 %
4 000-4 999	2 023	2,7 %	88,4 %	502	1 016	2,0 %	1 668	5,3 %
5 000-9 999	4 719	6,2 %	94,6 %	464	2 188	4,4 %	5 590	17,6 %
10 000-19 999	2 977	3,9 %	98,5 %	377	1 123	2,2 %	5 498	17,3 %
20 000-39 999	791	1,0 %	99,6 %	390	309	0,6 %	3 122	9,8 %
40 000-59 999	189	0,2 %	99,8 %	365	69	0,1 %	1 172	3,7 %
60 000-79 999	90	0,1 %	99,9 %	361	32	0,1 %	823	2,6 %
80 000-	46	0,1 %	100,0 %	286	13	0,0 %	420	1,3 %
<b>Total</b>	<b>76 046</b>			<b>518</b>	<b>50 149</b>		<b>31 714</b>	

**Test-modellberegninger - kun  $\ln(\text{ÅDT})$  eller  $\ln(\text{ÅDT})^2$  i tillegg?** Tabell 2.4.3 viser resultatene fra test-modellberegningen for to alternativer med ulike trafikkmengdeprediktorer:

- **Alternativ A:** Både  $\ln(\text{ÅDT})$  og  $\ln(\text{ÅDT})^2$  som prediktorer.
- **Alternativ B:** Kun  $\ln(\text{ÅDT})$  som prediktor.

Modellene inneholder alle prediktorvariablene som beskrevet i 3.1 (endelige modeller) med den eneste forskjellen at alternativ A inneholder  $\ln(\text{ÅDT})^2$  i tillegg til  $\ln(\text{ÅDT})$ .



Tabellen viser Pseudo-R<sup>2</sup> og MSPE (mean square prediction error; jf. avsnitt 3.2) som indikatorer for hvor gode modellene er. Pseudo-R<sup>2</sup> viser hvor stor andel av variansen i de registrerte antallene ulykker/skadde som forklares av modellen og bedre modeller har større verdier. MSPE viser det gjennomsnittlige kvadrerte avviket mellom predikerte og bedre modeller har følgelig mindre verdier. I tillegg viser tabellen prosentvise avvik mellom registrerte og predikerte ulykkestall, både for totale ulykkestall (alle år) og for årlige ulykkestall, og koeffisientene for trafikkmengdeprediktorene.

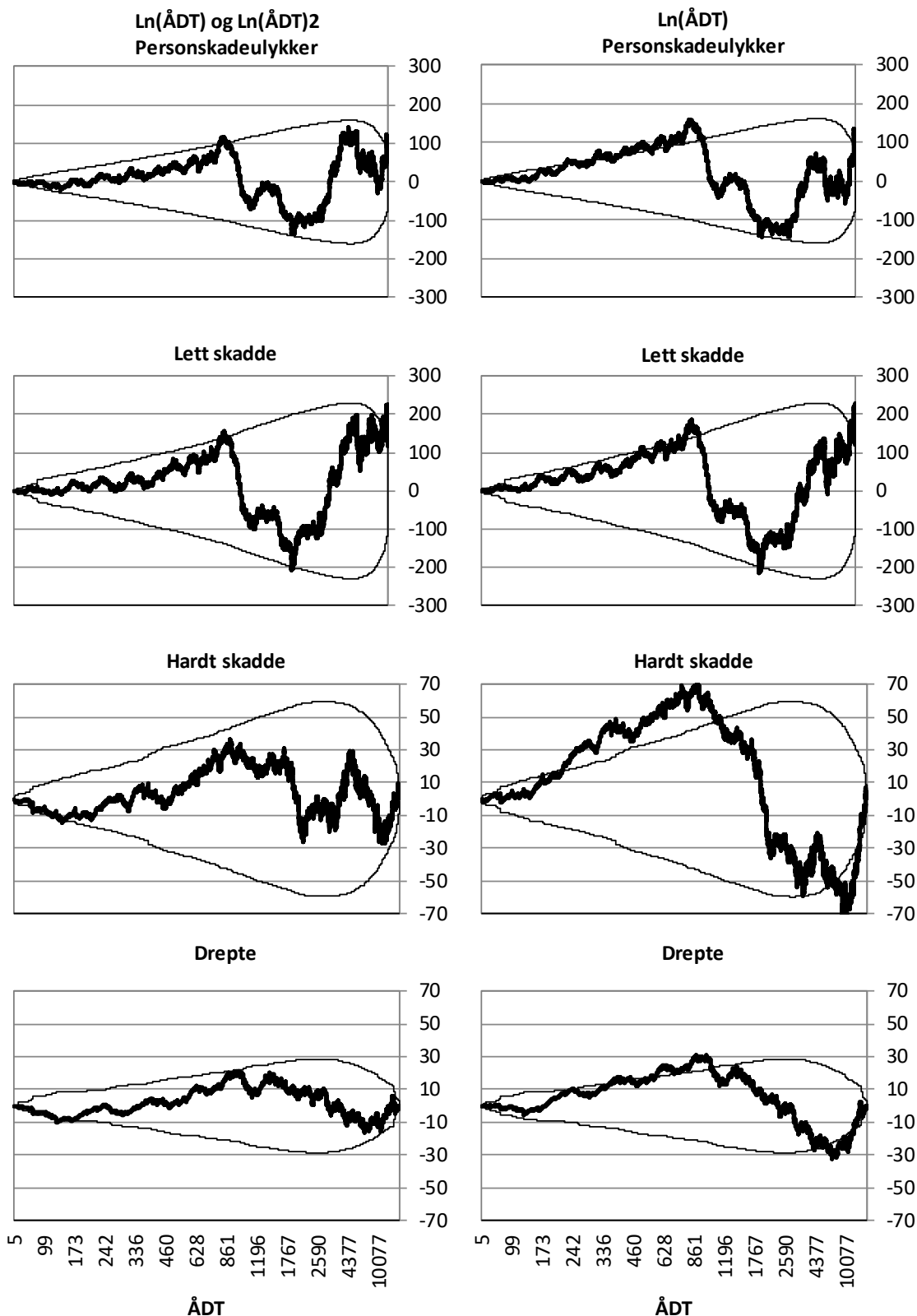
Tabell 2.4.3: Modell-fit og koeffisienter i modeller med ulike prediktorer for trafikkmengde (modell B i fet skrift er den som er brukt i de aktuelle modellene).

	Psu	LS	HS	Drepte	D/HS
	Koeff.	Koeff.	Koeff.	Koeff.	Koeff.
<b>Pseudo-R<sup>2</sup></b>					
A	0,1903	0,1397	0,0731	0,0642	0,0718
B	<b>0,1902</b>	<b>0,1396</b>	<b>0,0718</b>	<b>0,0626</b>	<b>0,0704</b>
Forrige	0,200	0,144	0,075	0,074	0,073
<b>MSPE</b>					
A	0,3374	0,6919	0,0463	0,0106	0,0630
B	<b>0,3383</b>	<b>0,6923</b>	<b>0,0464</b>	<b>0,0106</b>	<b>0,0631</b>
Forrige	0,544	1,286	0,062	0,015	0,086
<b>Avvik normale minus registrerte ulykkestall (alle år)</b>					
A	1,39 %	1,82 %	0,69 %	0,31 %	0,55 %
B	<b>1,54 %</b>	<b>1,86 %</b>	<b>0,75 %</b>	<b>0,38 %</b>	<b>0,64 %</b>
Forrige	1,00 %	2,45 %	0,66 %	0,28 %	0,17 %
<b>Avvik normale minus registrerte ulykkestall (per år)</b>					
A	1,57 %	2,40 %	2,65 %	0,91 %	2,23 %
B	<b>1,66 %</b>	<b>2,40 %</b>	<b>2,51 %</b>	<b>0,86 %</b>	<b>2,15 %</b>
Forrige	2,51 %	3,80 %	2,75 %	4,11 %	2,50 %
<b>Koeffisienter</b>					
A: Ln(ÅDT)	1,125	1,083	1,883	2,086	1,905
A: Ln(ÅDT) <sup>2</sup>	-0,013	-0,008	-0,071	-0,088	-0,073
B: Ln(ÅDT)	<b>0,929</b>	<b>0,962</b>	<b>0,843</b>	<b>0,812</b>	<b>0,838</b>
Forrige: Ln(ÅDT)	1,230	1,314	1,928	1,925	1,937
Forrige: Ln(ÅDT) <sup>2</sup>	-0,016	-0,018	-0,074	-0,070	-0,074

Som vist i tabell 2.4.3 har *modellene A med Ln(ÅDT) og Ln(ÅDT)<sup>2</sup>*, sammenlignet med modellene med kun Ln(ÅDT) som trafikkmengdeprediktor:

- Litt høyere Pseudo-R<sup>2</sup> og litt lavere MSPE, noe som tyder på at modellene med Ln(ÅDT)<sup>2</sup> er bedre, men forskjellene er små.
- Noe mindre avvik mellom summen av predikerte og registrerte totale ulykkestall («alle år»), men forskjellene er små.
- Dels mindre og dels større avvik mellom summen av predikerte og registrerte totale årlige ulykkestall («per år»), men forskjellene er små.

Figur 2.4.1 viser Cure-plott for modellene med Ln(ÅDT) og Ln(ÅDT)<sup>2</sup> (venstre) og med kun Ln(ÅDT) (høyre) som trafikkmengdeprediktorer. Cure-plott er nærmere forklart i avsnitt 3.2. Kort sagt viser plottene de summerte avvikene mellom registrerte og predikerte ulykkestall på alle segmentene som er sortert etter trafikkmengden. Disse avvikene bør ligge tettst mulig inntil null-linjen og helst innenfor de to smale linjene. Avvikene begynner alltid på null ved minste ÅDT og bør helst også slutte ved omtrent null ved høyeste ÅDT.



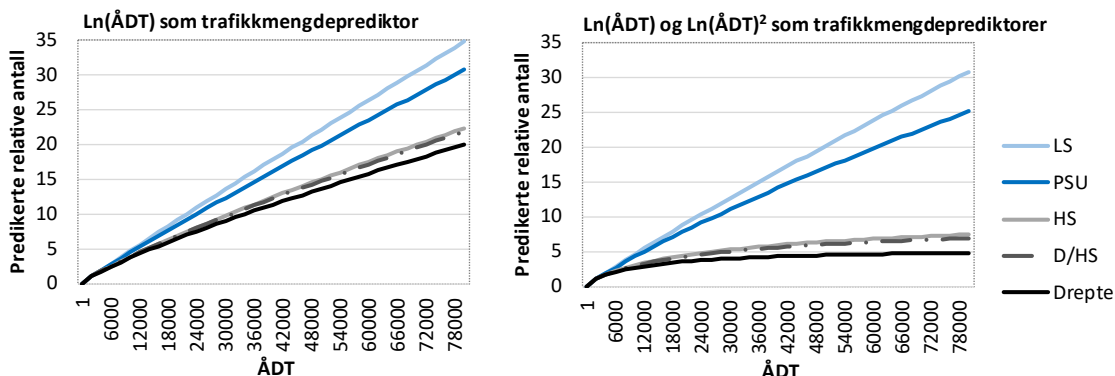
Figur 2.4.1: Cure plots for  $P_{su}$ , LS, HS og drepte, modeller med  $\ln(\text{ÅDT})$  og  $\ln(\text{ÅDT})^2$  (venstre) og med kun  $\ln(\text{ÅDT})$  (høyre) som trafikkmengdeprediktorer.

Som figur 2.4.1 viser er svingningene i Cure-plottene noe mindre og det er mindre andeler av avvikene som ligger utenfor de smale linjene i modellene med  $\text{Ln}(\text{ÅDT})^2$  som trafikkmengdeprediktor. Forskjellen er størst for HS. Forløpene viser imidlertid for alle modellene relativt store svingninger og selv om noen ligger utenfor de smale linjene i modellene med kun  $\text{Ln}(\text{ÅDT})$  som trafikkmengdeprediktor er forskjellene forholdsvis små og trolig av liten praktisk betydning.

Sammenligner man de aktuelle modellene med de forrige modellene (Høye, 2013) viser det seg at:

- Koeffisientene i de forrige modellene og alternativ A er relativt like.
- Pseudo- $R^2$  er noe mindre enn i de forrige modellene, noe som tyder på at de forrige modellene var noe bedre enn de aktuelle modellene. På den andre siden er MSPE også mindre i de aktuelle modellene, noe som tyder på at disse er bedre enn de forrige.

Sammenhengen mellom trafikkmengde og predikerte antall ulykker, skadde og drepte er vist i figur 2.4.2. for modellene med kun  $\text{Ln}(\text{ÅDT})$  som trafikkmengdeprediktor (venstre) og modellene med  $\text{Ln}(\text{ÅDT})$  og  $\text{Ln}(\text{ÅDT})^2$  som trafikkmengdeprediktorer (høyre).



Figur 2.4.2: Sammenhengen mellom trafikkmengde og den predikerte økningen av relative ulykkestall (relative antall ved  $\text{ÅDT}$  2000 er satt lik én) for modellene med kun  $\text{Ln}(\text{ÅDT})$  som trafikkmengdeprediktor (venstre) og modellene med  $\text{Ln}(\text{ÅDT})$  og  $\text{Ln}(\text{ÅDT})^2$  som trafikkmengdeprediktorer (høyre).

Som figur 2.4.2 viser er den predikerte økningen av ulykkestallene med økende  $\text{ÅDT}$  større i modellene med kun  $\text{Ln}(\text{ÅDT})$  som trafikkmengdeprediktor enn i modellene med både  $\text{Ln}(\text{ÅDT})$  og  $\text{Ln}(\text{ÅDT})^2$  som trafikkmengdeprediktorer, især for HS, drepte og D/HS.

Den prosentvise økningen av antall ulykker, skadde og drepte per prosent økning av trafikkmengden i modellene med kun  $\text{Ln}(\text{ÅDT})$  som trafikkmengdeprediktor er:

- 0,928% for Psu
- 0,962% for LS
- 0,841% for HS
- 0,811% for drepte
- 0,836% for D/HS.

I modellene med  $\text{Ln}(\text{ÅDT})$  og  $\text{Ln}(\text{ÅDT})^2$  som trafikkmengdeprediktorer endrer den prosentvise økningen av predikerte ulykkestall seg med økende  $\text{ÅDT}$  (økningen blir svakere med økende  $\text{ÅDT}$ ).

## 2.5 Fartsgrense

Fartsgrensen inngår i modellene med et sett dummyvariabler som prediktorvariabler. For hver fartsgrense er det definert en dummyvariabel som er 1 ved den respektive fartsgrensen og 0 ellers. Fartsgrensen 80 er referansekategorien (dvs. at ingen koeffisient blir beregnet for fartsgrense 80 km/t). I modellene for D, HS og D/HS er fartsgrensene 90 og 100 slått sammen til én dummyvariabel. Det er så få HS og drepte ved disse fartsgrensene at separate dummyvariabler ikke gir meningsfulle resultater.

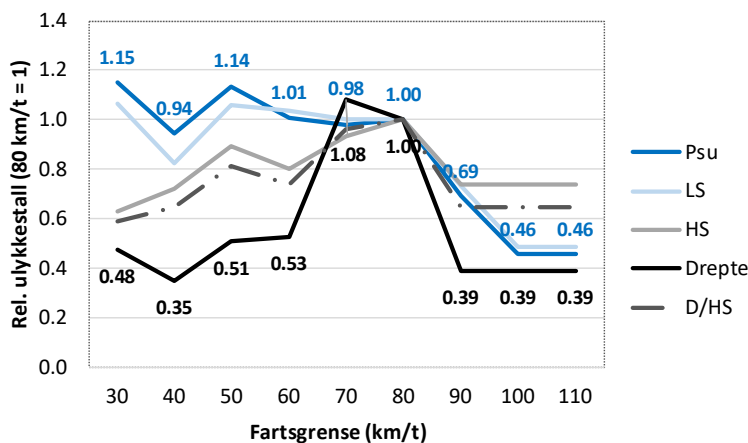
Dummyvariabler er valgt framfor en tallvariabel fordi det kan være generelle forskjeller mellom veger med ulike fartsgrenser som gjør at man ikke kan forvente at antall ulykker endrer seg monotont (eller til og med lineært) med fartsgrensen. Bl.a. er veger med ulike fartsgrenser ofte i ulike trafikkmiljøer. F.eks. er veger med fartsgrense 50 km/t eller lavere oftere enn andre veger i tettbygde strøk og har mer fotgjengertrafikk. En annen forklaring er at fartsgrensen ofte er satt ned på veger med spesielt mange ulykker (Strathman et al., 2001).

**Deskriptiv statistikk** for fartsgrensen er vist i tabell 2.5.1.

Tabell 2.5.1: Hyppighetsfordeling av fartsgrense, med lengde, gjennomsnitts-ÅDT (vektet med segmentlengde) og antall kjøretøykilometer.

Fartsgrense	Antall	Andel	Samlet lengde (km)	Andel av hele veglengde	ÅDT	Mill. kjøt.-km per år	Andel av alle kjøt.-km
30 km/t	825	1,1 %	276	0,5 %	2 038	205	1 %
40 km/t	2 549	3,4 %	864	1,7 %	3 177	1 002	3 %
50 km/t	10 205	13,4 %	4 524	9,0 %	2 352	3 883	12 %
60 km/t	14 671	19,3 %	8 060	16,1 %	1 941	5 711	18 %
70 km/t	3 405	4,5 %	1 849	3,7 %	5 728	3 866	12 %
80 km/t	42 745	56,2 %	33 458	66,7 %	1 055	12 882	41 %
90 km/t	1 109	1,5 %	847	1,7 %	5 084	1 571	5 %
100 km/t	397	0,5 %	205	0,4 %	25 273	1 888	6 %
110 km/t	140	0,2 %	68	0,1 %	28 409	706	2 %
<b>Total</b>	<b>76 046</b>		<b>50 149</b>		<b>1 733</b>	<b>31 714</b>	

Sammenhengen mellom fartsgrense og predikerte relative ulykkestall i de endelige modellene er vist i figur 2.5.1. Figuren viser at sammenhengen ikke er lineær, især for drepte og HS.



Figur 2.5.1: Sammenhengen mellom fartsgrense og relative ulykkestall (relative antall = 1 ved 80 km/t).

Sammenhengen mellom fartsgrense og relative ulykkestall har endret seg for HS, drepte og D/HS i forhold til de forrige modellene. I de forrige modellene var det en tendens til økende relative ulykkestall med synkende fartsgrense, mens det i de aktuelle modellene er en omvendt U-formet sammenheng mellom fartsgrense og relative antall HS, drepte og D/HS. For Psu og LS er sammenhengen relativt uendret.

## 2.6 Antall kjørefelt

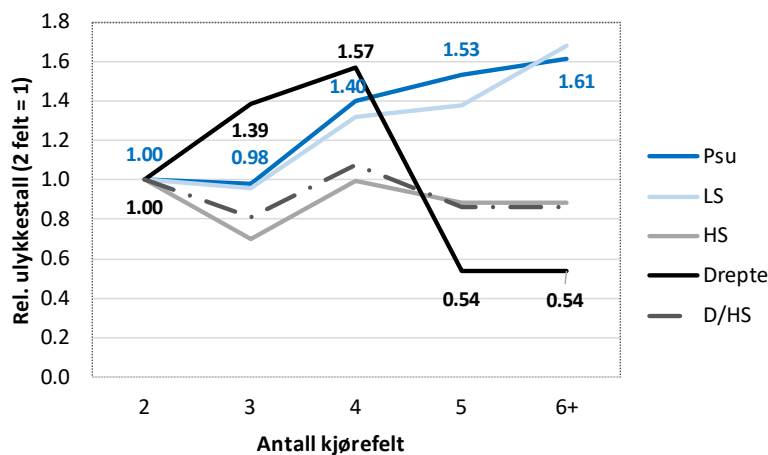
Prediktorvariabler for antall kjørefelt er definert på samme måte som for fartsgrense. For hvert antall kjørefelt er det definert en dummyvariabel. Segmenter med kun ett kjørefelt er utelatt fra datasettet fordi dette er en liten og svært heterogen gruppe av helt ulike typer veier og dermed ikke noen meningsfylt kategori. I modellene for Psu og LS er alle segmenter med seks eller flere kjørefelt slått sammen i én dummyvariabel og i modellene for HS og drepte er alle segmenter med fem eller flere kjørefelt slått sammen i én dummyvariabel.

*Deskriptiv statistikk* for antall kjørefelt er vist i tabell 2.6.1.

Tabell 2.6.1: Hyppighetsfordeling av antall kjørefelt, med lengde, gjennomsnitts-ÅDT (veket med segmentlengde) og antall kjøretøykilometer.

Antall kjørefelt	Antall	Andel	Samlet lengde (m)	Andel av veglengden	ÅDT	Mill. kjøretøy-km per år	Andel av alle kjtkm
2	73 810	97,1 %	49 340	98,4 %	1 386	24967,78	78,7 %
3	603	0,8 %	189	0,4 %	10 255	708,15	2,2 %
4	1 381	1,8 %	557	1,1 %	24 020	4881,05	15,4 %
5	110	0,1 %	25	0,0 %	40 030	365,86	1,2 %
6+	142	0,2 %	38	0,1 %	56 573	791,07	2,5 %
<b>Total</b>	<b>76 046</b>		<b>50 149</b>		<b>1 733</b>	<b>31 714</b>	

Sammenhengen mellom antall kjørefelt og predikerte relative ulykkestall i de endelige modellene er vist i figur 2.6.1. Figuren viser at det predikerte antall Psu og LS øker med økende antall kjørefelt (hvis alt annet er likt, inklusive trafikkmengde) og at sammenhengen ikke er lineær. Det predikerte antall drepte øker opp til fire kjørefelt, men er betydelig lavere på veier med fem eller flere kjørefelt. Mellom antall hardt skadde og antall kjørefelt er det kun en svak sammenheng. Sammenhengen mellom antall D/HS og antall kjørefelt er i hovedsak påvirket av HS og kun i liten grad av drepte. For Psu, LS og drepte er sammenhengen med antall kjørefelt omtrent som i Høyre (2014), mens antall HS i Høyre (2014) økte med økende antall kjørefelt.



Figur 2.6.1: Sammenhengen mellom antall kjørefelt og det predikerte relative ulykkestall (relative antall = 1 ved to kjørefelt).

## 2.7 Kryss, rundkjøringer og ramper

Prediktorvariabler for X-kryss, T-kryss, rundkjøringer, av-ramper, på-ramper og uspesifiserte ramper er den naturlige logaritmen av antallet kryss/rundkjøringer/ramper per kilometer pluss én (pluss én for å unngå å ta logaritmen av null), dvs. f.eks. for X-kryss:

$$\ln((\text{Antall X-kryss per km}) + 1)$$

Variabler som er basert på den naturlige logaritmen har vist seg å gi mer logiske resultater enn andre typer variabler (f.eks. uttransformert eller kvadrert antall). Med den naturlige logaritmen unngår man at antall predikerte ulykker øker eksponentielt med økende antall kryss.

I de forrige modellene var  $\ln(\text{antall} + 1)$  prediktorvariabler (uten «per km»). Ulempen med å bruke absolutte antall (istedenfor antall per km) er at antall ulykker kan bli grovt overestimert. F.eks. medfører dette at en 10 km lang strekning med 10 kryss har et langt høyere normalt antall ulykker enn 10 strekninger på 1 km med ett kryss per strekning. Dermed er modellene vanskelige å bruke for datasett med en annen fordeling av strekningslengder enn i datasettet som ble brukt til modellutviklingen. Med antall per kilometer unngår man dette problemet. Den eneste ulempen kan være at korte strekninger vil kunne ha absurd mange kryss per kilometer. På lengre strekninger vil imidlertid resultatene bare i liten grad bli påvirket av slike korte strekninger og normale ulykkestall vil kunne summeres over flere kortere strekninger til det samme antallet som på én lengre strekning.

**Plankryss (X- og T-kryss):** X-kryss er firearmede plankryss og T-kryss er trearmede plankryss. En del av T-kryssene i datamaterialet kan være avkjørsler. Det skilles ikke mellom ulike typer kryssregulering. Kryss er i datamaterialet alltid lagt til hovedvegen i krysset. Det betyr at sidevegene i datamaterialet ikke har noen kryss (ulykkene er som regel uansett stedfestet på hovedvegen).

Den originale datafilen inneholdt en del plankryss på motorveger og på tofeltsveg med planskilte kryss. Slike kryss er som regel ramper (ofte i forbindelse med f.eks. rasteplasser). Kryssene er derfor endret slik:

- T-kryss på motorveger og på tofeltsveg med planskilte kryss: Antallet er satt lik null og for hvert slikt T-kryss er det lagt til en «Uspesifisert rampe»
- X-kryss på motorveger og på tofeltsveg med planskilte kryss: Antallet er satt lik null og for hvert slikt X-kryss er det lagt til én avrampe og én pårampe.

**Rundkjøringer:** Rundkjøringer omfatter alle typer rundkjøringer. Rundkjøringene inngår ikke i datamaterialet, men det er gjort en opptelling av antall rundkjøringer som ligger på hvert segment. På samme måte som for plankryss er rundkjøringene alltid lagt til hovedvegen. Ulykkene som skjer i rundkjøringer er også lagt til segmentet på hovedvegen.

**Ramper:** Av- og påramper er ramper til og fra (hoved-)veger med planskilte kryss. Ramper finnes både på motorveger/tofeltsveger med planskilte kryss og på andre typer veger. Uspesifiserte ramper har i den opprinnelige datafilen vært definert som T-kryss. Som for plankryss og rundkjøringer beskriver rampevariablene antall ramper som ligger i tilknytning til hvert segment, selve rampen inngår ikke i datamaterialet.

Sammenhengen av plankryss, rundkjøringer og ramper med antall ulykker kan være forskjellig avhengig av bl.a. fartsgrense, vegtype, om vegen har skille mellom kjøreretningene eller ikke og kurver og stigninger. Slike interaksjonseffekter er imidlertid ikke tatt hensyn til i modellene.

**Deskriptiv statistikk** er vist i tabell 2.7.1 for plankryss og rundkjøringer og i tabell 2.7.2 for planskilte kryss (ramper), samt fordelingen av kryssene, rundkjøringene og rampene på de ulike vegtypene (vegtypene er definert i neste avsnitt). Som beskrevet ovenfor er X- og T-kryss på motorveger og tofeltsveger med planskilte kryss omdefinert som ramper.

Tabell 2.7.1: Deskriptiv statistikk og hyppighetsfordeling for plankryss og rundkjøringer.

	X-kryss			T-kryss			Rundkjøringer		
	Antall	Andel	Per km	Antall	Andel	Per km	Antall	Andel	Per km
Alle veger	1409		0,028	31375		0,626	1235		0,025
Motorveg	0	0 %	-	0	0 %	-	1	0 %	0,003
Tofelts-planskilt	0	0 %	-	0	0 %	-	3	0 %	0,009
Øvrig TEN-T	46	3 %	0,013	1950	6 %	0,546	135	11 %	0,038
Øvrig Ev/Rv	127	9 %	0,029	2910	9 %	0,656	210	17 %	0,047
Fylkesveg	1236	88 %	0,030	26515	85 %	0,640	886	72 %	1236

Tabell 2.7.2: Deskriptiv statistikk og hyppighetsfordeling for toplanekryss (ramper).

	Av-ramper			På-ramper			Uspes. ramper		
	Antall	Andel	Per km	Antall	Andel	Per km	Antall	Andel	Per km
Alle veger	689		0,014	636		0,013	136		0,003
Motorveg	207	30 %	0,554	205	32 %	0,549	66	49 %	0,177
Tofelts-planskilt	150	22 %	0,450	145	23 %	0,435	70	51 %	0,210
Øvrig TEN-T	139	20 %	0,039	121	19 %	0,034		0 %	-
Øvrig Ev/Rv	144	21 %	0,032	124	19 %	0,028		0 %	-
Fylkesveg	49	7 %	0,001	41	6 %	0,001		0 %	-

**Test-modellberegninger:** Tabell 2.7.3 viser resultatene fra test-modellberegningen for to alternativer med ulike prediktorer for ramper:

- **Alternativ A:** Tre prediktorvariabler for ramper (av-, på- og uspesifiserte ramper).
- **Alternativ B:** De tre rampetyperne er slått sammen til «ramper».

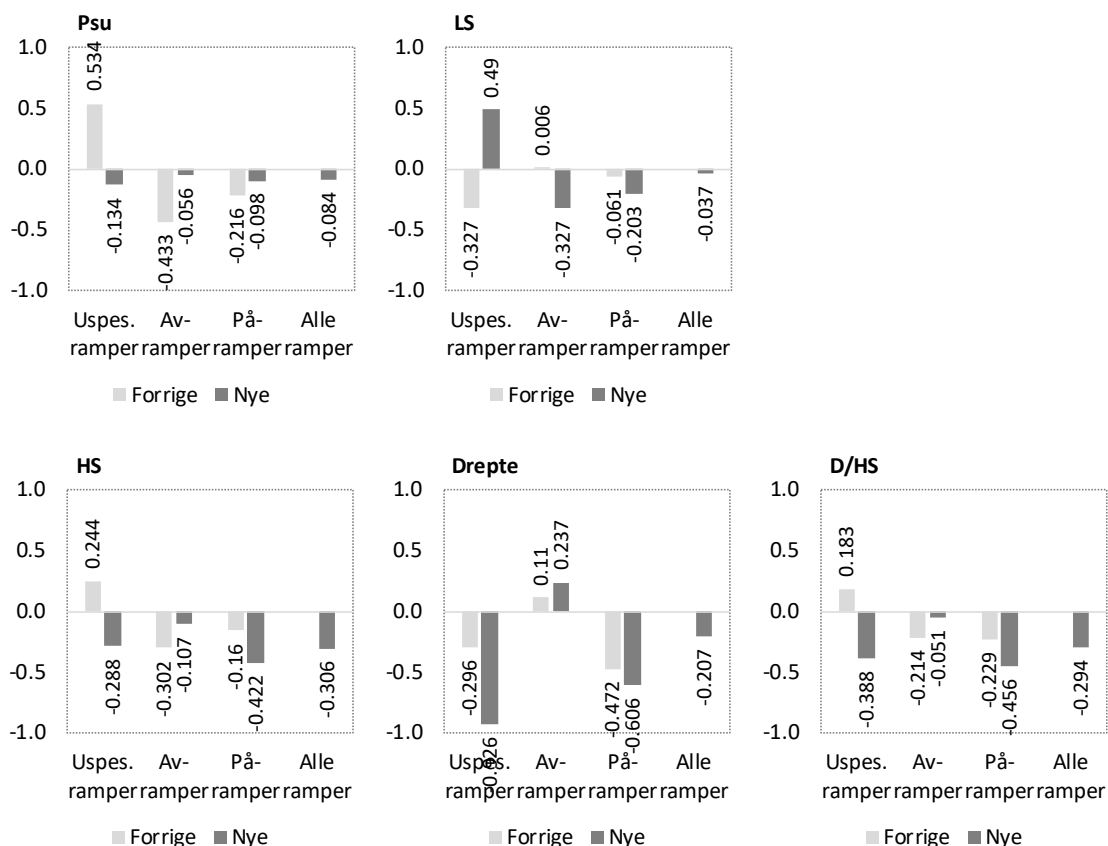
I tillegg viser tabellen Pseudo-R<sup>2</sup> og MSPE (jf. avsnitt 2.4 og 3.2) som indikatorer for hvor gode modellene er (Pseudo-R<sup>2</sup>: bedre modeller har større verdier; MSPE: bedre modeller har lavere verdier). Øvrige prediktorer i modellen er som i de endelige modellene (unntak: fire prediktorer for ATK, jf. avsnitt 2.11 og en prediktor for broer, jf. avsnitt 2.13).

Tabell 2.7.3: Modell-fit og koeffisienter i modeller med ulike prediktorer for ramper (test-beregninger, koeffisientene kan avvike noe fra de endelige modellene).

	Psu		LS		HS		Drepte		D/HS	
	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p
<b>Alternativ A: Tre rampetyper (av, på, uspesifisert)</b>										
Pseudo-R <sup>2</sup>	0,1902		0,1396		0,0718		0,0626		0,0704	
MSPE	0,3383		0,6923		0,0464		0,0106		0,0631	
Av-ramper	<b>-0,056</b>	0,300	<b>0,006</b>	0,906	<b>-0,107</b>	0,519	<b>0,237</b>	0,470	<b>-0,051</b>	0,738
På-ramper	<b>-0,098</b>	0,098	<b>-0,061</b>	0,288	<b>-0,422</b>	0,027	<b>-0,606</b>	0,141	<b>-0,456</b>	0,011
Uspes. ramper	<b>-0,134</b>	0,314	<b>-0,327</b>	0,017	<b>-0,288</b>	0,476	<b>-0,926</b>	0,422	<b>-0,388</b>	0,319
Motorveg	<b>-0,749</b>	0,000	<b>-0,680</b>	0,000	<b>-0,689</b>	0,010	<b>-1,176</b>	0,057	<b>-0,726</b>	0,004
Tofelts planskilt	<b>-0,721</b>	0,000	<b>-0,661</b>	0,000	<b>-0,828</b>	0,000	<b>0,044</b>	0,900	<b>-0,592</b>	0,003
<b>Alternativ B: Alle ramper slått sammen</b>										
Pseudo-R <sup>2</sup>	0,1880		0,1384		0,0716		0,0621		0,0702	
MSPE	0,3386		0,6919		0,0464		0,0106		0,0631	
Ramper	<b>-0,084</b>	0,029	<b>-0,037</b>	0,331	<b>-0,306</b>	0,008	<b>-0,207</b>	0,359	<b>-0,294</b>	0,006
Motorveg	<b>-0,857</b>	0,000	<b>-0,782</b>	0,000	<b>-0,719</b>	0,007	<b>-1,191</b>	0,053	<b>-0,755</b>	0,003
Tofelts planskilt	<b>-0,770</b>	0,000	<b>-0,724</b>	0,000	<b>-0,845</b>	0,000	<b>0,011</b>	0,975	<b>-0,614</b>	0,002
<b>Forrige modellene (Høye 2014)</b>										
Pseudo-R <sup>2</sup>	0,200		0,144		0,075		0,074		0,073	
MSPE	0,544		1,286		0,062		0,015		0,086	
Uspes. ramper	<b>0,534</b>	0,000	<b>0,490</b>	0,000	<b>0,244</b>	0,482	<b>-0,296</b>	0,765	<b>0,183</b>	0,573
Av-ramper	<b>-0,433</b>	0,000	<b>-0,327</b>	0,000	<b>-0,302</b>	0,204	<b>0,110</b>	0,808	<b>-0,214</b>	0,319
På-ramper	<b>-0,216</b>	0,007	<b>-0,203</b>	0,014	<b>-0,160</b>	0,510	<b>-0,472</b>	0,338	<b>-0,229</b>	0,302

Koeffisientene for de ulike rampetyperne i de forrige modellene (Høye, 2014) og i de aktuelle modellene (A og B) er vist i figur 2.7.1.





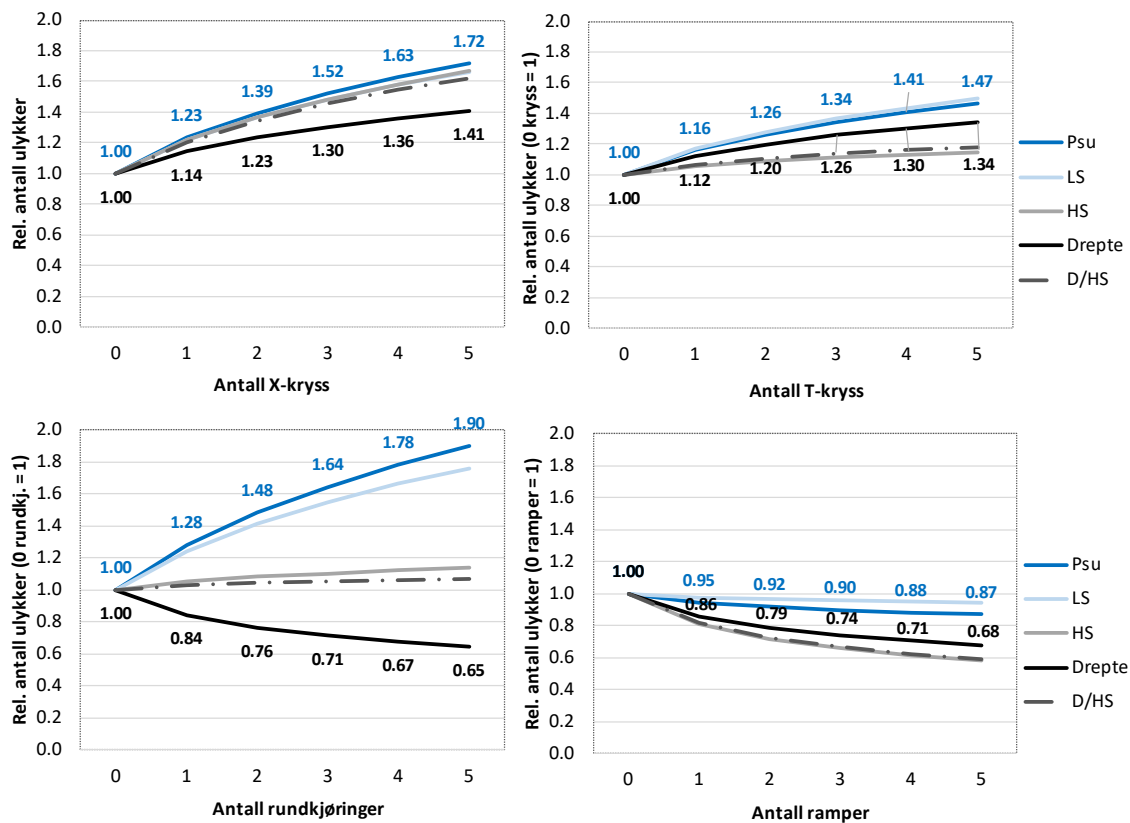
Figur 2.7.1: Koeffisientene for de ulike rampetyperne i de forrige modellene (Høye, 2014) og i de aktuelle modellene (A: Uspes., av- og på-ramper; B: Alle ramper).

Tabell 2.7.3 og figur 2.7.1 viser følgende:

- **A vs. B:** Pseudo- $R^2$  og MSPE er nesten identiske mellom A og B på andre eller tredje desimal. Forskjellene mellom koeffisientene for av- og på-rampene er ikke konsistente mellom skadegradene. Koeffisientene for motorveg og tofeltsveg med planskilte kryss er også nesten like i A og B. En relativt stor andel av koeffisientene er ikke statistisk signifikante. Motorveger og tofeltsveger med planskilte kryss har noe gunstigere effekter i B enn i A.
- **A/B vs. forrige modell:** Forskjellene mellom koeffisientene for av- og på-rampene mellom de aktuelle og de forrige modellene er usystematiske.

Siden forskjellene mellom de enkelte rampetyperne er lite konsistente, både mellom skadegradene og mellom de aktuelle og de forrige modellene, konkluderes at de ulike rampetyperne slås sammen i de endelige modellene.

Sammenhengen mellom antall plankryss, rundkjøringer, ramper og predikerte relative ulykkestall i de endelige modellene er vist i figur 2.7.2.



Figur 2.7.2: Sammenhengen mellom antall plankryss, rundkjøringer og ramper og predikerte relative ulykkestall.

Figur 2.7.2 viser at de predikerte ulykkestall øker med økende antall X- og T-kryss og synker med økende antall ramper. Økende antall rundkjøringer medfører flere Psu og LS, en liten økning av antall HS og færre drepte. Den prosentvise endringen av ulykkestallene med økende antall kryss/rundkjøringer/ramper

- Blir mindre for hver ytterligere kryss/rundkjøring/rampe
- Er uavhengig av segmentlengden.

Sistnevnte betyr at f.eks. en dobling av antall kryss medfører samme endring i antall ulykker på en 10 km lang veg som på en 1 km lang veg. Predikerte ulykkestall kan også summeres for flere korte vegsegmenter (f.eks. 10 segmenter på 1 km lengde hver) slik at summen er den samme som det predikerte antallet på alle segmentene til sammen (ett segment på 10 km).

## 2.8 Type veg

Det er definert fem dummyvariabler for ulike typer veg som er de samme som i de forrige modellene:

1. *Motorveg*
2. *Tofeltsveg med planskilte kryss* (motortrafikkveg i tidligere terminologi)
3. *TEN-T-veg for øvrig* (ikke motorveg eller tofeltsveg med planskilte kryss)
4. *Europa-/riksveg (Ev/Rv) for øvrig* (ikke motorveg eller tofeltsveg med planskilte kryss eller TEN-T veg)
5. *Fylkesveg* (referansekategori)

Disse vegtypene er valgt fordi det finnes generelle forskjeller i vegstandard, bl.a. mht. kjørefelt- og skulderbredde. Eksempelvis har alle motorveger midtrekkverk og TEN-T veger har generelt en noe høyere standard enn øvrige Europa- og riksveger.

Tofeltsveger med planskilte kryss er i hovedsak tofeltsveger, men 29% av segmentene og 25% av veglengden på slike veger har flere enn to felt, derav de aller fleste tre eller fire felt.

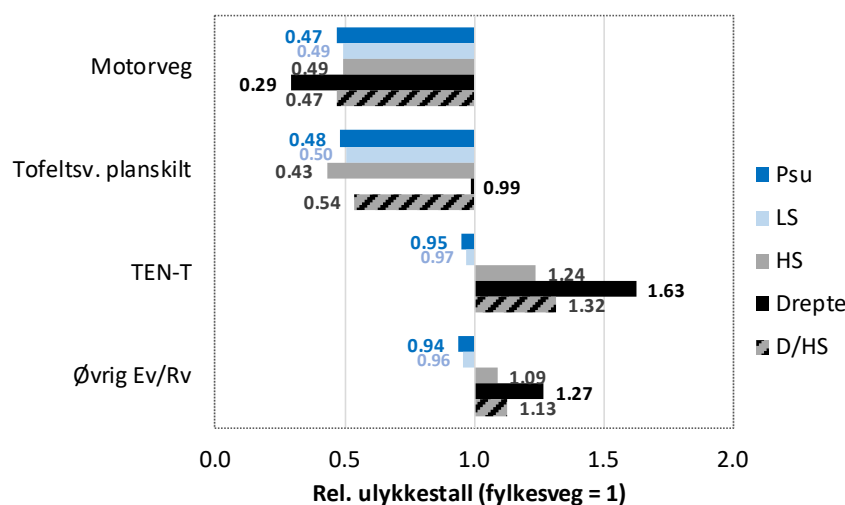
**Deskriptiv statistikk** for type veg er vist i tabell 2.8.1.

Tabell 2.8.1: Type veg, sammenheng med segmentlengde og ÅDT (vektet med segmentlengde).

	Lengde			Andel av veglengden	ÅDT		Mill. kjt.km per år	
	N	Gjennom- snitt (m)	Sum (km)		Gjennom- snitt	N	%*	
<b>1 Motorveg</b>	807	463	374	0,7 %	30 012	4 094	12,9 %	
<b>2 Tofeltsv. planskilt</b>	677	492	333	0,7 %	12 198	1 483	4,7 %	
<b>3 TEN-T</b>	5 542	645	3 573	7,1 %	3 839	5 007	15,8 %	
<b>4 Øvrig Ev/Rv</b>	6 998	634	4 438	8,8 %	3 377	5 471	17,2 %	
<b>5 Fylkesveg</b>	62 022	668	41 431	82,6 %	1 036	15 659	49,4 %	
<b>Alle</b>		<b>659</b>	<b>50 149</b>		<b>1 733</b>	<b>31 714</b>		

\* Andel av alle kjt-km

Sammenhengen mellom type veg og relative ulykkestall i de endelige modellene er vist i figur 2.8.1.



Figur 2.8.1: Sammenhengen mellom vegtype og predikerte ulykkestall (relative antall på fylkesveger = 1).

Figur 2.8.1 viser at motorveger har betydelig færre ulykker enn andre typer veg. Det samme gjelder for tofeltsveger med planskilte kryss, unntatt at disse har omtrent like mange drepte som fylkesveger. Dette gjelder selv om det er kontrollert for om vegene har midtdeler eller midtrekkverk (disse er egne prediktorvariabler i modellene). Øvrige TEN-T og Ev/Rv har noen færre Psu og LS, og flere HS, drepte og D/HS enn andre veger. Resultatene er omtrent like som i Høye (2014), unntatt at tofeltsveger med planskilte kryss i Høye (2014) hadde flere drepte enn fylkesveger.

## 2.9 Midtdeler og midtrekkverk

For midtdeler og midtrekkverk er det definert fire dummyvariabler som er prediktorvariabler i modellene:

- Verken midtdeler eller midtrekkverk (referansekategori)
- Midtdeler, ikke midtrekkverk
- Midtrekkverk, ikke midtdeler
- Midtdeler og -rekkverk

Det er definert fire dummyvariabler istedenfor én for med/uten midtdeler og én for med/uten midtrekkverk fordi det kan være interaksjonseffekter mellom midtdeler og midtrekkverk, dvs. at virkningen av midtrekkverk kan være forskjellig på veger med og uten midtdeler og at virkningen av midtdeler kan være forskjellig på veger med og uten midtrekkverk.

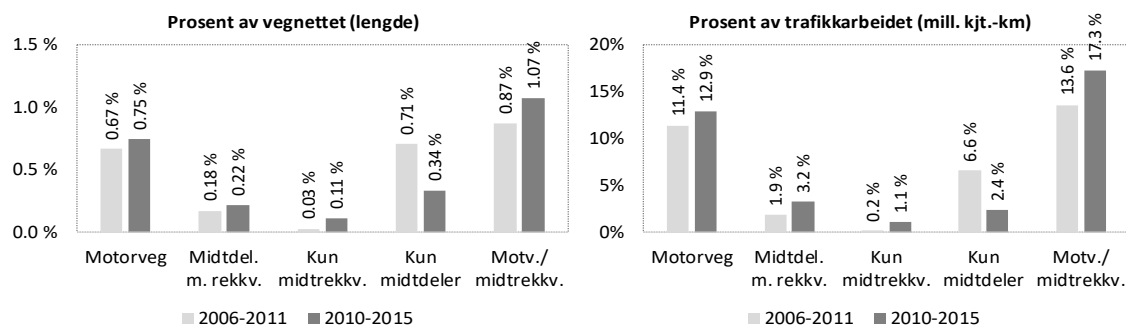
På motorvegene har alle fire variablene verdien null. Resultatene er nøyaktig de samme som om alle motorvegene hadde hatt verdien én på variabelen midtdeler.

**Deskriptiv statistikk** for midtdeler og midtrekkverk er vist i tabell 2.9.1. Tallene gjelder kun veger som ikke er motorveger.

Tabell 2.9.1: Midtdeler og midtrekkverk, deskriptiv statistikk og sammenheng med ÅDT (uten motorveger).

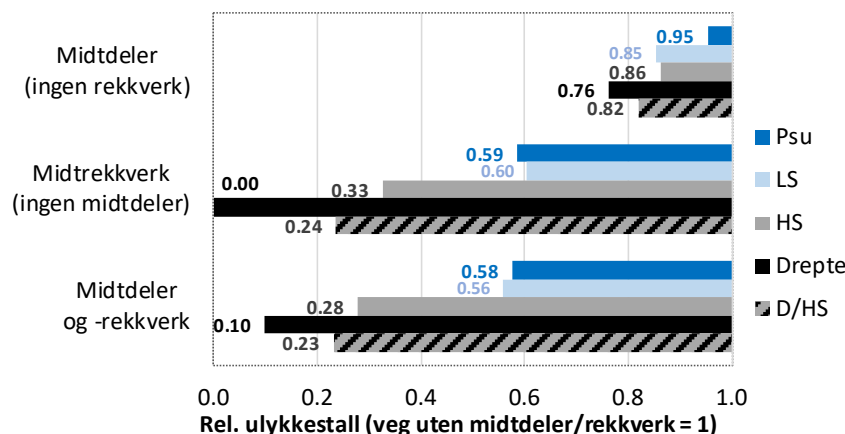
	N	Lengde			ÅDT Gjennom- snitt	Mill. kjt-km per år	
		Gjennom- snitt (m)	Sum (km)	Andel av veg lengden		N	%
<b>Uten midtdeler/-rekkv.</b>	74 097	667	49 443	99,33 %	1 411	25 464	92,19 %
<b>Kun midtdeler</b>	717	236	169	0,34 %	12 422	766	2,77 %
<b>Kun midtrekkverk</b>	142	386	55	0,11 %	18 095	362	1,31 %
<b>Midtdeler m. rekkverk</b>	283	386	109	0,22 %	25 790	1 029	3,72 %
<b>Alle</b>	<b>75 239</b>	<b>662</b>	<b>49 776</b>		<b>1 520</b>	<b>27 620</b>	

Figur 2.9.1 viser fordelingen av veglengden og av trafikkarbeidet på motorveger og andre veger med midtdeler og/eller midtrekkverk i den aktuelle analysen (2010-2015) og i Høyre (2014; år 2006-2011). Figuren viser at det har vært en økning av både veglengden og trafikkarbeid på motorveger og veger med midtrekkverk (både med og uten midtdeler). Relativt sett har økningen vært størst for veger med midtrekkverk som ikke er motorveger (+257% av veglengden og +362% av trafikkarbeidet på veger med kun midtrekkverk, unntatt motorveger; +59% av veglengden og +100% av trafikkarbeidet på veger med midtrekkverk, inkl. veger med midtrekkverk og midtdeler, unntatt motorveger).



Figur 2.9.1: Fordelingen av veglengden og av trafikkarbeidet på motorveger og andre veger med midtdeler og/eller midtrekkverk.

Figur 2.9.2 viser predikerte relative ulykkestall på veger med og uten midtdeler / midtrekkverk i de endelige modellene.



Figur 2.9.2: Sammenhengen mellom midtdeler, midtrekkverk og predikerte relative ulykkestall (relative antall på veger uten midtdeler eller midtrekkverk = 1).

Figur 2.9.2. viser at det på veger med midtrekkverk (med eller uten midtdeler) er langt færre skadde og drepte enn på veger uten midtdeler og -rekkverk. På veger med kun midtdeler er det også færre ulykker enn på veger uten midtdeler og -rekkverk, men forskjellen er mindre enn for midtrekkverk. I Høye (2014) har det i motsetning til resultatene i figur 2.9.2 vært noen flere skadde og især drepte på veger med midtdeler enn på veger uten midtdeler og -rekkverk.

Resultatene i figur 2.9.2 gjelder kun veger som ikke er motorveger. Virkningen av midtdeler og -rekkverk på motorveg er i sin helhet fanget opp av vegtype.

## 2.10 Forsterket midtoppmerking

Forsterket midtoppmerking er kombinasjonen av vanlig oppmerking og rumleriller. I den opprinnelige datafilen og i Høye (2014) er det skilt mellom to typer forsterket midtoppmerking:

- Bred forsterket midtoppmerking (over 1 m bred)
- Smal forsterket midtoppmerking (opptil 1 m bred)

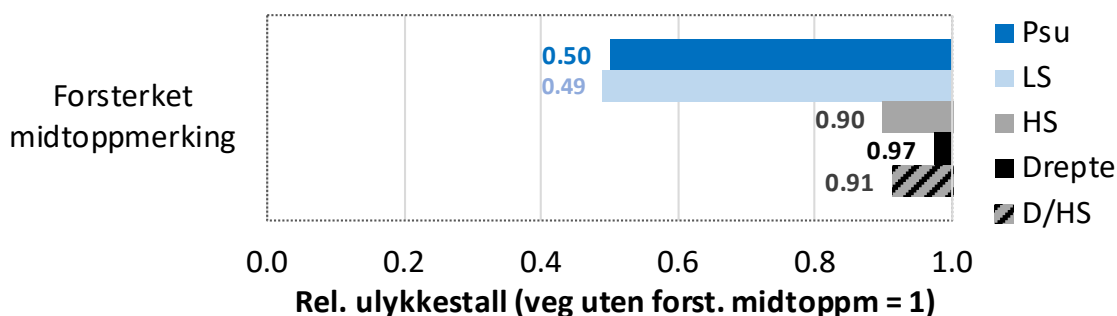
I Høye (2014) var det 111 km veg med bred og 141 km veg med smal forsterket midtoppmerking. I den aktuelle datafilen er det kun registrert 22 km veg med bred forsterket midtoppmerking og 291 km veg med smal forsterket midtoppmerking. Det er ukjent hva den tilsynelatende nedgangen av bred forsterket midtoppmerking skyldes. Siden 22 km (fordelt på 48 segmenter, vs. 522 segmenter med smal forsterket midtoppmerking) er for lite for å kunne få meningsfulle resultater er bred og smal forsterket midtoppmerking i de aktuelle modellberegningene slått sammen til én dummyvariabel (med vs. uten forsterket midtoppmerking). På motorveger og andre veger med midtdeler og/eller midtrekkverk er variabelen satt lik null (dette handler om 49 segmenter eller 7,9% av alle segmentene med forsterket midtoppmerking).

**Deskriptiv statistikk** for forsterket midtoppmerking er vist i tabell 2.10.1. Tabellen viser resultatene for variablene som er brukt i modellberegningene (forsterket midtoppmerking satt lik null på motorveger og veger med midtdeler og/eller -rekkverk).

Tabell 2.10.1: Forsterket midtoppmerking, deskriptiv statistikk og sammenheng med ÅDT (vektet med segmentlengde).

Forst. midt- Oppmerking	N	Lengde			ÅDT Gjennom- snitt	Mill. kjt-km per år	
		Gjennom- snitt (m)	Sum (km)	Andel av veglengden		N	%
Bred	48	454	22	0,0 %	12 599	100	0,3 %
Smal	522	557	291	0,6 %	6 435	683	2,2 %
Bred/smal	570	549	313	0,6 %	6 865	783	2,5 %
Uten	75 476	660	49 837	99,4 %	1 700	30 931	97,5 %
<b>Alle</b>	<b>76 046</b>	<b>659</b>	<b>50 149</b>		<b>1 733</b>	<b>31 714</b>	

Figur 2.10.1 viser predikerte relative ulykkestall på veger med og uten forsterket midtoppmerking i de endelige modellene.



Figur 2.10.1: Sammenhengen mellom forsterket midtoppmerking og predikerte relative ulykkestall (relative antall på veger uten forsterket midtoppmerking = 1).

Figur 2.10.1 viser at forsterket midtoppmerking medfører færre Psu og LS og noen færre HS, drepte og D/HS. I Høye (2014) ble det funnet noen større effekter for bred (>1 meter) enn for smal forsterket midtoppmerking, men forskjellene var ikke store. Virkningen var også mindre for de mer alvorlige skadegradene, men det var mindre forskjeller mellom skadegradene enn i resultatene i figur 2.10.1.

## 2.11 Automatisk trafikkontroll (ATK)

Automatisk trafikkontroll (ATK) omfatter punkt-ATK (PATK) og streknings-ATK (SATK).

I NVDB er det definert influensstrekninger for både PATK og SATK, dvs. strekninger hvor man antar at trafikantene påvirkes av ATK. Influensstrekningene i datafilen for modellberegningene er hentet direkte fra NVDB. For ATK-punkter som mangler influensstrekningen i NVDB er det beregnet en teoretisk influensstrekning på 100 meter foran punktet og 3000 meter etter punktet. Noen få influensstrekninger fra NVDB er blitt forkastet fordi de ikke har noe ATK-punkt i nærheten.

I den opprinnelige datafilen er det skilt mellom influensstrekninger i én og begge retninger. Influensstrekninger i én retning er strekninger som etter definisjonen som er gitt ovenfor kun er påvirket av ATK i én retning, og tilsvarende for influensstrekninger i begge retninger.

**Deskriptiv statistikk** for P- og SATK er vist i tabell 2.11.1. At det finnes langt flere PATK-strekninger med PATK i én retning enn med PATK i begge retninger skyldes at ATK-punktene ofte er satt opp i nærheten av hverandre slik at det lite overlapp mellom influensområdene.

Tabell 2.11.1: PATK og SATK, deskriptiv statistikk og sammenheng med ÅDT (vektet med segmentlengde).

	Lengde			Andel av veglengden	ÅDT	Mill. kjt-km per år	
	N	Gjennomsnitt (m)	Sum (km)		Gjennomsnitt	N	%
Ingen ATK	74 093	665	49 280	98,3 %	1 618	29 104	91,8 %
PATK en retning	1 310	408	534	1,1 %	8 813	1 719	5,4 %
PATK begge retn.	380	480	182	0,4 %	9 116	607	1,9 %
SATK en retning	139	472	66	0,1 %	5 878	141	0,4 %
SATK begge retn.	124	700	87	0,2 %	4 554	144	0,5 %
<b>Alle</b>	<b>76 046</b>	<b>659</b>	<b>50 149</b>		<b>1 733</b>	<b>31 714</b>	

**Test-modellberegninger:** Tabell 2.11.2 viser resultatene fra test-modellberegningen for seks alternativer med ulike prediktorer for ATK:

- **Alternativ A:** Fire dummyvariabler: PATK i én retning, PATK i begge retninger, SATK i én retning, SATK i begge retninger
- **Alternativ B:** To dummyvariabler: PATK (én eller begge retninger), SATK (én eller begge retninger)
- **Alternativ C:** To dummyvariabler: PATK/SATK i én retning, PATK/SATK i begge retninger
- **Alternativ D:** Én dummyvariabel: ATK vs. ikke ATK (det skilles ikke mellom PATK og SATK eller mellom én eller begge retninger)
- **Alternativ E:** Ingen ATK-prediktorer
- **Alternativ F:** Én dummy for PATK (en eller begge retninger) og to dummier for SATK (SATK i én retning, SATK i begge retninger)

I tillegg viser tabellen Pseudo-R<sup>2</sup> og MSPE (jf. avsnitt 2.4 og 3.2) som indikatorer for hvor gode modellene er (Pseudo-R<sup>2</sup>: bedre modeller har større verdier; MSPE: bedre modeller har lavere verdier). Øvrige prediktorer i modellen er som i de endelige modellene (unntak: en prediktor for broer er med i modellene, jf. avsnitt 2.13; tre prediktorer for rampetyper, jf. avsnitt 2.7).

Tabell 2.11.2: Modell-fit og koeffisienter i modeller med ulike prediktorer for ATK (test-beregninger, koeffisientene kan avvike noe fra de endelige modellene; koeffisientene for PATK og SATK i fet skrift).

	Psu		LS		HS		Drepte		D/HS	
	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p
<b>Alternativ A: Fire dummier</b>										
Pseudo-R <sup>2</sup>	0,1902		0,1396		0,0718		0,0626		0,0704	
MSPE	0,3383		0,6923		0,0464		0,0106		0,0631	
PATK i 1 retning	<b>-0,009</b>	0,832	<b>0,000</b>	0,997	<b>-0,149</b>	0,225	<b>-0,156</b>	0,516	<b>-0,153</b>	0,180
PATK i begge retn.	<b>0,097</b>	0,132	<b>0,092</b>	0,222	<b>-0,018</b>	0,925	<b>-0,044</b>	0,897	<b>-0,013</b>	0,942
SATK i 1 retning	<b>-0,175</b>	0,403	<b>-0,159</b>	0,470	<b>0,598</b>	0,144	<b>0,456</b>	0,518	<b>0,589</b>	0,113
SATK i begge retn.	<b>-0,629</b>	0,017	<b>-0,727</b>	0,009	<b>-1,928</b>	0,061	<b>-0,871</b>	0,416	<b>-1,515</b>	0,042
<b>Alternativ B: To dummier (PATK SATK)</b>										
Pseudo-R <sup>2</sup>	0,1901		0,1396		0,0713		0,0624		0,0700	
MSPE	0,3380		0,6917		0,0464		0,0106		0,0631	
PATK alle	<b>0,020</b>	0,578	<b>0,025</b>	0,552	<b>-0,113</b>	0,287	<b>-0,121</b>	0,553	<b>-0,114</b>	0,249
SAKT alle	<b>-0,369</b>	0,023	<b>-0,389</b>	0,023	<b>-0,138</b>	0,703	<b>-0,066</b>	0,908	<b>-0,081</b>	0,800
<b>Alternativ C: To dummier (én retning begge retninger)</b>										
Pseudo-R <sup>2</sup>	0,1901		0,1395		0,0713		0,0624		0,0700	
MSPE	0,3382		0,6922		0,0464		0,0106		0,0631	
P/SATK i en retning	<b>-0,016</b>	0,688	<b>-0,009</b>	0,845	<b>-0,099</b>	0,403	<b>-0,096</b>	0,672	<b>-0,097</b>	0,376
P/SATK begge retn.	<b>0,048</b>	0,436	<b>0,030</b>	0,671	<b>-0,103</b>	0,567	<b>0,005</b>	0,988	<b>-0,059</b>	0,720
<b>Alternativ D: Én dummy (ATK)</b>										
Pseudo-R <sup>2</sup>	0,1901		0,1395		0,0713		0,0624		0,0700	
MSPE	0,3380		0,6920		0,0464		0,0106		0,0631	
ATK (alle)	<b>0,001</b>	0,970	<b>0,003</b>	0,947	<b>-0,115</b>	0,263	<b>-0,115</b>	0,552	<b>-0,111</b>	0,242
<b>Alternativ E: Ingen ATK-prediktor</b>										
Pseudo-R <sup>2</sup>	0,1901		0,1395		0,0713		0,0624		0,0699	
MSPE	0,3380		0,6920		0,0464		0,0106		0,0631	
<b>Alternativ F: Tre dummier (PATK SATK én retning SATK begge retninger)</b>										
Pseudo-R <sup>2</sup>	0,1902		0,1396		0,0717		0,0626		0,0704	
MSPE	0,3379		0,6917		0,0464		0,0106		0,0631	
PATK alle	<b>0,020</b>	0,577	<b>0,025</b>	0,549	<b>-0,112</b>	0,289	<b>-0,121</b>	0,551	<b>-0,114</b>	0,250
SATK i 1 retning	<b>-0,176</b>	0,400	<b>-0,160</b>	0,466	<b>0,597</b>	0,145	<b>0,455</b>	0,519	<b>0,587</b>	0,114
SATK i begge retn.	<b>-0,629</b>	0,017	<b>-0,728</b>	0,009	<b>-1,928</b>	0,061	<b>-0,872</b>	0,416	<b>-1,515</b>	0,042

Tabell 2.11.2 viser følgende:

- De aller fleste koeffisientene for ATK-variablene er veldig små og ikke statistisk signifikante, unntatt koeffisientene for SATK i begge retninger (for alle skadegrader unntatt drepte). Koeffisientene for SATK viser at det er færre ulykker på segmenter med SATK i begge retninger enn og flere på segmenter med SATK i én retning enn på andre segmenter.
- Det er praktisk talt ingen forskjeller mellom Pseudo-R<sup>2</sup> og MSPE mellom modellene A-F. For Psu og LS er MSPE lavest for alternativ F, fulgt av B, drepte og E. Den mest detaljerte modellen (A) predikerer dårligere enn de mindre detaljerte modellene, inkludert den uten ATK-prediktor. Pseudo-R<sup>2</sup> er høyest (på fjerde desimal) for modell A og F.

Tabell 2.11.3 viser de prosentvise avvikene mellom predikerte og registrerte ulykkes- og skadetallene på segmentene med ATK.



Tabell 2.11.3: MSPE og avvik mellom predikerte og registrerte ulykkes- og skadetallene på segmentene med ATK (fet skrift og grå bakgrunn: de to «beste» modellene per skadegrad; kursiv grå skrift: de to «dårligste» modellene per skadegrad).

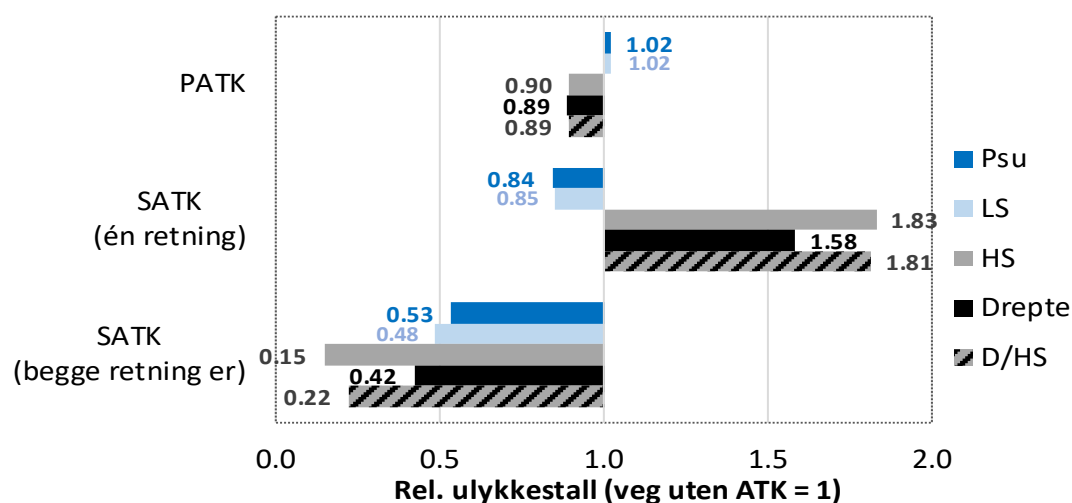
	A Fire dummier	B To dummier: PATK, SATK	C To dummier: 1 retn., 2 retn.	D Én dummy	E Ingen ATK- prediktor	F Tre dummier
<b>MSPE</b>						
<b>Psu</b>	1,1731	<b>1,1585</b>	1,1678	1,1591	1,1591	<b>1,1571</b>
<b>LS</b>	2,5353	<b>2,5097</b>	2,5293	2,5196	2,5200	<b>2,5093</b>
<b>HS</b>	<b>0,1075</b>	0,1081	0,1081	0,1081	0,1085	<b>0,1073</b>
<b>Drepte</b>	<b>0,0283</b>	0,0284	0,0284	0,0284	0,0284	<b>0,0283</b>
<b>D/HS</b>	<b>0,1589</b>	0,1594	0,1596	0,1594	0,1601	<b>0,1586</b>
<b>Avvik</b>						
<b>Psu</b>	0,26 %	<b>0,12 %</b>	<b>-0,09 %</b>	-0,20 %	-0,30 %	0,13 %
<b>LS</b>	<b>-0,65 %</b>	-0,81 %	-1,33 %	-1,31 %	-1,53 %	<b>-0,77 %</b>
<b>HS</b>	<b>0,93 %</b>	1,24 %	2,46 %	1,23 %	11,41 %	<b>0,78 %</b>
<b>Drepte</b>	<b>-0,11 %</b>	0,18 %	4,55 %	0,18 %	10,10 %	<b>-0,04 %</b>
<b>D/HS</b>	<b>1,37 %</b>	1,69 %	3,90 %	1,70 %	11,58 %	<b>1,21 %</b>

Tabell 2.11.3 viser at modell F produserer nesten gjennomgående de minste avvikene, fulgt av modell A. Det er relativt små forskjeller mellom modell A/F og B. Modellene uten ATK-prediktor og modell C er nesten gjennomgående dårligere enn andre modeller.

Basert på resultatene i tabell 2.11.2 og 2.11.3 konkluderes at modellene A og F er de «beste» modellene. Disse har de minste MSPE på segmentene med ATK, de største Pseudo-R<sup>2</sup> for hele modellen, og minst avvik i de predikerte ulykkestallene. Resultatene tyder på at det ikke er mye å vinne ved å skille mellom PATK i én vs. to retninger, men at man bør skille mellom PATK og SATK. At koeffisientene for SATK i begge retninger er betydelig større (og signifikante) enn koeffisientene for SATK i én retning tyder på at det er forskjeller mellom segmentene med SATK i én vs. begge retninger.

I praksis er en fordel med modell F framfor A at man unngår å vurdere om et segment har PATK i én eller begge retninger, noe som kan være vanskelig især på veger med mange ATK-punkter. Derfor benyttes alternativ F med tre dummivariabler (PATK, SATK i én retning, SATK i begge retninger) i de endelige modellene.

Figur 2.11.1 viser predikerte relative ulykkestall på veger med og uten PAT/SATK i de endelige modellene for alle skadegradene.



Figur 2.11.1: Sammenhengen mellom ATK og predikerte relative ulykkestall (relative antall på veger uten ATK = 1).

De relative ulykkestallene i figur 2.11.1 må tolkes som forskjeller mellom ulike veger. Tallene kan *ikke* tolkes som effekter av ATK. Dette fordi ATK som regel installeres på veger med høyere normale ulykkestall enn andre veger. Selv om ATK reduserer antall ulykker, kan ulykkestallene med ATK fortsatt være høyere enn på andre ellers sammenlignbare veger (endogenitet). Forutsetter man at ATK reduserer ulykker, betyr dette at segmenter med ATK hadde hatt enda flere ulykker enn de har, hvis de ikke hadde hatt ATK.

Figur 2.11.1 viser at segmenter med PATK har noen flere Psu og LS, samt noen færre HS, drepte og D/HS enn segmenter uten PATK. Veger med SATK i én retning har færre Psu og LS, men flere HS, drepte og D/HS enn veger uten SATK, mens veger med SATK i begge retninger har flere ulykker av alle skadegrader enn veger uten SATK.

Hvorvidt segmenter med PATK/SATK har flere eller færre ulykker enn segmenter uten ATK kan på grunn av endogenitetseffekten som beskrevet ovenfor ikke tolkes som effekt av PATK/SATK.

## 2.12 Vegbelysning

Vegbelysning har ikke vært med som prediktor i modellene til Høye (2014). Vegbelysning er definert i en dummyvariabel som indikerer om vegen er belyst eller ikke. Det er ikke skilt mellom ulike typer belysning eller om det er satt opp belysning på en eller begge sider av vegen. Lysstrekningene er beregnet ved å bruke alle eksisterende belysningsstrekninger + 50 meter i hver ende samt alle belysningspunkt + 50 meter til hver side.

**Deskriptiv statistikk** for vegbelysning er vist i tabell 2.12.1. Tabellen viser at nesten to tredjedeler av all trafikkarbeid foregår på belyst veg. På motorveger og tofeltsveger med planskilte kryss er det over 90% av trafikkarbeidet og mellom 50% og 60% på øvrige veger.

Tabell 2.12.1: Vegbelysning, deskriptiv statistikk og sammenheng med ÅDT (vektet med segmentlengde).

	N	Lengde			ÅDT Gjennomsnitt	Mill. kjt-km per år	
		Gjennomsnitt (m)	Sum (km)	Andel av veglengden		N	%
<b>Ingen vegbelysning</b>	48 901	790	38 616	77,0 %	891	12 560	39,6 %
<b>Vegbelysning (alle)</b>	27 145	425	11 533	23,0 %	4 550	19 154	60,4 %
<b>Alle</b>	<b>76 046</b>	<b>659</b>	<b>50 149</b>		<b>1 733</b>	<b>31 714</b>	
<b>Vegbelysning på</b>							
- Motorveg	705	467	329	88 %	32 160	3 864	94 %
- Tofelts-plansk.	605	506	306	92 %	12 593	1 407	95 %
- Øvrig TENT	2 329	412	960	27 %	8 342	2 924	58 %
- Øvrig Ev/Rv	2 675	398	1 064	24 %	7 503	2 914	53 %
- Fv	20 831	426	8 874	21 %	2 484	8 045	51 %

**Test-modellberegninger:** Tabell 2.12.2 viser resultatene fra test-modellberegningen for to alternativer med ulike prediktorer for vegbelysning:

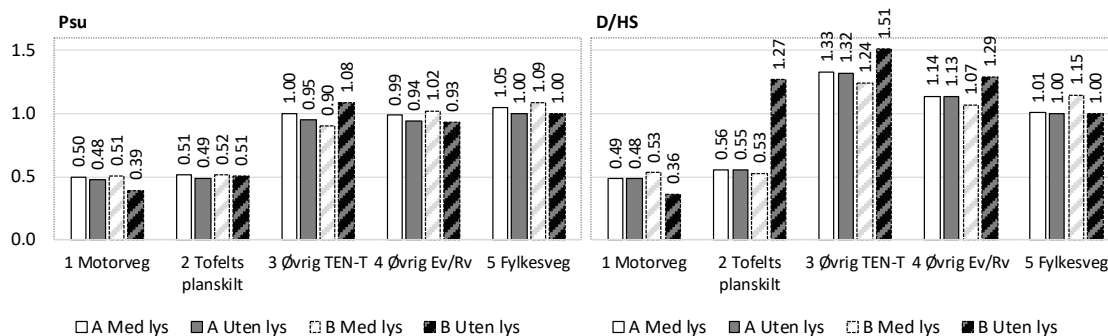
- **Alternativ A:** En dummyvariabel (vegbelysning vs. ikke vegbelysning)
- **Alternativ B:** Fem dummyvariabler for vegbelysning, én for hver vegtype (med verdien én på segmenter av den respektive vegtypen dersom disse har vegbelysning, null på alle andre segmentene)

I tillegg viser tabellen Pseudo-R<sup>2</sup> og MSPE (jf. avsnitt 2.4 og 3.2) som indikatorer for hvor gode modellene er (Pseudo-R<sup>2</sup>: bedre modeller har større verdier; MSPE: bedre modeller har lavere verdier). Øvrige prediktorer i modellen er som i de endelige modellene (unntak: tre prediktorer for rampetyper, jf. avsnitt 2.7; en prediktor for broer er med i modellene, jf. avsnitt 2.13).

Tabell 2.12.2: Modell-fit og koeffisienter i modeller med ulike prediktorer for vegbelysning (test-beregninger, koeffisientene kan avvike noe fra de endelige modellene; koeffisientene for vegbelysning i fet skrift).

	Psu		LS		HS		Drepte		D/HS	
	Koeff.	p	Koeff.	P	Koeff.	p	Koeff.	p	Koeff.	p
<b>Alternativ A: Én prediktor</b>										
Pseudo-R <sup>2</sup>	0,1902		0,1396		0,0717		0,0626		0,0704	
MSPE	0,3379		0,6917		0,0464		0,0106		0,0631	
Vegbelysning	<b>0,048</b>	0,053	<b>0,095</b>	0,001	<b>0,050</b>	0,455	<b>-0,181</b>	0,151	<b>0,005</b>	0,931
<b>Alternativ B: Vegbelysning per vegtype (fem dummier)</b>										
Pseudo-R <sup>2</sup>	0,1905		0,1399		0,0722		0,0637		0,0710	
MSPE	0,3371		0,6907		0,0464		0,0106		0,0630	
Vegbelysning på...										
... motorveg	<b>0,248</b>	0,254	<b>0,399</b>	0,061	<b>0,548</b>	0,382	<b>-0,628</b>	0,583	<b>0,379</b>	0,495
... tofelts planskilt	<b>0,019</b>	0,948	<b>0,339</b>	0,293	<b>-0,615</b>	0,368	<b>-1,373</b>	0,103	<b>-0,875</b>	0,103
... øvrig TER-N	<b>-0,179</b>	0,001	<b>-0,148</b>	0,013	<b>-0,141</b>	0,278	<b>-0,337</b>	0,132	<b>-0,196</b>	0,098
... øvrig Ev/Rv	<b>0,086</b>	0,061	<b>0,159</b>	0,003	<b>-0,103</b>	0,398	<b>-0,476</b>	0,036	<b>-0,191</b>	0,089
... Fv	<b>0,082</b>	0,003	<b>0,122</b>	0,000	<b>0,151</b>	0,053	<b>0,049</b>	0,749	<b>0,136</b>	0,059

Figur 2.12.1 viser predikerte relative antall Psu og D/HS i modellene A (én prediktor for vegbelysning) og B (én prediktor for vegbelysning per type veg) på de ulike typer veg med og uten belysning.



Figur 2.12.1: Predikerte relative antall Psu (venstre) og D/HS (høyre) i modellene A (én prediktor for vegbelysning) og B (én prediktor for vegbelysning per type veg); relative antall på ubelyst fylkesveg er satt lik én (testberegninger; resultatene kan avvike fra de endelige modellene).

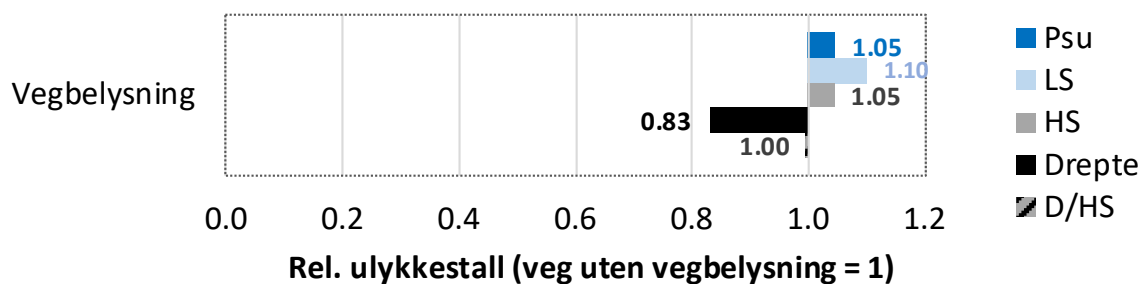
Tabell 2.12.2 og figur 2.12.1 viser at:

- Pseudo-R<sup>2</sup> er større og MSPE er mindre (eller like stor) i alternativ B, sammenlignet med A, noe som tyder på at B er «bedre». Forskjellene er imidlertid små (på tredje eller fjerde desimal).
- Mange av koeffisientene for vegbelysning er ikke signifikante, især for HS, drepte og D/HS, selv om en del av disse er forholdsvis store.

- Forskjellene mellom veger med og uten belysning er de samme for alle vegtypene i A. Her har ubelyste veger 5% færre Psu og 0,5% færre D/HS enn belyste veger. I modell B endrer forskjellene mellom belyste og ubelyste veger seg både i størrelse og fortegn. En del av endringene virker ulogiske, f.eks. at antall HS og D/HS er høyere på belyste enn på ubelyste motorveger mens antall drepte er lavere, samtidig som både antall HS, drepte og D/HS er lavere på belyste enn på ubelyste tofeltsveger med planskilte kryss.
- Figur 2.12.1 viser at de relative forskjellene i predikerte ulykkestall mellom vegtypene er relativt uendret mellom A og B. Dette tyder på at vegbelysningsvariablene ikke «spiser opp» uforholdsmessig mye av forskjellene mellom vegtypene. Dette til tross for at mesteparten av motorvegene og tofeltsvegene med planskilte kryss er belyst.

Siden modellen med én vegbelysningsprediktor per vegtype kun medfører små forbedringer av modellene framfor modellen med kun én vegbelysningsprediktor velges sistnevnte til de endelige modellene. Fordelen med modellen med kun én prediktor for vegbelysning er at modellen er enklere og at man unngår tilsynelatende ulogiske forskjeller i hvordan vegbelysning påvirker normale ulykkestall mellom vegtypene.

Figur 2.12.2 viser predikerte relative ulykkestall på veger med og uten vegbelysning. Som for ATK kan koeffisientene for vegbelysning ikke tolkes som effekt av vegbelysning. Dette fordi det kan være andre forskjeller mellom veger med og uten belysning.



Figur 2.12.2: Relative predikerte ulykkestall på veger med og uten vegbelysning (relative antall på veger uten vegbelysning = 1).

Figur 2.11.2 viser at veger med belysning har flere Psu, LS og HS, men færre drepte enn veger uten belysning. Som beskrevet ovenfor kan disse forskjellene ikke tolkes som effekter av vegbelysning.

## 2.13 Broer

Broer har ikke vært med som prediktor i modellene til Høye (2014). For alle segmentene er det definert om de ligger på en bro eller ikke. Alle broer på minst 100 meter lengde er tatt med. Kortere broer er ikke tatt med.

**Deskriptiv statistikk** for broer er vist i tabell 2.13.1. Det er kun en liten andel av veglengden som ligger på broer.

Tabell 2.13.1: Broer, deskriptiv statistikk og sammenheng med ADT (vektet med segmentlengde).

	N	Lengde			ÅDT	Mill. kjt-km per år	
		Gjennomsnitt (m)	Sum (km)	Andel av veglengden	Gjennomsnitt	N	%
Ingen bro	75 281	664	49 993	99,7 %	1 710	31 197	98,4 %
Bro (alle)	765	205	157	0,3 %	9 023	517	1,6 %
<b>Alle</b>	<b>76 046</b>	<b>659</b>	<b>50 149</b>		<b>1 733</b>	<b>31 714</b>	

**Test-modellberegninger:** Tabell 2.13.2 viser resultatene fra test-modellberegningen for to alternativer med ulike prediktorer for broer:

- **Alternativ A:** Én dummyvariabel (bro vs. ikke bro)
- **Alternativ B:** Ingen prediktor for bro.

I tillegg viser tabellen Pseudo-R<sup>2</sup> og MSPE (jf. avsnitt 2.4 og 3.2) som indikatorer for hvor gode modellene er (Pseudo-R<sup>2</sup>: bedre modeller har større verdier; MSPE: bedre modeller har lavere verdier). Øvrige prediktorer i modellen er som i de endelige modellene (unntak: tre prediktorer for ulike rampetyper, jf. avsnitt 2.7).

Tabell 2.13.2: Modell-fit og koeffisienter i modeller med ulike prediktorer for vegbehysning (test-beregninger, koeffisientene kan avvike noe fra de endelige modellene).

	Psu		LS		HS		Drepte		D/HS	
	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p
<b>Alternativ A: Én prediktor for broer</b>										
Pseudo-R <sup>2</sup>	0,1902		0,1396		0,0717		0,0626		0,0704	
MSPE	0,3379		0,6917		0,0464		0,0106		0,0631	
Bro	<b>-0,075</b>	0,413	<b>0,012</b>	0,899	<b>-0,451</b>	0,112	<b>-0,524</b>	0,398	<b>-0,478</b>	0,071
<b>Alternativ B: Ingen prediktor for broer</b>										
Pseudo-R <sup>2</sup>	0,1902		0,1396		0,0716		0,0625		0,0702	
MSPE	0,3378		0,6917		0,0464		0,0106		0,0631	

Tabell 2.13.3 viser MSPE og de prosentvise avvikene mellom predikerte og registrerte ulykkes- og skadetallene på segmentene med bro.

Tabell 2.13.3: MSPE og avvik mellom predikerte og registrerte ulykkes- og skadetallene på segmentene med bro (testberegninger; resultatene kan avvike fra de endelige modellene).

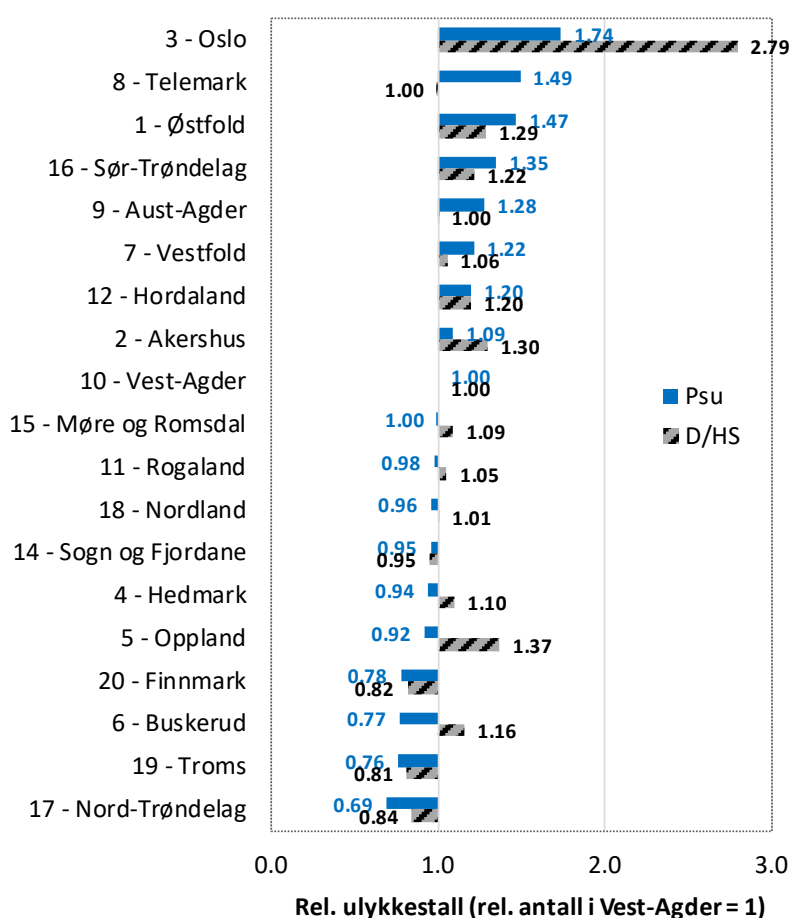
	MSPE		Avvik	
	Med bro-prediktor	Uten bro-prediktor	Med bro-prediktor	Uten bro-prediktor
Psu	0,292	0,289	0,59 %	0,51 %
LS	0,941	0,943	0,73 %	0,74 %
HS	0,037	0,038	0,21 %	-0,14 %
Drepte	0,006	0,006	-0,05 %	-0,34 %
D/HS	0,058	0,060	0,13 %	-0,23 %

Tabell 2.13.3 viser at MSPE og avvikene på segmentene med bro er dels større og dels mindre i modellene med bro-prediktor enn i modellene uten bro-prediktor. Siden prediktorvariabelen for broer ikke ser uten til å gjøre modellene bedre, droppes denne prediktoren fra de endelige modellene.

## 2.14 Fylke

For hvert fylke er det definert én dummyvariabel. Fylkesvariablene fanger opp generelle forskjeller mellom fylkene som f.eks. forskjeller i topografi, vær og befolkningstetthet. Referansekategori er fylke 10 (Vest-Agder), da dette er fylket med laveste avvik av antall Psu og antall D/HS per million kjøretøykilometer fra de gjennomsnittlige antallene Psu og D/HS per million kjøretøykilometer. I Høye (2014) var det Fylke 12 (Hordaland) som var referansekategori.

Relative predikerte ulykkestallantall i de ulike fylkene er vist i figur 2.14.1. De relative ulykkestallene i Fylke 10 (Vest-Agder) er satt lik 1 (dette er referansekategori). Fylkene er sortert etter relative antall Psu i figur 2.14.1. Relative antall LS i fylkene ligner på de relative antallene Psu og relative antall drepte og HS ligner på de relative antallene D/HS.



Figur 2.14.1: Relative ulykkestall i ulike fylker (relative antall i Vest-Agder = 1; sortert etter relativt antall Psu).

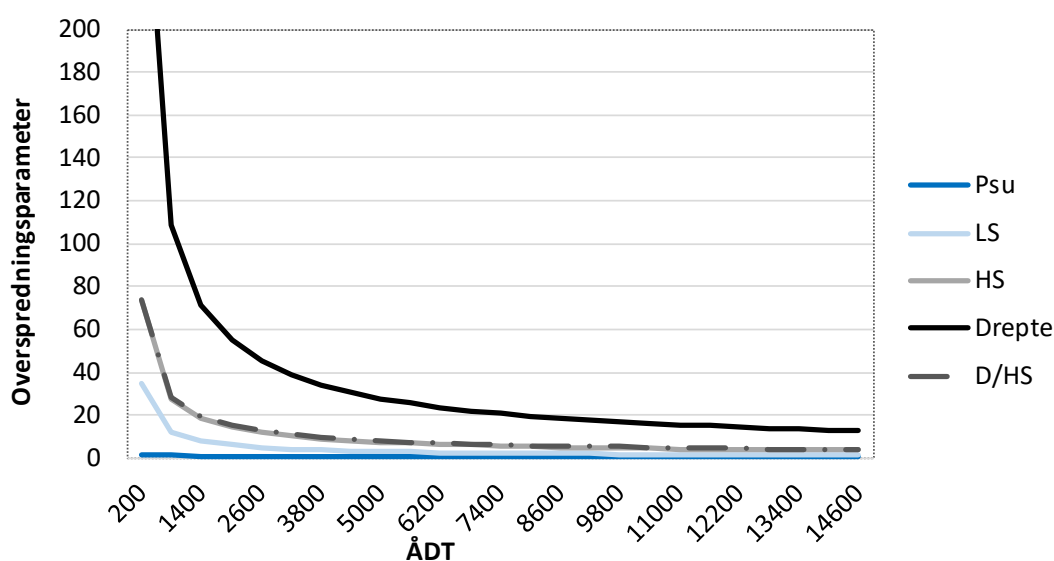
Relative antall LS i fylkene ligner på de relative antallene Psu og relative antall drepte og HS ligner på de relative antallene D/HS. Rekkefølgene fylkene imellom er omtrent lik som i Høye (2014). De største endringene er Oslo som kom seks plasser opp (flere Psu) og Finnmark som kom seks plasser ned (færre Psu). De øvrige fylkene har ikke flyttet seg med mer enn fire plasser, de fleste mindre.

## 2.15 Konstantterm

Alle modellene inneholder en konstantterm. Denne er beregnet automatisk av statistikkprogrammet.

## 2.16 Overspredningsparameter

Alle modellene inneholder koeffisienter for å beregne overspredningsparameteren som en funksjon av segmentlengde, antall år og trafikkmengde. Dermed tar man hensyn til at overspredningen synker med økende segmentlengde, antall år og trafikkmengde. Figur 2.16.1 viser overspredningsparameteren som en funksjon av trafikkmengden. Resultatene gjelder et vegsegment på 1 km lengde og ett år.



Figur 2.16.1: Overspredningsparameter som en funksjon av trafikkmengden (gjelder 1 km og ett år).

Overspredningsparameteren kan beregnes med hjelp av koeffisientene som er vist i neste avsnitt (tabell 3.1.1). For Psu beregnes overspredningsparameteren for eksempel etter følgende formel:

$$\text{Overspredningsparameter} = e^{5,920 + \ln(\text{lengde} * \text{år}) * -0,601 + \ln(\text{ÅDT}) * -0,240}$$

Sammenhengen mellom ÅDT og overspredningsparameteren som er vist i figur 2.16.1 ligner for alle skadegrader på sammenhengen som ble funnet i Høye (2014).

## 3 Modellene

### 3.1 Koeffisientene og beregningseksempel

Tabell 3.1.1 viser koeffisientene for alle prediktorvariablene i modellene for alle skadegrader. De fleste variabler er dummyvariabler. For slike variabler er alltid en av kategoriene referansekategori, dvs. at det i modellen ikke er beregnet noen koeffisient. For alle øvrige kategoriene settes variabelen lik 1 hvis et segment er i denne kategorien og null ellers (f.eks. settes variabelen «fartsgrense 70» lik én på en veg med fartsgrense 70 km/t mens alle andre fartsgrensevariablene settes lik null. Koeffisientene for dummyvariablene viser forholdet mellom antall ulykker i denne kategorien og i referansekategorien.

Variabler som ikke er dummyvariabler er ÅDT (her benyttes  $\ln(\text{ÅDT})$  som prediktor) og antall kryss, rundkjøringer og ramper (her benyttes  $\ln(\text{antall per kilometer} + 1)$  som prediktor).

Segmentlengde og antall år inngår som den naturlige logaritmen av henholdsvis antall år og antall meter med koeffisienten lik én i modellen.

For å vise hvordan modellen brukes for å beregne normale ulykkestall er det lagd et beregningseksempel. Eksempelet viser beregningen av det normale antall  $\text{Psu}$  i løpet av ett år for en vegstrekning med følgende egenskaper:

Lengde (meter):	5 000
Trafikkmengde:	10 000
Fartsgrense:	70
Antall kjørefelt:	2
Antall X-kryss:	1
Antall kurver:	2
Vegklasse:	Riksveg, ikke motorveg og ikke del av TEN-T vegnettet
Fylke:	Hedmark (fylke 4)

Ved å sette inn de respektive verdiene for de enkelte prediktorvariablene og koeffisientene i modellformelen beregnes normale antall ulykker og skadde. Modellformelen er

$$E(n) = e^{\sum_i \text{Prediktor}_i * \text{Koeff}_i}$$

Det normale antall  $\text{Psu}$  er følgende

$$PSU = e^{-16,584 + \ln(5000*1) + \ln(10000)*0,982 - 0,021 + \ln(1/5+1)*0,302 - 0,063 - 0,062} = 2,427$$

Siden vegen har to kjørefelt, som er referansekategorien for antall kjørefelt, settes det ikke inn noen koeffisient for antall kjørefelt. Antall T-kryss, rundkjøringer mv. som er null, settes heller ikke inn i formelen. Av antall X-kryss settes den naturlige logaritmen av antallet per kilometer pluss én inn i formelen.



Tabell 3.1.1: Koeffisienter i de endelige modellene for personskadenulykkes (Psu), lett skadde (LS), hardt skadde (HS), drepte og drepte eller hardt skadde (D/HS).

	Psu		LS		HS		Drepte		D/HS	
	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p
Ln(ADT)	0,928	0,000	0,962	0,000	0,841	0,000	0,811	0,000	0,836	0,000
Fgr, 30 km/t	0,140	0,094	0,062	0,567	-0,462	0,124	-0,739	0,256	-0,522	0,063
Fgr, 40 km/t	-0,058	0,257	-0,189	0,002	-0,324	0,037	-1,054	0,005	-0,438	0,003
Fgr, 50 km/t	0,128	0,000	0,060	0,120	-0,111	0,230	-0,676	0,001	-0,208	0,016
Fgr, 60 km/t	0,009	0,748	0,035	0,258	-0,223	0,002	-0,641	0,000	-0,301	0,000
Fgr, 70 km/t	-0,021	0,527	0,005	0,900	-0,069	0,430	0,080	0,590	-0,037	0,641
Fgr 80 km/t		(ref.)		(ref.)		(ref.)		(ref.)		(ref.)
Fgr 90 km/t	-0,369	0,000	-0,310	0,000						
Fgr, 100-110 km/t	-0,785	0,000	-0,713	0,000						
Fgr, 90-110 km/t					-0,299	0,127	-0,940	0,021	-0,437	0,016
Kjørefelt: 2		(ref.)		(ref.)		(ref.)		(ref.)		(ref.)
Kjørefelt: 3	-0,018	0,810	-0,041	0,620	-0,351	0,127	0,327	0,374	-0,207	0,310
Kjørefelt: 4	0,338	0,000	0,278	0,000	-0,007	0,965	0,448	0,202	0,076	0,632
Kjørefelt: 5	0,425	0,003	0,321	0,028						
Kjørefelt: 6+	0,478	0,000	0,520	0,000						
Kjørefelt: 5+					-0,126	0,650	-0,625	0,446	-0,151	0,574
X-kryss <sup>1</sup>	0,302	0,000	0,284	0,000	0,285	0,004	0,192	0,410	0,271	0,004
T-kryss <sup>1</sup>	0,214	0,000	0,224	0,000	0,077	0,064	0,165	0,053	0,093	0,016
Rundkjøringer <sup>1</sup>	0,359	0,000	0,315	0,000	0,072	0,422	-0,244	0,335	0,038	0,662
Ramper <sup>1</sup>	-0,078	0,041	-0,032	0,404	-0,302	0,009	-0,217	0,337	-0,292	0,006
Motorveg	-0,761	0,000	-0,706	0,000	-0,710	0,008	-1,235	0,045	-0,755	0,003
Tofelts planskilt	-0,729	0,000	-0,686	0,000	-0,843	0,000	-0,010	0,977	-0,618	0,002
Øvrig TEN-T	-0,049	0,112	-0,028	0,431	0,215	0,007	0,486	0,000	0,276	0,000
Øvrig Ev/Rv	-0,063	0,018	-0,043	0,170	0,086	0,229	0,239	0,067	0,122	0,065
Fylkesveg		(ref.)		(ref.)		(ref.)		(ref.)		(ref.)
Ikke midtd./rekkv.		(ref.)		(ref.)		(ref.)		(ref.)		(ref.)
Kun midtdeler	-0,048	0,526	-0,160	0,053	-0,149	0,504	-0,271	0,578	-0,199	0,342
Kun midtrekkverk	-0,535	0,002	-0,503	0,003	-1,122	0,039	-15,509	0,986	-1,443	0,008
Midtdeler+rekkv.	-0,551	0,000	-0,583	0,000	-1,280	0,001	-2,322	0,012	-1,466	0,000
Forst. midtoppm.	-0,693	0,000	-0,714	0,000	-0,106	0,662	-0,026	0,947	-0,091	0,675
SATK (én retn.)	-0,173	0,407	-0,161	0,462	0,603	0,141	0,459	0,515	0,595	0,110
SATK (begge retn.)	-0,627	0,017	-0,727	0,009	-1,923	0,062	-0,866	0,419	-1,509	0,043
PATK	0,020	0,573	0,023	0,585	-0,111	0,296	-0,118	0,562	-0,111	0,259
Vegbelysning	0,047	0,058	0,095	0,001	0,045	0,504	-0,186	0,139	-0,001	0,991
1 - Østfold	0,385	0,000	0,424	0,000	0,235	0,119	0,272	0,354	0,253	0,071
2 - Akershus	0,087	0,086	0,099	0,103	0,251	0,074	0,222	0,426	0,260	0,047
3 - Oslo	0,553	0,000	0,487	0,000	1,043	0,000	0,980	0,126	1,027	0,000
4 - Hedmark	-0,062	0,248	-0,035	0,586	0,082	0,571	0,123	0,655	0,094	0,485
5 - Oppland	-0,086	0,121	-0,071	0,277	0,342	0,017	0,205	0,456	0,315	0,018
6 - Buskerud	-0,259	0,000	-0,169	0,010	0,101	0,499	0,287	0,306	0,146	0,287
7 - Vestfold	0,195	0,000	0,260	0,000	0,088	0,589	-0,247	0,483	0,055	0,718
8 - Telemark	0,400	0,000	0,525	0,000	-0,031	0,845	0,083	0,783	-0,002	0,992
9 - Aust-Agder	0,245	0,000	0,325	0,000	-0,056	0,757	0,193	0,557	0,004	0,980
10 - Vest-Agder		(ref.)		(ref.)		(ref.)		(ref.)		(ref.)
11 - Rogaland	-0,023	0,670	0,045	0,481	-0,028	0,853	0,322	0,256	0,052	0,710
12 - Hordaland	0,181	0,001	0,234	0,000	0,190	0,207	0,086	0,774	0,180	0,197
14 - Sogn og Fjor.	-0,046	0,473	0,017	0,825	-0,008	0,965	-0,235	0,478	-0,051	0,747
15 - Møre og Rom.	0,000	0,998	0,031	0,634	0,120	0,411	-0,108	0,712	0,087	0,523
16 - Sør-Trøndelag	0,301	0,000	0,372	0,000	0,205	0,170	0,155	0,590	0,197	0,154
17 - Nord-Trøndel.	-0,370	0,000	-0,350	0,000	-0,279	0,101	0,150	0,611	-0,169	0,268
18 - Nordland	-0,042	0,441	0,065	0,320	-0,044	0,769	0,162	0,556	0,006	0,963
19 - Troms	-0,272	0,000	-0,165	0,024	-0,341	0,049	0,179	0,551	-0,205	0,185
20 - Finnmark	-0,254	0,001	-0,213	0,017	-0,276	0,170	0,015	0,967	-0,202	0,264
Konstantterm	-16,584	0,000	-16,736	0,000	-17,703	0,000	-18,769	0,000	-17,423	0,000

<sup>1</sup> Prediktorvariabelen i modellene er Ln(antall per kilometer + 1).

Tabell 3.1.2 viser koeffisientene for beregning av overspredningsparameteren. Overspredningsparameteren beregnes på samme måte som normale ulykkestall som beskrevet ovenfor.

Tabell 3.1.2: Koeffisienter for beregning av overspredningsparameteren i de endelige modellene for personskadenulykker (Psu), lett skadde (LS), hardt skadde (HS), drepte og drepte eller hardt skadde (D/HS).

	Psu		LS		HS		Drepte		D/HS	
	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p	Koeff.	p
Ln(lengde * år)	-0,601	0,000	-0,674	0,000	-0,598	0,000	-1,024	0,000	-0,654	0,000
Ln(ÅDT)	-0,240	0,000	-0,749	0,000	-0,708	0,000	-0,742	0,000	-0,685	0,000
Konstantterm	5,920	0,000	12,165	0,000	12,181	0,000	16,719	0,000	12,453	0,000

### 3.2 Hvor gode er modellene?

Tabell 3.2.1 viser ulike goodness-of-fit (GOF) statistikker som beskriver hvor gode modellene predikerer registrerte ulykkestall:

- **Pseudo-R<sup>2</sup>**: Denne statistikken som er beregnet av statistikkprogrammet (Stata), er basert på loglikelihood og indikerer hvor «sannsynlig» et sett med modellparametre er, gitt de registrerte ulykkestallene. Jo større Pseudo-R<sup>2</sup>, desto bedre er modellen. Denne statistikken anvendes av statistikkprogrammet som kriterium i modellberegningen, dvs. at koeffisientene settes slik at R<sup>2</sup> blir maksimert.
- **MSPE (mean square prediction error)**: Dette er det gjennomsnittlige kvadrerte avviket mellom registrerte og predikerte antall ulykker / skadde / drepte. Jo mindre MSPE, desto bedre er modellen.
- **Elvik-indeks**: Elvik indeks er en funksjon av overspredningen i de predikerte og registrerte ulykkestallene og indikerer hvor stor andel av variansen i de registrerte ulykkestallene som er forklart av modellen.

Tabell 3.2.1 viser i tillegg til GOF indikatorene deskriptiv statistikk for registrerte og normale ulykkestall samt avvik mellom normale og registrerte ulykkestall for hele datamaterialet (alle år). For antall normale og registrerte ulykkestall per år vises summene og avvikene mellom normale og registrerte ulykkestall. Det prosentvise avviket i de årlige ulykkestallene er større enn i de totale ulykkestallene. Dette skyldes at normale ulykkestall for segmenter som inngår med flere år i datafilen er mer nøyaktige enn normale ulykkestall for segmenter med færre år med data.

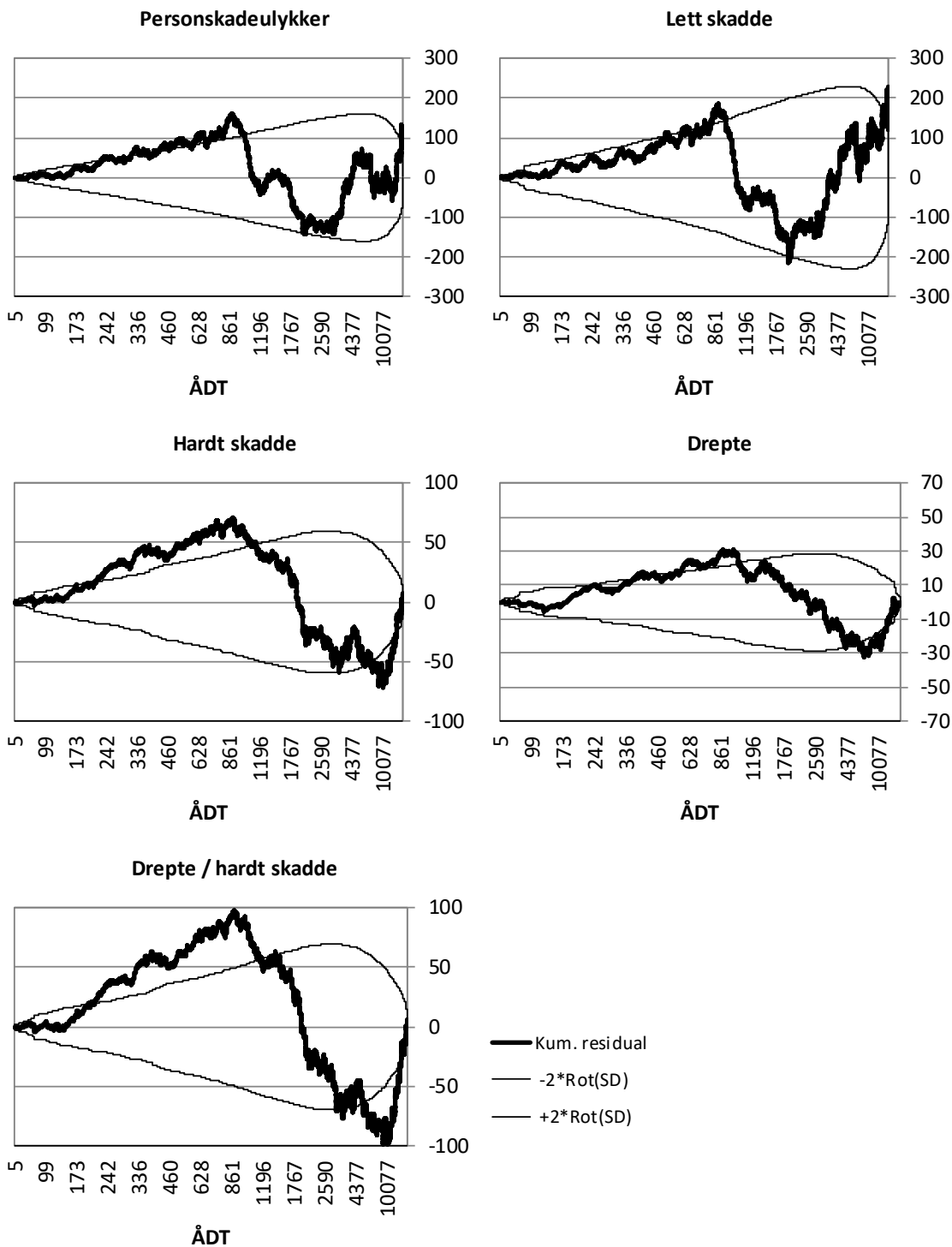
Den nederste delen av tabellen viser at avvikene for årlige ulykkestall varierer mellom segmenter med ulik ÅDT. At noen av avvikene er store, især i den høyeste ÅDT-klassen, skyldes små ulykkestall.

Tabell 3.2.1: GOF for de endelige modellene og avvik mellom normale og registrerte ulykkestall.

	km veg	Psu	LS	HS	D	D/HS
<b>De aktuelle modellene</b>						
Pseudo R <sup>2</sup>		0,1901	0,1395	0,0715	0,0621	0,0701
MSPE <sup>1</sup>		0,3379	0,6924	0,0464	0,0106	0,0631
Elvik-indeks		0,9424	0,8366	3,0060	4,8655	2,5293
<b>De forrige modellene (Høye, 2014)</b>						
Pseudo R <sup>2</sup>		0,200	0,144	0,075	0,074	0,073
MSPE <sup>1</sup>		0,544	1,286	0,062	0,015	0,086
Elvik-indeks		0,600	0,605	2,895	4,559	2,393
<b>Totale ulykkestall (alle år): Deskriptiv statistikk for registrerte og normale ulykkestall</b>						
Registrerte (sum)	50 150	18 141	22 153	2 537	655	3 192
Registrerte (maks.)	50 150	27	33	6	5	9
Registrerte (varians)	50 150	0,554	1,064	0,049	0,011	0,067
Registrerte (overspred.)	50 150	5,390	8,884	13,527	29,183	13,807
Normale (sum)	50 150	18 421	22 565	2 556	658	3 213
Normale (maks.)	50 150	17,6	23,5	1,1	0,3	1,4
Normale (varians)	50 150	0,260	0,425	0,003	0,000	0,005
Normale (overspred.)	50 150	0,310	1,452	-27,136	-112,809	-21,116
<b>Totale ulykkestall (alle år): Avvik normale minus registrerte ulykkestall</b>						
Avvik (antall)	50 150	279,9	411,6	19,1	2,5	20,6
Avvik (%)	50 150	<b>1,54 %</b>	<b>1,86 %</b>	<b>0,75 %</b>	<b>0,38 %</b>	<b>0,64 %</b>
<b>Per år: Antall normale og registrerte ulykkestall</b>						
Registrerte (sum)	50 150	3 347	4 092	456	118	574
Predikerte (sum)	50 150	3 403	4 190	467	119	586
<b>Per år: Avvik normale minus registrerte ulykkestall</b>						
Avvik (antall)	50 150	55,6	98,2	11,5	1,0	12,3
Avvik (%)	50 150	<b>1,66 %</b>	<b>2,40 %</b>	<b>2,51 %</b>	<b>0,86 %</b>	<b>2,15 %</b>
<b>Prosentvise avvik mellom normale og registrerte antall per år på veger med ulik ÅDT</b>						
ÅDT ≤ 1000	33 026	4,1 %	3,6 %	12,1 %	19,4 %	13,3 %
1 000 < ÅDT ≤ 1 500	4 337	-5,5 %	-7,4 %	-11,5 %	-13,1 %	-12,0 %
1 500 < ÅDT ≤ 2 000	2 703	-7,8 %	-6,4 %	-14,6 %	-21,0 %	-16,1 %
2 000 < ÅDT ≤ 5 000	6 350	3,6 %	5,1 %	-3,2 %	-8,2 %	-4,4 %
5 000 < ÅDT ≤ 10 000	2 187	-0,8 %	1,6 %	-2,6 %	-6,8 %	-3,3 %
10 000 ≤ ÅDT	1 546	4,7 %	4,6 %	23,3 %	33,5 %	25,3 %

<sup>1</sup>Ikke sammenlignbar mellom modeller for ulike skadegrader

Figur 3.2.1 viser *Cure-plots* for alle skadegrader. Cure står for cumulative residuals. En CURE-plot viser hvor godt modeller predikerer registrerte ulykkestall ved ulike verdier av en eller flere prediktorvariabler (Caliendo et al., 2007), i dette tilfelle trafikkmengde. For å lage en Cure plot sorteres dataene etter den aktuelle prediktorvariabelen (her trafikkmengde) og man beregner kumulative residualer, dvs. at man for hver datapunkt beregner summen av alle avvik mellom registrerte og predikerte ulykkestall for alle datapunktene fra den minste trafikkmengden til trafikkmengden på det aktuelle datapunktet. Når de kumulative residualene (de brede linjene) ligger innenfor ± to standardavvik (de smale linjene) anser man som regel modellene som «gode» dvs. at de predikerte tallene ikke viser store systematiske avvik. De kumulative residualene begynner alltid på null (ved den minste trafikkmengden) og dersom det totale antall predikerte er lik det totale antall registrerte ulykker ender linjen som viser de kumulative residualene også på null (ved den høyeste trafikkmengden).



Figur 3.2.1: Cure plots for alle skadegrader, endelige modeller.

Cure-plottene i figur 3.2.1 viser at de kumulative residualene delvis ligger utenfor de smale linjene, men slutter forholdsvis nær null. Sistnevnte viser at summen av de predikerte ulykestallene er omtrent lik de registrerte. Sammenlignet med modellene til Høye (2014) er det større svingninger i de kumulative residualene utenfor de smale linjene, men avvikene mellom summen av predikerte og registrerte er betydelig mindre, især for LS.

## 4 Trendfaktor

Normale ulykkestall som er beregnet med ulykkesmodellene kan tolkes som resultater for 2012. For å omregne resultatene til andre år må man justere to tall: Det normale ulykkestall og overspredningsparameteren. For å trendjustere normale ulykkestall er det utviklet en trendfaktor som er basert på informasjon om transportarbeid i Norge i 1997 til 2014 med alle typer motorkjøretøy (Farstad, 2015) og antall Psu, LS, HS, drepte og D/HS i 1997 til 2014 (SSB). Trendjusteringen av overspredningsparameteren gjøres ut fra sammenhengen mellom normale ulykkestall og overspredningsparameteren. Trendjusteringen av normale ulykkestall er beskrevet i avsnitt 4.2, trendjusteringen av overspredningsparameteren i avsnitt 4.3.

Når man trendjusterer normale ulykkestall som er beregnet med ulykkesmodellen, må man i modellberegningen oppgi trafikkmengden for det året det skal trendjusteres til. Når man kun har trafikkmengden for et annet år kan man beregne trafikkmengden for det året som skal trendjusteres til som beskrevet i avsnitt 4.3.

### 4.1 Trendjustering av de forrige modellene

Tabell 4.1.1 viser registrerte og predikerte ulykkestall per million kjøretøykilometer i modellene fra 2014 og de aktuelle modellene, samt trendfaktoren fra 2014 og ulykkestallene som skulle ha vært funnet i 2016 dersom utviklingen hadde fulgt trendfaktoren. Den høyre kolonnen i tabellen viser avvikene. Avvikene er på over 10% for LS. I Høye (2014) hadde antall LS vist seg å være systematisk overpredikert av modellene, relativt store avvik i negativ retning er derfor som forventet for de normale antallene (ikke for de registrerte).

Tabell 4.1.1: Ulykkestall per million kjøretøykilometer i modellene fra 2014 og 2016, trendfaktor fra 2014, ulykkestall hvis utviklingen hadde fulgt trendfaktoren og avvik.

		Per mill. kjøretøykilometer				Avvik
		Trend-faktor	2014-modell	2016-modell	2016 med trendfaktor	
<b>Registrerte</b>	<b>Psu</b>	0,749	0,157	0,112	0,117	-4,3%
	<b>LS</b>	0,751	0,204	0,137	0,154	-10,7%
	<b>HS</b>	0,731	0,020	0,016	0,015	6,1%
	<b>Drept</b>	0,738	0,006	0,004	0,004	-1,6%
	<b>D/HS</b>	0,733	0,026	0,020	0,019	4,5%
<b>Predikerte</b>	<b>Psu</b>	0,749	0,158	0,113	0,118	-4,7%
	<b>LS</b>	0,751	0,209	0,138	0,157	-12,2%
	<b>HS</b>	0,731	0,020	0,016	0,015	5,3%
	<b>Drept</b>	0,738	0,006	0,004	0,004	-2,2%
	<b>D/HS</b>	0,733	0,026	0,020	0,019	4,0%

## 4.2 Trendjustering av predikerte (normale) ulykkestall

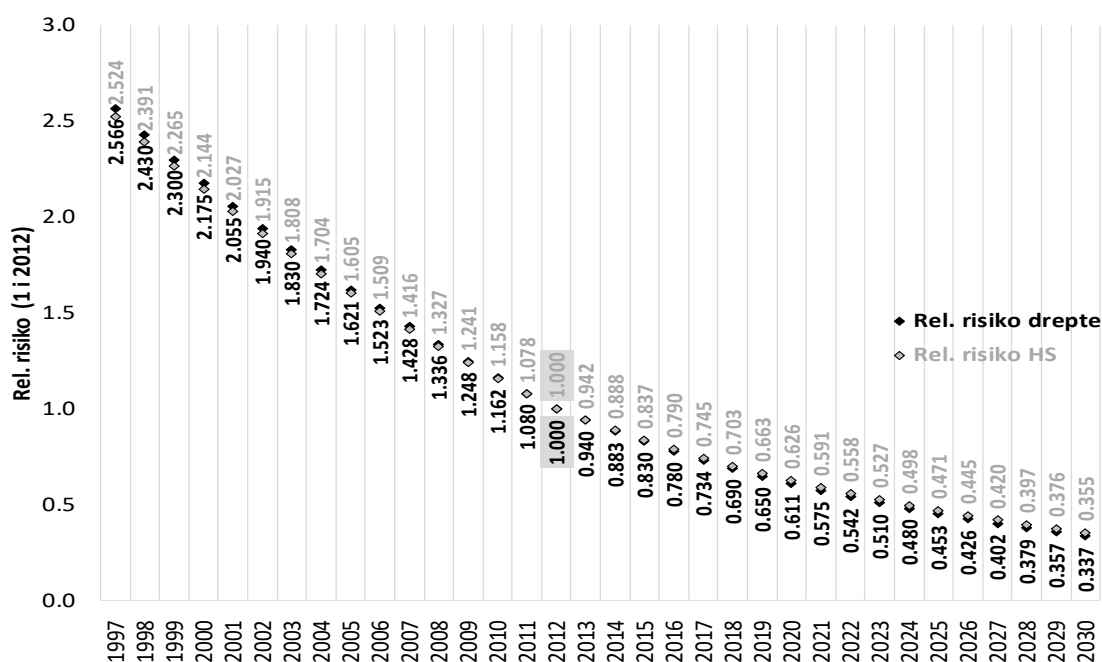
For å trendjustere normale ulykkestall til et år X ganges disse med det relative risikotall for året X som er vist i figur 4.2.1 for drepte og HS, i figur 4.2.2 for D/HS og i figur 4.2.3 for LS og Psu. De relative risikotallene for hvert år i perioden 1997-2030 viser risikoen i år X i forhold til risikoen i år 2012. Risikoen for hvert år er beregnet som antall skadde / drepte som er beregnet med en trendfunksjon delt på den totale trafikkmengden. Dette betyr at de relative risikotallene viser hvor mange ulykker det hadde vært i et år X, i forhold til 2012, hvis alt annet inkludert trafikkmengden hadde vært likt. Når man trendjusterer normale ulykkestall forutsettes dermed at man har oppgitt trafikkmengden for det året man trendjusterer til.

Risikotallene for 1997-2030 er beregnet med to forskjellige trendfunksjoner for antall skadde / drepte:

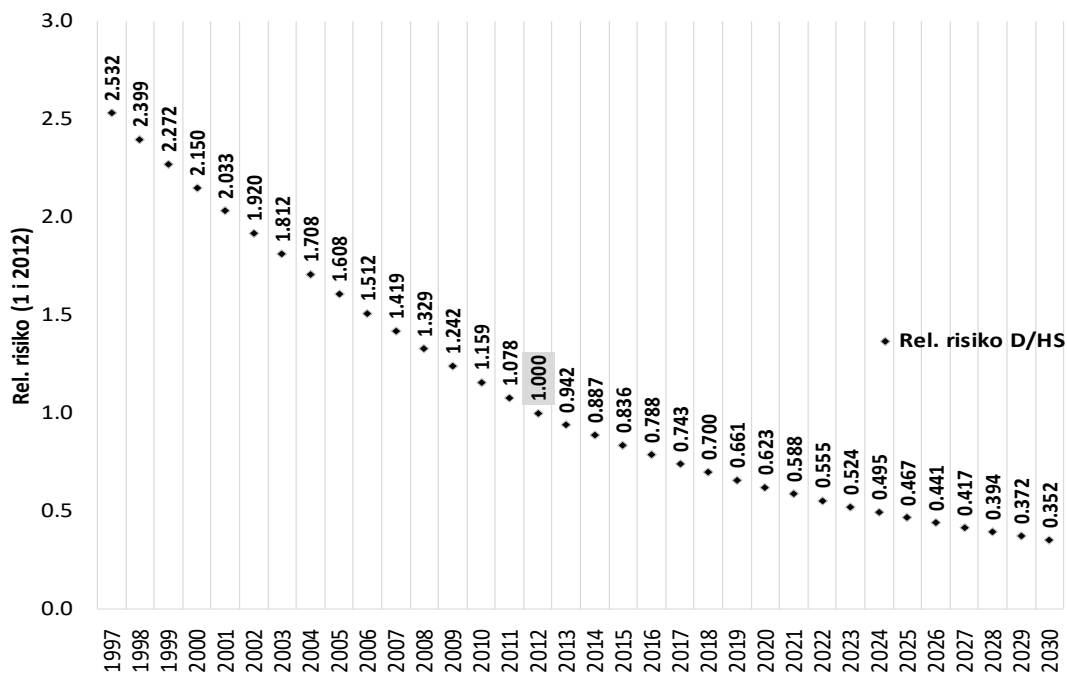
- For årene 1997-2014 brukes for hver skadegrad den trendfunksjonen som beskriver de faktiske antallene skadde / drepte i årene 1997-2014 best (lineær trend for drepte og HS, polynom trend for LS og Psu).
- For årene 2015-2020 brukes for alle skadegradene en eksponentiell trendfunksjon (som er basert på faktiske risikotall i 1997-2014); dermed unngår man å predikere en urealistisk stor nedgang av ulykkestallene i fremtidige år (med den første trendfunksjonen ville den relative risikoen krysse null-linjen allerede få år etter 2020).

For trafikkmengden er det beregnet en lineær trendfunksjon basert på transportytelsene ifølge Farstad (2015). Risikotallene for drepte, HS og D/HS følger ikke en lineær trend, selv om de er beregnet med en lineær trend for antall drepte, HS og D/HS.

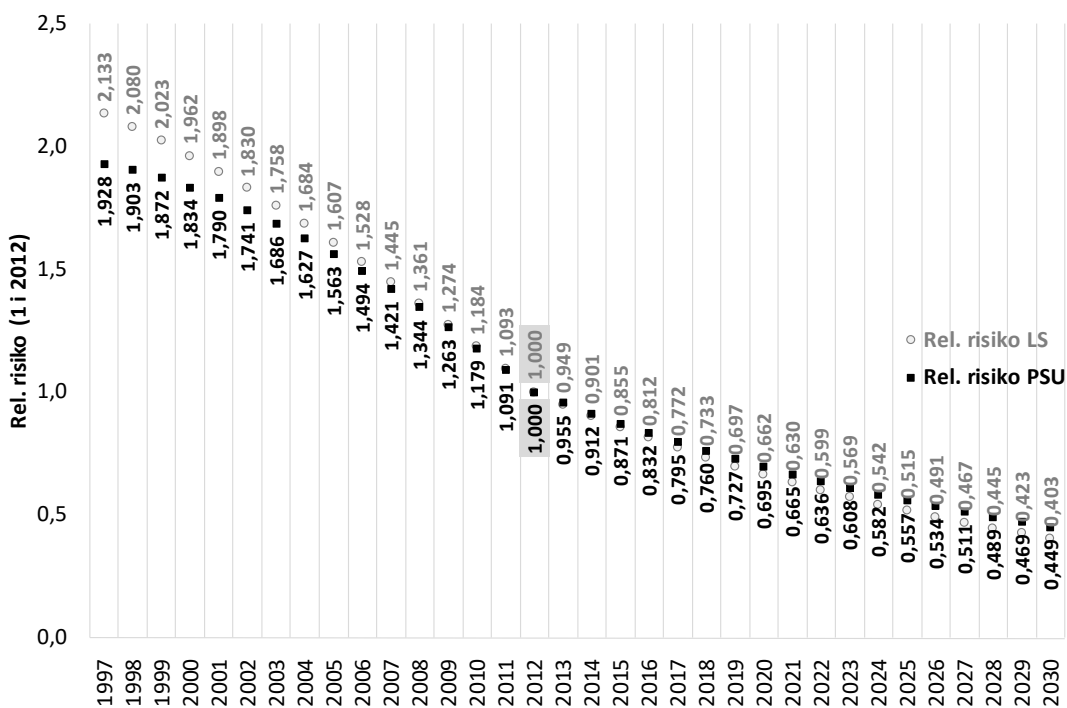
Referanseåret for den relative risikoen er 2012, dvs. at den relative risikoen for hver skadegrad er satt lik én i 2012. 2012 er det «gjennomsnittlige» året i datafilen som er lagt til grunn for beregningen av ulykkesmodellene.



Figur 4.2.1: Relativ risiko for antall drepte og HS i 1997-2030 (relativ risiko i 2012 = 1).



Figur 4.2.2: Relativ risiko for antall D/HS i 1997-2030 (relativ risiko i 2012 = 1).



Figur 4.2.3: Relativ risiko for antall LS og Psu i 1997- (relativ risiko i 2012 = 1).

Den store nedgangen av den relative risikoen i løpet av årene 1997 til 2012 (og videre til 2020) kan forklares med at det har vært en stor økning av trafikkarbeidet, samtidig som antall ulykker, skadde og drepte har gått ned:

- Trafikkarbeidet med alle typer motorkjøretøy har økt fra ca. 33000 mill. kjøretøykilometer i 1997 til ca. 45 946 mill. kjøretøykilometer i 2015 (+37%)
- Antall Psu har gått ned fra 8765 i 1997 til 4972 i 2012 (-43%)
- Antall drepte har gått ned fra 303 i 1997 til 147 i 2012 (-51%)
- Antall HS har gått ned fra 1,341 i 1997 til 674 i 2012 (-50%)

### 4.3 Trendjustering av overspredningsparameteren

Når man trendjusterer normale ulykkestall må man, i en EB-analyse, i tillegg til de normale ulykkestallene også justere overspredningsparameteren. Den trendjusterte overspredningsparameteren beregnes som en funksjon av det normale ulykkestall for det året man har data for og det året det skal trendjusteres til, samt to koeffisienter, etter følgende formel (a kan droppes fra beregningen av trendfaktoren):

$$Faktor = \frac{a * x_1^b}{a * x_2^b}$$

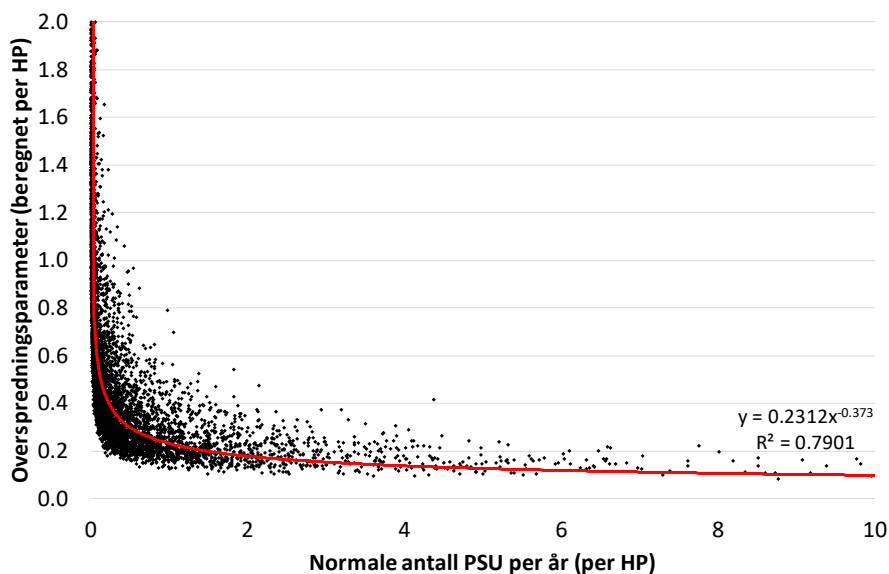
I formelen er  $x_1$  det normale ulykkestallet i året det skal justeres til og  $x_2$  er det normale ulykkestallet som er beregnet med modellen (og som gjelder året 2012), a og b er koeffisienter. Hvilke koeffisienter som skal brukes avhenger av skadegraden og er vist i tabell 4.3.1. I utgangspunktet anbefales koeffisientene som står under «Anbefalte koeffisienter». De alternative koeffisientene kan ev. brukes for korte segmenter (under 1,2 km, se forklaring nedenfor).

Tabell 4.3.1: Koeffisienter for beregning av overspredningsparameteren som en funksjon av normale ulykkes- og skadetall med en formel av typen  $y = a * x^b$  hvor y er overspredningsparameteren, x er normalt ulykkestall (ikke trendjustert) og a og b er koeffisientene.

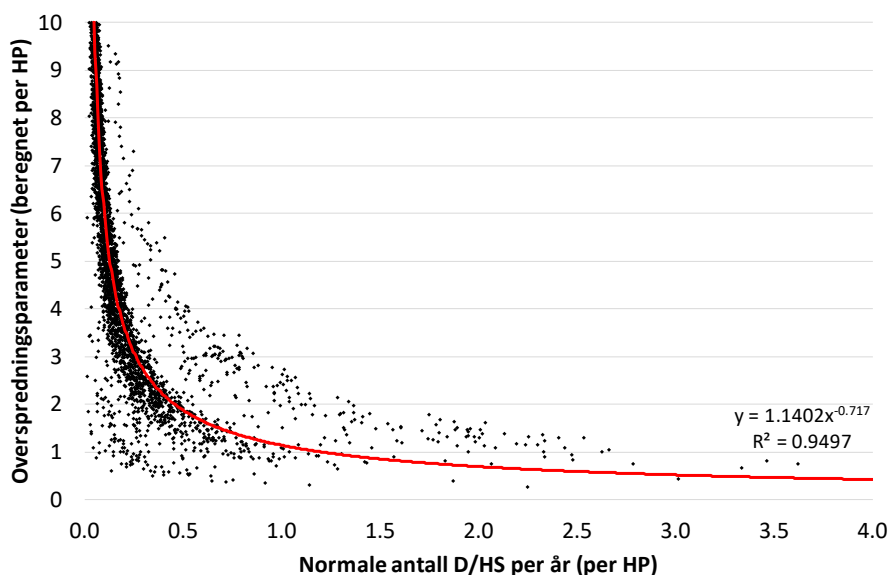
Koeffisient	Psu	LS	HS	D	D/HS
<b>Anbefalte koeffisienter</b>					
a	0,2312	1,2724	1,0545	0,411	1,1402
b	-0,373	-0,715	-0,713	-0,824	-0,717
<b>Alternative koeffisienter (for korte segmenter, &lt; 1,2 km)</b>					
a	0,5065	1,1744	0,5532	1,449	0,7222
b	-0,309	-0,715	-0,736	-0,690	-0,734

Trendjusteringen med de anbefalte koeffisientene er basert på sammenhengen mellom predikerte ulykkestall og overspredningsparameter i en aggregert datafil hvor én datapunkt er en hel hovedparsell på én veg i ett fylke i ett år. Sammenhengen for Psu er vist i figur 4.3.1 og sammenhengen for antall D/HS er vist i figur 4.3.2. De røde linjene viser trendfunksjonene (potensfunksjon) som ligger til grunn for koeffisientene i tabell 4.3.1. Funksjonene med LS, HS og drepte ligner på funksjonene for Psu og D/HS.





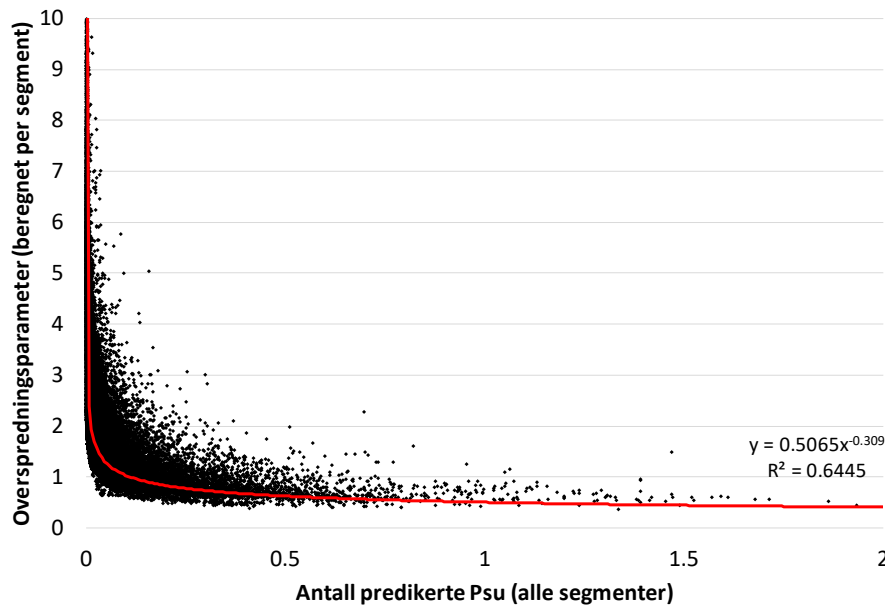
Figur 4.3.1: Sammenhengen mellom normale antall Psu per år og overspredningsparameter i aggregert datafil (segmentlengde i gjennomsnitt 7,65 km, maks. 59,14 km).



Figur 4.3.2: Sammenhengen mellom normale antall D/HS per år og overspredningsparameter i aggregert datafil (segmentlengde i gjennomsnitt 7,65 km, maks. 59,14 km).

En potensfunksjon av formen  $y = a * x^b$  (y er overspredningsparameteren og x er normalt ulykkestall) er funksjonen som beskriver sammenhengen best. Figurene viser likevel at det er forholdsvis stor spredning rundt trendlinjene.

Beregner man slike funksjoner basert på originaldatafilen hvor alle segmentene er 1,2 km eller kortere er det større spredning av datapunktene rundt trendlinjen og for det samme normale ulykkestallene er overspredningsparameteren i gjennomsnitt noe høyere (figur 4.3.3 viser funksjonen for Psu som eksempel). Derfor anbefales i utgangspunktet å bruke koeffisientene som er beregnet med den aggregerte datafilen. Disse gjelder for forholdsvis lange segmenter; i den aggregerte datafilen er segmentene i gjennomsnitt 7,65 km lange (opptil 59,14 km). For segmenter under 1,2 km lengde kan man likevel bruke resultatene fra originaldatafilen. Dermed vil man unngå å underestimere overspredningsparameteren.



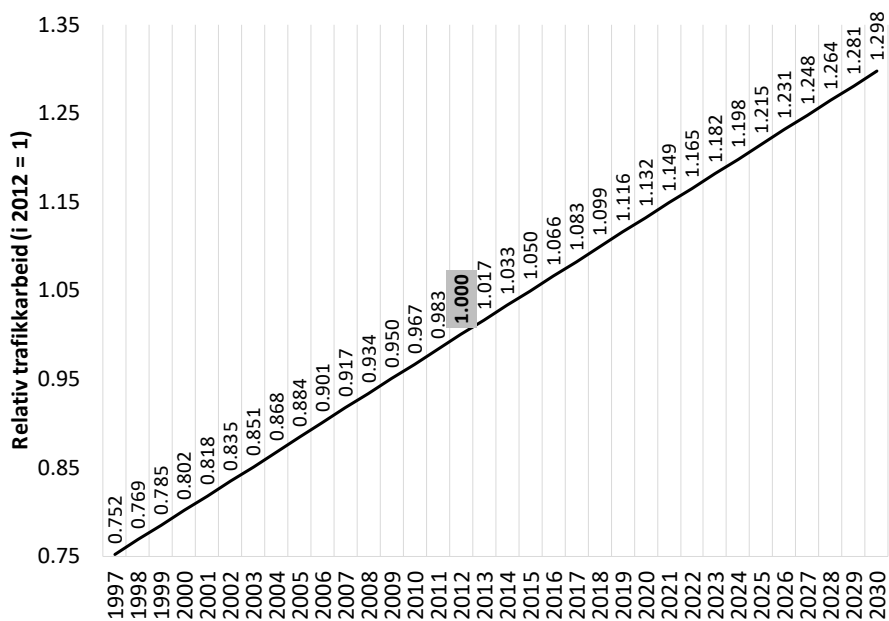
Figur 4.3.3: Sammenhengen mellom normale antall Psu per år og overspredningsparameter i original datafil (segmentlengde i gjennomsnitt 0,66 km, maks. 1,2 km).

Sammenhengen mellom normalt antall ulykker og overspredningsparameteren avhenger av hvilket datagrunnlag som brukes for å beregne sammenhengen. Dermed vil sammenhengen strengt tatt ikke være overførbart til andre datagrunnlag og trendjusteringen av overspredningsparameteren vil alltid være unøyaktig. Det finnes imidlertid ingen mulighet for å trendjustere overspredningsparameteren på en standardisert måte som ikke medfører noe unøyaktighet. Beregner man den trendjusterte overspredningsparameteren basert på originaldatafilen er unøyaktigheten større enn når den beregnes basert på den aggregerte datafilen, især for drepte, HS og D/HS, men man unngår å underestimere overspredningsparameteren. Totalt sett har det imidlertid forholdsvis liten betydning for den trendjusterte overspredningsparameteren og dermed for EB-vektene hvilken trendjustering som velges for overspredningsparameteren.

#### 4.4 Trendjustering av trafikkmengden

Trendfaktoren som er beskrevet i forrige avsnitt tar ikke hensyn til endringer av trafikkmengden i den forstand at de relative risikotallene er beregnet som antall ulykker i forhold til den faktiske trafikkmengden i de respektive årene. Dersom man justerer et antall ulykker til et annet år med hjelp av trendfaktoren forutsettes derfor implisitt at trafikkmengden i det året man justerer til, er nøyaktig det samme som i det året hvor ulykkestallet ble beregnet. Hvis man vil ta hensyn til at også trafikkmengden endrer seg over tid, kan man beregne ulykkestallet med den antatte trafikkmengden i det året det skal justeres til. Figur 4.4.1 viser den relative totale trafikkmengden (antall kjøretøykilometer med både person- og godstransport) i 1997 til 2030. Den relative trafikkmengden i 2012 er satt lik én. Dette er den samme relative trafikkmengden som ble lagt til grunn i beregningen av de relative risikotallene for trendfaktoren.

Hvis man for eksempel har trafikkmengden for året 2012 og vil justere denne til året 2020, ganger man trafikkmengden med 1,132. Hvis man har trafikkmengden for 2020 og vil justere denne til året 2025, ganger man trafikkmengden først med 1,215 (den relative trafikkmengden i 2025) og deler resultatet på 1,132 (den relative trafikkmengden i 2020).



Figur 4.4.1: Relativ trafikkmengde (antall kjøretøykilometer, person- og godstransport; relativt antall i 2012 = 1).

## 5 Brukerveiledning for Ulykkesmodeller.xlsx

I vedlagte xls-regneark Ulykkesmodeller.xlsx kan man beregne:

- **Normale antall** Psu, LS, HS, drepte og D/HS som en funksjon av prediktorvariablene som er beskrevet ovenfor
- **Overspredningsparameteren** for hver skadegrad
- **EB-vekten** for hver skadegrad som kan brukes til å beregne forventede ulykkes/-skadetall
- **Forventede antall** Psu, LS, HS, drepte og D/HS (hvis man i tillegg oppgir registrerte ulykkes- og skadetall)

Beregningene kan gjøres for én vegstrekning om gangen. Vegstrekningen har ideelt sett konstante vegegenskaper på hele strekningen (for eksempel samme fartsgrense, trafikkmengde mv.). Fremgangsmåten for heterogene vegstrekninger er beskrevet nedenfor. Alle tallene beregnes både uten trendjustering og med trendjustering til et valgfritt år mellom 1997-2030. Resultater uten trendjustering gjelder året 2012.

Brukeren må oppgi følgende informasjon i de gule cellene i regnearket:

- **Alle vegegenskapene som er prediktorer i ulykkesmodellene:**
  - **Trafikkmengde (ÅDT)** (skal man bruke trendjustering oppgis trafikkmengden for det året man skal justere til)
  - **Fartsgrense (km/t)**
  - **Antall kjørefelt**
  - **Antall T-kryss**
  - **Antall X-kryss**
  - **Antall rundkjøringer**
  - **Antall ramper** (av- eller påramper)
  - **Type veg** (velges fra rullegardinmeny)
  - **ATK** (velges fra rullegardinmeny)
  - **Vegbelysning**
  - **Midtdeler, midtrekkverk** (ikke nødvendig å oppgi på motorveger<sup>1</sup>)
  - **Forsterket midtoppmerking** (kun på veger uten midtdeler/-rekkverk; velg "Nei" hvis vegen også har midtdeler/-rekkverk)
  - **Fylke** (velges fra rullegardinmeny)
- **Lengde (meter)**
- **Antall år:** Dette er antall år som man vil beregne normale ulykkestall for (obs: trendjustering kan kun gjøres for ett år om gangen, se nedenfor)

---

<sup>1</sup> Virkningen av midtdeler og midtrekkverk på motorveg er fanget opp av vegtypevariabelen (motorveg); resultatene i regnearket er derfor uavhengige av hva som er oppgitt for midtdeler og midtrekkverk når vegtypen er motorveg.

- **År for trendjustering:** Dette er året det skal trendjusteres til; år mellom 1997 og 2030 kan oppgis (eller utelates)
- **Justering av overspredningsparameteren:** Man kan velge om overspredningsparameteren skal trendjusteres for lange eller korte segmenter; det gis følgende anbefaling:
  - Generelt anbefales trendjustering for lange segmenter; denne er, især for drepte, HS og D/HS mer nøyaktig enn for korte segmenter
  - For korte strekninger som omfatter kun en liten andel av en hovedparsell (under 1,2 km) kan trendjusteringen for korte segmenter brukes

Noen praktiske problemstillinger er beskrevet i det følgende.

**Heterogene strekninger (strekninger med varierende vegegenskaper):** Ideelt sett har hele strekningen de samme vegegenskapene. For å beregne normale ulykkestall mv. for strekninger med varierende egenskaper kan man dele strekningen opp i homogene segmenter med konstante egenskaper og summere resultatene. Overspredningsparameter og EB-vekt lar seg ikke summere. Skal man beregne forventede ulykkestall for heterogene strekninger må man først beregne det forventede ulykkestall for hver (homogen) delstrekning (som funksjon av normalt ulykkestall, overspredningsparameter og registrerte ulykkestall for hver delstrekning). De forventede ulykkestallene kan man summere til det forventede ulykkestallet for hele strekningen.

**Analysert for flere år:** Skal man trendjustere resultatene må man gjøre én beregning for hvert år som skal inngå i analysen (trendjusteringen kan kun gjøres for ett år om gangen). Normale ulykkestall kan summeres for flere år. Forventede ulykkestall for flere år beregnes på samme måte som forventede ulykkestall for heterogene strekninger (se forrige punkt).

**Trendjustering og trafikkmengde:** Trendfaktoren (de relative risikotallene) justerer kun for endringer av ulykkesrisikoen over tid (dvs. antall ulykker / skadde / drepte per million kjøretøykilometer). Trafikkmengdeendringer over tid er ikke tatt hensyn til i trendjusteringen. Trafikkmengden må derfor alltid oppgis for det aktuelle året (året det skal trendjusteres til). Skal man for eksempel beregne det normale antall ulykker på en veg i 2020 oppgir man trafikkmengden på denne vegen i 2020. De beregnede normale ulykkestallene viser hvor mange ulykker man normalt kunne ha forventet på en veg med den gitte trafikkmengden i 2012.

## 6 Referanser

- Caliendo, C., Guida, M. & Parisi, A. (2007). A crash-prediction model for multilane roads. *Accident Analysis and Prevention*, 39, 657-670.
- Elvik, R. (2012). Oppdatering av ulykkesmodeller: Foreløpig analyse av data for Sør-Trøndelag. TØI-Arbeidsdokument 50195. Oslo: Transportøkonomisk institutt.
- Farstad, E. (2015). Transportytelser i Norge 1946-2014. TØI-Rapport 1454/2015. Oslo: Transportøkonomisk institutt.
- Geedipally, S. R., Lord, D., & Park, B.-J. (2009). Analyzing different parameterizations of the varying dispersion parameter as a function of segment length. *Transportation Research Record*, 2103, 108-118.
- Hauer, E. (2001). Overdispersion in modelling accidents on road sections and in empirical Bayes estimation. *Accident Analysis and Prevention*, 33, 799-808.
- Hauer, E., Harwood, D. W., Council, F. M., & Griffith, M. S. (2002). Estimating safety by the empirical Bayes method: A tutorial. *Transportation Research Record*, 1784, 126-131.
- Høye, A. (2013). Ulykkesmodeller. TØI-Arbeidsdokument 50429. Oslo: Transportøkonomisk institutt.
- Høye, A. (2014). Utvikling av ulykkesmodeller for ulykker på riks- og fylkesvegnettet i Norge. TØI-rapport 1323/2014. Oslo: Transportøkonomisk institutt.
- Lord, D., & Park, P. Y.-J. (2008). Investigating the effects of the fixed and varying dispersion parameters of Poisson-gamma models on empirical Bayes estimates. *Accident Analysis & Prevention*, 40(4), 1441-1457.
- Miaou, S.-P., & Lord, D. (2003). Modeling traffic crash-flow relationships for intersections. *Transportation Research Record*, 1840, 31-40.
- Ragnøy, A. & Elvik, R. (2003). Trafikksikkerhetsanalyse av stamvegnettet i Norge. TØI-rapport 649/2003. Oslo: Transportøkonomisk institutt.
- Strathman, J. G., Dueker, K. J., Zhang, J., & Williams, T. (2001). Analysis of design attributes and crashes on the Oregon highway system. Report FWHA-OR-RD-02-01. Center for Urban Studies. College of Urban and Public Affairs. Portland State University. Portland, Oregon.

### Flere relevante referanser

- Abdel-Aty, M., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, 32, 633-642.
- Anastasopoulos, P. C., & Mannering, F. (2011). An empirical assessment of fixed and random parameter logit models using crash- and non-crash-specific injury data. *Accident Analysis & Prevention*, 43, 1140-1147.

- Anastasopoulos, P. C., Tarko, A. P., & Mannering, F. L. (2008). Tobit analysis of vehicle accident rates on interstate highways. *Accident Analysis & Prevention*, 40(2), 768-775.
- Bonneson, J. A., & Pratt, M. P. (2008). Procedure for developing accident modification factors from cross-sectional data. *Transportation Research Record*, 2083, 40-48.
- Cenek, P. D., Davies, R. B., & Henderson, R. J. (2012). Crash risk relationships for improved road safety management. NZ Transport Agency research report 488. Wellington, New Zealand.
- Couto, A., & Ferreira, S. (2011). A note on modeling road accident frequency: A flexible elasticity model. *Accident Analysis & Prevention*, 43(6), 2104-2111.
- Dietze, M. & Weller, G. (2011). Applying speed prediction models to define road sections and to develop accident prediction models: A German case study and a Portuguese exploratory study. Deliverable 6.2 of ERANET project RISMET. Dresden, Technical University of Dresden, 2011.
- Dinu, R. R., & Veeraragavan, A. (2011). Random parameter models for accident prediction on two-lane undivided highways in India. *Journal of Safety Research*, 42(1), 39-42.
- Findley, D. J., Hummer, J. E., Rasdorf, W., Zegeer, C. V., & Fowlere, T. J. (2012). Modeling the impact of spatial relationships on horizontal curve safety. *Accident Analysis & Prevention*, 45, 296-304.
- Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., & Krogsgård Thomsen, L. (1995). Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. *Accident Analysis & Prevention*, 27(1), 1-20.
- Garay, A. M., Hashimoto, E. M., Orteg, E. M. M., & Lachos, V. H. (2011). On estimation and influence diagnostics for zero-inflated negative binomial regression models. *Computational Statistics and Data Analysis*, 55, 1304-1318.
- Geedipally, S. R., Lord, D., & Park, B.-J. (2009). Analyzing different parameterizations of the varying dispersion parameter as a function of segment length. *Transportation Research Record*, 2103, 108-118.
- Geedipally, S. R., Lord, D., & Dhavala, S. S. (2012). The negative binomial-Lindley generalized linear model: Characteristics and application using crash data. *Accident Analysis & Prevention*, 45(0), 258-265.
- Hanley, K. A., Gibby, A., & Ferrara, T. (2000). Analysis of accident reduction factors on California state highways. *Transportation Research Record*, 1717, 37-45.
- Ivan, J. N., & O'Mara, P. J. (1997). Prediction of traffic accident rates using Poisson regression. Transportation Research Board 76th Annual Meeting, Washington DC.
- Ivan, J. N., Wang, C., & Bernardo, N. R. (2000). Explaining two-lane highway crash rates using land use and hourly exposure. *Accident Analysis & Prevention*, 32, 787-795.
- Kaminska, J., & Elvik, R. (2012). Choice of functional form and type of model to describe the relationship between independent variables and the dependent variable in accident prediction models. Unpublished Manuscript. Oslo, Institute of Transport Economics.
- Khattak, A. J., Khattak, A. J., & Council, F. M. (2002). Effects of work zone presence on injury and non-injury crashes. *Accident Analysis & Prevention*, 34, 19-29.
- Kim, D.-G., & Washington, S. (2006). The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. *Accident Analysis & Prevention*, 38(6), 1094-1100.

- Lord, D., & Mahlawat, M. (2009). Examining application of aggregated and disaggregated Poisson-gamma models subjected to low sample mean bias. *Transportation Research Record*, 2136, 1-10.
- Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, 44(5), 291-305.
- Lord, D., & Persaud, B. (2000). Accident prediction models with and without trend. *Transportation Research Record*, 1717, 102-108.
- Lord, D., Washington, S., & Ivan, J. N. (2007). Further notes on the application of zero-inflated models in highway safety. *Accident Analysis & Prevention*, 39(1), 53-57.
- Lord, D., Washington, S. P., & Ivan, J. N. (2005). Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: Balancing statistical fit and theory. *Accident Analysis & Prevention*, 37(1), 35-46.
- Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, 44(5), 291-305.
- Miaou, S.-P. (1996). Measuring the goodness of fit of accident prediction models. FHWA-RD-96-040.
- Miaou, S.-P., Lu, A., & Lum, H. S. (1996). Pitfalls of using  $R^2$  to evaluate goodness of fit of accident prediction models. *Transportation Research Record*, 1542, 6-13.
- Milton, J., & Mannering, F. (1998). The relationship among highway geometries, traffic-related elements and motor-vehicle accident frequencies. *Transportation*, 25, 395-413.
- Mitra, S., & Washington, S. (2007). On the nature of over-dispersion in motor vehicle crash prediction models. *Accident Analysis & Prevention*, 39(3), 459-468.
- Noland, R. B., & Oh, L. (2004). The effect of infrastructure and demographic change on traffic-related fatalities and crashes: A case study of Illinois county-level data. *Accident Analysis & Prevention*, 36, 525-532.
- Oh, J., Lyon, C., Washington, S., Persaud, B., & Bared, J. (2003). Validation of FHWA crash models for rural intersections. *Transportation Research Record*, 1840, 41-49.
- Poch, M., & Mannering, F. (1996). Negative binomial analysis of intersection-accident frequencies. *Journal of Transportation Engineering*, 122(2), 105-113.
- Ragnøy, A., Christensen, P., & Elvik, R. (2002). Skadegradstetthet - SGT. Et nytt mål på hvor farlig en vegstrekning er. TØI-Rapport 618/2002. Oslo: Transportøkonomisk institutt.
- Sawalha, Z., & Sayed, T. (2001). Evaluating safety of urban arterial roadways. *Journal of Transportation Engineering*, 127(2), 151-158.
- Shankar, V. N., Albin, R. B., Milton, J. C., & Mannering, F. L. (1998). Evaluating median cross-over likelihoods with clustered accident counts: An empirical inquiry using random effects negative binomial model. *Transportation Research Record*(1635), 44-48.
- Shankar, V., Milton, J., & Mannering, F. (1997). Modeling accident frequencies as zero-altered probability processes: An empirical inquiry. *Accident Analysis & Prevention*, 29(6), 829-837.



- Shankar, V. N., Ulfarsson, G. F., Pendyala, R. M., & Nebergall, M. B. (2003). Modeling crashes involving pedestrians and motorized traffic. *Safety Science*, 41, 627-640.
- Vogt, A., & Bared, J. (1998). Accident models for two-lane rural roads: Segments and intersections. Report FHWA-RD-98-133.
- Wang, X., & Abdel-Aty, M. A. (2007). Investigation of signalized intersection right-angle crash occurrence at intersection, roadway, and approach levels. Paper presented at the 84th Annual Meeting of the TRB, Washington, D.C.
- Washington, S., Persaud, B., Lyon, C., & Oh, J. (2005). Validation of accident models for intersections. Federal Highway Administration, Final Report, FHWA-RD-03-037.
- Zhang, C., & Ivan, J. N. (2005). Effects of geometric characteristics on head-on crash incidence on two-lane roads in Connecticut. *Transportation Research Record*, 159-164.

## Transportøkonomisk institutt (TØI)

### Stiftelsen Norsk senter for samferdselsforskning

TØI er et anvendt forskningsinstitutt, som mottar basisbevilgning fra Norges forskningsråd og gjennomfører forsknings- og utredningsoppdrag for næringsliv og offentlige etater. TØI ble opprettet i 1964 og er organisert som uavhengig stiftelse.

TØI utvikler og formidler kunnskap om samferdsel med vitenskapelig kvalitet og praktisk anvendelse. Instituttet har et tverrfaglig miljø med rundt 70 høyt spesialiserte forskere.

Instituttet utgir tidsskriftet Samferdsel med 10 nummer i året og driver også forskningsformidling gjennom TØI-rapporter, artikler i vitenskapelige tidsskrifter, samt innlegg og intervjuer i media. TØI-rapportene er gratis tilgjengelige på instituttets hjemmeside [www.toi.no](http://www.toi.no).

TØI er partner i CIENS Forskningscenter for miljø og samfunn, lokalisert i Forskningsparken nær Universitetet i Oslo (se [www.ciens.no](http://www.ciens.no)). Instituttet deltar aktivt i internasjonalt forsknings-samarbeid, med særlig vekt på EUs rammeprogrammer.

TØI dekker alle transportmidler og temaområder innen samferdsel, inkludert trafiksikkerhet, kollektivtransport, klima og miljø, reiseliv, reisevaner og reiseetterspørsel, arealplanlegging, offentlige beslutningsprosesser, næringslivets transporter og generell transportøkonomi.

Transportøkonomisk institutt krever opphavsrett til egne arbeider og legger vekt på å opptre uavhengig av oppdragsgiverne i alle faglige analyser og vurderinger.

**Besøks- og postadresse:**  
Transportøkonomisk institutt  
Gaustadalléen 21  
NO-0349 Oslo

22 57 38 00  
[toi@toi.no](mailto:toi@toi.no)  
[www.toi.no](http://www.toi.no)