CEDR Transnational Road Research Programme
Call 2013: Safety

funded by Germany, Ireland, UK
and Netherlands

CEDR
Conférence Européenne des Directeurs des Routes
Conference of European Directors of Roads

PRACT
Predicting Road ACcidents - a Transferable methodology across Europe

Development of new Crash Modification Factors/Functions per key safety treatments

Deliverable D2

July, 2015

Università degli Studi Firenze DICEA Dipartimento di Ingegneria Civile e Ambientale

Technische Universität Berlin

Imperial College London

With the support of:

autostrade per l'italia

IRF

European Union Road Federation ERF

TRB ANB25 Committee

CEDR
Conférence Européenne des Directeurs des Routes
Conference of European Directors of Roads

CEDR Call 2013: Safety
PRACT
Predicting Road ACcidents - a Transferable methodology across Europe

Development of new Crash Modification Factors/Functions per key safety treatments

Due date of deliverable: 31/07/2015
Actual submission date: 1/08/2015

Start date of project: 01/04/2014 End date of project: 31/03/2016

Authors of this deliverable:
Niovi Karathodorou, Daniel Graham, Jun (Simon) Hu ICL
Thomas Richter, Stephan Ruhl, TUB
George Yannis, Anastasios Dragomanovits, Alexandra Laiou, NTUA
Francesca La Torre, Lorenzo Domenichini UNIFI

PEB Project Manager: Elizabeth Mathie, Highways Agency, UK

Version: Final, 02.2016

TABLE OF CONTENTS

List of Figures

List of Tables

Glossary of Terms

- **Road Safety Measures / Treatments / Countermeasures / Interventions**: any modifications in road design, maintenance and equipment, traffic control, vehicle design, inspection and protective devices, driver training, public education, enforcement and post-accident care, that aim at reducing accident frequency or severity;

- **Crash Modification Factor (CMF)** or Function, or Accident Modification Factor: the relative change in accident frequency due to a change in one specific condition (when all other conditions and site characteristics remain constant). CMF is the ratio of the expected accident frequency after a modification or measure is implemented to the estimated accident frequency if the change does not take place;

Executive Summary

The PRACT Project (Predicting Road ACcidents - a Transferable methodology across Europe) aims at developing a European accident prediction model structure that could be applied to different European road networks with appropriate calibration. PRACT is funded by the National Road Authorities of Germany, Ireland, UK and Netherlands within the Conference of European Directors of Roads (CEDR) 2013 Transnational Research Programme - Safety.

The research partners of the PRACT project are:

- Università degli Studi di Firenze (Italy) - Project Leader,

- National Technical University of Athens (Greece),

- Technische Universität Berlin (Germany), and

- Imperial College London (UK).

The key aim of the project is to develop a procedure that will enable Accident Prediction Models (APMs) and Crash Modification Factors (CMFs) to be transferred to conditions different to the conditions for which they were developed. This will be implemented into a practical guideline and a user friendly tool that will allow the different road administrations to identify suitable APMs and CMFs and adapt them for use in local conditions. The project focuses in particular on motorways and two-way two-lane rural roads.

CMFs are indicators that quantify the reduction in accident rates resulting from safety treatments. They are an important tool used by road agencies and researchers to better understand the effectiveness of safety interventions and form the basis for evidence based safety policies. CMFs allow the estimation of safety benefits in economic analyses of safety policies and hence can enable optimal use of resources.

The PRACT project comprises 5 work packages. Work package 2 (WP2) focuses on reviewing and identifying gaps in the existing literature on CMFs. The specific objectives of the second work package of the project (WP2) are:

- To review the recent and salient literature on CMFs, including the background and development of the CMF and various approaches for developing CMFs. Results of the review will be used to develop an online web repository of CMFs that will accompany the transferability tool developed as part of the project (WP3, WP4).
- To identify key CMFs which have not been fully studied or have been omitted in the literature.
- If possible, to develop new missing CMFs.

Deliverable D2 presents the results of the CMF development undertaken as part of the WP2 work package of the PRACT project. The first section details the CMFs that were estimated as part of the PRACT project. Chapter 2 presents the methodologies used. Chapter 3 describes the corresponding data sources. Results are presented and discussed in Chapter 4 before some concluding remarks in Chapter 5.

CMFs for the following countermeasures were estimated within PRACT: the presence of a work zone, average speed enforcement (section control) and high friction wearing course for Italian motorways; traffic composition, road width, horizontal curvature and vertical gradient for German two-way two-lane rural roads; and traffic composition, horizontal curvature and vertical gradient for English two-way two-lane rural roads.

Two distinct approaches were used to estimate CMFs. The choice of methodology for each CMF depended on the type of CMF to be estimated as well as data availability. The Empirical Bayes Before-After (EB) method was used to estimate the effect of work zones, high friction wearing courses and average speed enforcement (section control) on accident rates on rural motorways based on Italian data. The advantage of the EB approach is that it controls for the effects of regression to the mean, which arise from the fact that treatments tend to be implemented at accident blackspots. On the other hand, CMFs for England and Germany were derived by estimating Negative Binomial models. The methodology can still provide unbiased estimates of the effect of traffic composition, road width, horizontal curvature and vertical gradient on accident rates as such features are unlikely to depend on accident rates. A comprehensive set of explanatory variables was included in the models to avoid omitted variable bias.

Table 1: *Summary of CMFs estimated within PRACT (F=Fatal, I=Injury, ROR=Run-off-Road)*

Countermeasure/ feature	Country	Methodology	Value/ function	Injury Type	Crash Type
Presence of a work zone	Italy	Empirical-Bayes	0.84-3.11 depending on work zone layout (1.33 on average)	F+I	ALL
Speed enforcement (section control)	Italy	Empirical-Bayes	0.52-1.55 depending on injury & crash type Insignificant effect for some injury & crash types	Various	Various
High friction wearing course	Italy	Empirical-Bayes	0.27	F+I	ROR, wet pavement
Horizontal curvature	England	Negative Binomial	no significant effect	F+I	ALL
Vertical gradient (V)	England	Negative Binomial	CMF = $e^{0.09*\Delta V}$	F+I	ALL
% HGV (HGV)	England	Negative Binomial	CMF = $e^{-7.58*\Delta HGV}$	F+I	ALL
% two wheel traffic	England	Negative Binomial	no significant effect	F+I	ALL
Road width	Germany	Negative Binomial	CMF = $e^{-0.17*\Delta RW}$	F+I	ALL
Horizontal curvature (HC)	Germany	Negative Binomial	CMF = $e^{0.003*\Delta HC}$	F+I	ALL
Vertical gradient (V)	Germany	Negative Binomial	no significant effect	F+I	ALL
% HGV (HGV)	Germany	Negative Binomial	no significant effect	F+I	ALL

1 Introduction

1.1 CMF Development

WP2 reviewed the CMF literature for 92 countermeasures or road features for motorways and two-way two-lane rural roads (more details can be found in Deliverable D4). The CMFs developed as part of the project aim to fill gaps identified during the literature review within the limits imposed by data availability issues. Selection of the CMFs to develop also took into account results from a survey of 22 National Road Authorities and other relevant institutions[1] from Europe, US and Australia that was conducted as part of the first work package of the project (WP1). In general, the estimated CMFs were identified by road agencies as valuable but not so readily available.

CMFs are estimated based on data from three countries: Italy, England and Germany. Different CMFs are developed for each country depending on data availability. Table 1 summarises the CMFs developed for each country. The remainder of this section explains in more detail the choice of CMFs to develop for each country and the contribution made to existing literature.

Table 1.1: *CMFs developed in WP2*

Country	Crash Modification Factor	Road Type
Italy	Work zones	Motorway (rural)
Italy	Average speed enforcement (section control)	Motorway (rural)
Italy	High friction wearing course	Motorway (rural)
Germany	Traffic composition	Two-way two-lane rural road
Germany	Number of lanes	Two-way two-lane rural road
Germany	Road width	Two-way two-lane rural road
Germany	Horizontal curvature	Two-way two-lane rural road
Germany	Vertical gradient	Two-way two-lane rural road
England	Traffic composition	Two-way two-lane rural road
England	Horizontal curvature	Two-way two-lane rural road
England	Vertical gradient	Two-way two-lane rural road

[1] Survey participants included mostly National Road Authorities, but also Road Managing Companies, Academia/Research Institutes and Highway Consultants.

1.2 Italy

Firstly, we estimate CMFs for **work zones for rural motorways**. Work zones are critical parts of the road network in terms of safety. The drivers have to face additional choices because of the temporary and unfamiliar road layouts; furthermore, the new manoeuvres to deal with may cause additional conflict points between the vehicle paths. Roadway work zones are hazardous, both for workers and motorists, who drive through the complex array of signs, delineators and lane changes. A study for the years 2003-2007 in Sweden (Liljegren, 2008) reported that work zone injury accidents were 0.6% of the annual injury accidents and the associated fatalities were 0.9%. Statistics from the U.S. (FHWA, 2015) count 669 killed in 2014 in road accidents within work zones, which is 2% of the total 32,645 fatalities. According to FHWA data, out of a total of 87,606 crashes recorded at work zones in 2010 in the U.S. (1.6% of the total number of roadway crashes in 2010) only 0.6% led to fatalities, whereas the 30% were injury crashes and 69% were property damage only crashes.

Very few studies exist on the topic of quantifying safety impacts of roadworks. The Highway Safety Manual (HSM) synthesized previous research (Khattak and Council, 2002) in order to provide a way to quantitatively evaluate work zone safety but this is limited to the evaluation of the effect of changing work zone length and duration (AASHTO, 2010). A preliminary study to evaluate work zone CMFs with a Naive before/after approach has been conducted in the ASAP Project (Saleh et al, 2013). However, the methodology does not account for regression to the mean and hence, results may not be reliable. In the PRACT project a set of CMFs for single vehicle and multi-vehicle fatal and injury crashes for different work zones layout are derived by means of an Empirical Bayes before-after analysis.

The survey conducted as part first work package (WP1) of the PRACT project showed that CMFs for work zones for rural motorways are useful for road agencies. 86.7% of participating road agencies responded that CMFs for work zones for motorways are highly required. Despite their usefulness, 64.3% of respondents reported that there was a low availability of CMFs relating to work zones for motorways. This is corroborated by the literature review conducted in WP2.

We also estimate CMFs for **average speed enforcement** based on the time of travel over a given segment length **on rural motorways (section control)**. Driving speed is considered to have a strong effect on crashes and crash severities but the effect of driving above the posted speed limit is often assessed indirectly by evaluating the effectiveness of enforcement methods that reduce the percentage of speeders and the speed distributions' standard deviations with respect to the mean.

A very extensive evaluation of available studies relating speed enforcement techniques with crash counts and severity can be found in Elvik et al.,2009. This evaluation shows that the effectiveness of this treatments varies considerably with the technique applied (laser, radar, section control). Only one study is available based on section control enforcement (Stefan, 2006). The study shows a potential crash reduction of 30% but it is based in a limited amount of data and therefore, results may be somewhat unreliable.

The survey conducted as part of the first work package (WP1) of the PRACT project showed that CMFs for average speed enforcement for rural motorways are useful for road agencies. 64.7% of participating road agencies responded that CMFs for speed enforcement on motorways are highly required.

In the PRACT project a set of CMFs for single vehicle and multi-vehicle fatal+injury crashes and property damage only crashes for sections with average speed enforcement and different traffic volumes are derived by means of an Empirical Bayes before after analysis.

Lastly, we estimate CMFs related to the adoption of **high friction wearing courses in rural motorways**. Several studies have proven the effectiveness of the use of high friction wearing courses: the "Departments Of Transportation" (DOTs) of Pennsylvania, Kentucky and South Carolina conducted before after studies considering 3-5 years before and after the treatment showing crash reductions of 100%, 90% and 57% respectively (FHWA, XX). Florida DOT, in 2005, obtained a reduction of 20%-30% considering all the crashes, and a reduction of 50% considering only wet pavement crashes (Gan et al, 2015). This type of intervention is more effective that a simple increase in friction in low friction situations as it has also a visual impact (being typically of a different colour as compared to asphalt concrete).

There is a lack of studies at a European level. Moreover, the survey conducted as part first work package (WP1) of the PRACT project showed that CMFs for high friction wearing course for rural motorways are useful for road agencies. 71.4% of participating road agencies responded that CMFs for high friction wearing courses for motorways are highly required. Despite their usefulness, 63.9% of respondents reported that there was a low availability of CMFs relating to the use of high friction wearing courses on motorways.

In the PRACT Project a CMFs for single vehicle **Run Off Road crashes on wet pavements** are derived by means of an Empirical-Bayes before/after analysis.

1.3 England

For England, we estimate **CMFs for traffic composition (% of heavy goods vehicles (HGV) and two-wheeled vehicles in traffic)** by developing negative binomial models for accident rates that include explanatory variables relating to traffic composition. CMFs are estimated **for two-lane two-way rural roads**. Although total traffic volume is typically included in accident prediction models, variables representing traffic composition are rarely used. Our literature review found only 5 CMF estimates relating to traffic composition for two-way two-lane rural roads (Dinu et al, 2011; Vogt and Bared, 1998), and none for rural motorways. Moreover, there appears to be a lack of estimates at the European level. Reviewed estimates are based on data from the US and India only. As CMFs may not be transferable to other countries, estimating a CMF based on UK data is pertinent.

The survey conducted in WP1 of this project showed that CMFs for traffic composition are useful for road agencies. 69.2% of participating road agencies responded that CMFs for two-way two-lane rural roads relating to traffic characteristics, including traffic composition, are highly required. Despite the apparent usefulness of CMFs on traffic characteristics, 81.8% of respondents reported that there was a low availability of such CMFs. The survey results are corroborated by the literature review conducted in WP2.

The models developed to estimate CMFs on traffic composition also allow the estimation of CMFs for **horizontal curvature and vertical gradient**. In the WP1 survey, 69.2% of road authorities suggested that curvature CMFs for two-way two-lane rural roads are needed for implementation in their network; 63.6 % indicated that few estimates were available. Vertical gradient CMFs appear to be equally valuable in particular for two-lane rural roads; 64.3% of road authorities considered them useful while 72.7% reported a lack of suitable estimates.

Given the availability of a centralised database of all pavements works in England and annual SCRIM measurements from 2005 onwards, we also investigated the possibility of estimating CMFs for skid resistance for motorways and two-way two-lane rural roads using the Empirical Bayes Before-After methodology. Such CMFs were identified both as valuable for road agencies based on the WP1 survey and as lacking in the literature. However, this was not possible as insufficient road segments were available where resurfacing works satisfying the required properties had been completed. The properties required were defined as works conducted on;

- rural motorways or two-way two lane rural roads,
- on all lanes and for the entire length of the segment;
- works happened as standalone projects, i.e. without any other major projects completed at the same time, and
- that occurred after 2005 so that SCRIM measurements are also available.

1.4 Germany

For Germany we develop CMFs on **AADT**, **traffic composition (% of heavy goods vehicles in traffic)**, **number of lanes, horizontal curvature and vertical gradient** for two-way two-lane rural roads using negative binomial models. As explained in the previous section, CMFs for these road characteristics are highly desirable and often lacking based on the results of the WP1 survey. The development of CMFs using generalised linear modelling is a current trend in Germany. The approach was first applied to a sample of the German rural road network by Maier et al (2013). Maier el al (2013) used comprehensive road design information for parts of the rural road network based on both digital road information databases (*Straßeninformationsbank – SIB*) and own surveys. The current project used a similar approach but focused on the state of Brandenburg, as explained in more detail in the data section.

2 Methodology

Two distinct approaches are used for CMF development in the project. The choice of methodology for each CMF depends on the type of CMF to be estimated as well as data availability. Data limitations mean that for some countermeasures the choice of methodologies is limited.

When data on the year/date a countermeasure was implemented, as well as data on accident rates and traffic volumes are available both for the period before and after application of the countermeasure, CMFs are developed using an Empirical Bayes Before-After (EB) approach (Hauer, 1997). The advantage of the approach is that it controls for the effects of regression to the mean. Countermeasures tend to be implemented at sites where high accident rates have been recorded. This non-random allocation of countermeasures can lead to self-selection bias, including the so-called regression to the mean (RTM) effect. The RTM effect arises because observed high accident rates may simply be due to random variation. If this is the case, they will tend to be closer to the mean value in future observations. Thus, a reduction in accident rates may be observed that is however random rather than due to the implemented countermeasure. Because of its ability to deal with RTM, the EB approach is currently widely used for CMF development (example of studies include Harkey et al 2008; Khan et al 2015; Lyon et al 2008; Park et al 2012; Patel et al 2007; Persaud et al 2004; Persaud et al 2012). In the PRACT project it is used to estimate the effect for work zones, high friction wearing courses and average speed enforcement (section control) on accident rates on rural motorways based on Italian data.

When no suitable data is available to employ the Empirical Bayes approach, multivariate regression models can also be used to estimate CMFs. The approach is useful when only cross-sectional data are available for estimation. However, it is not suitable for countermeasures that have been implemented to road segments because of high accident rates. When countermeasures have been allocated at accident blackspots, the countermeasure variable will be endogenous in the model (i.e. correlated with the error term) leading to biased estimates; more advanced modelling techniques (e.g. instrumental variables) are needed to obtain unbiased estimates of the effect of the treatment. On the other hand, multivariate regression models are suitable for CMF estimation when countermeasures are independent of accident rates (e.g. blanket treatments) and for road features that do not depend on accidents such as the number of lanes or traffic composition. Care should be taken to include a detailed set of explanatory variables in the model to avoid issues relating to omitted variable bias: Variables omitted from the model that affect accidents and are also correlated with the error term can lead to biased estimates. An advantage of using multivariate regression models for CMF estimation is that they can provide CMF estimates as a function of the countermeasure of interest. This can be helpful for countermeasures/road features that are represented by continuous variables such as the % of heavy goods vehicles in traffic.

For this project, negative binomial models are estimated using data from England and Germany to obtain CMFs for traffic composition, lane width, horizontal curvature and vertical gradient. Such features are unlikely to depend on accident rates and hence the methodology should provide unbiased estimates of their effect on accident rates. We use as detailed a specification as possible given the data available to avoid the issue of omitted variable bias.

Table 2.1 presents the methodology used for the CMFs developed in the PRACT project. More details on each methodology, including details of the model specification, can be found in Sections 2.1 and 2.2.

Table 2.1: *Methodologies employed in CMF estimation*

Country	CMF	Road Type	Methodology
Italy	Work zones	Rural motorway	Empirical Bayes Before-After
Italy	Average speed enforcement (section control)	Rural motorway	Empirical Bayes Before-After
Italy	High friction wearing course	Rural motorway	Empirical Bayes Before-After
England	Traffic composition	Two-way two-lane rural road	Negative Binomial Model
England	Horizontal curvature	Two-way two-lane rural road	Negative Binomial Model
England	Vertical gradient	Two-way two-lane rural road	Negative Binomial Model
Germany	Traffic composition	Two-way two-lane rural road	Negative Binomial Model
Germany	Lane width	Two-way two-lane rural road	Negative Binomial Model
Germany	Horizontal curvature	Two-way two-lane rural road	Negative Binomial Model
Germany	Vertical gradient	Two-way two-lane rural road	Negative Binomial Model

2.1 Empirical Bayes Before-After

Assume a countermeasure is implemented at a number of sites at some point t in time. We will refer to the sites that received the treatment as the treated sites, the time before the treatment occurred as the 'before period' and the time after the treatment occurred as the 'after period'. The EB approach estimates the expected number of accidents at the treated sites in the before period as a weighted sum of the observed number of accidents at the treated sites in the before period and the number of accidents predicted for the before period for untreated sites with similar characteristics to the treated sites. In particular,

$$N_{EXP,B} = N_{PRED,B} \cdot w + N_{OBS,B} \cdot (1 - w) \qquad (1)$$

where

$N_{EXP,B}$ is the expected number of accidents in the before period (also called the Empirical Bayes estimate of the number of accidents in the before period)

$N_{PRED,B}$ is the predicted number of accidents for reference entities that are similar to the entities in the treatment group in the before period

$N_{OBS,B}$ is the observed number of accidents in the treatment group in the before period

And w is a weight that is chosen so that $Var(N_{EXP,B})$ is minimised (Hauer, 1997).

To estimate $N_{PRED,B}$, the EB approach uses Safety Performance Functions (SPFs) that have been estimated using data from untreated reference sites with similar characteristics to the treated sites. For the evaluation of the expected crash frequencies for work zones and average speed enforcement CMFs, the SPF model proposed in the NCHRP 17-45 project (Bonneson et. al., 2012) and published in the Highway Safety Manual Supplement (AASHTO, 2014) is adopted. In particular,

$$N_{PRED,B} = N_{SPF} \cdot (CMF_1 \cdot CMF_2 \cdot \dots \cdot CMF_m) \cdot C \qquad (2)$$

where $N_{SPF} = L \cdot \exp[a + b \cdot \ln(c \cdot AADT)]$ is an accident prediction function developed for some base conditions, L is segment length and $AADT$ is Annual Average Daily Traffic. Parameter estimates from the Highway Safety Manual Supplement are used for a, b and c (AASHTO, 2014, 2010)

CMF_1, ..., CMF_m are CMFs that reflect how reference sites vary compared to the base conditions for which N_{SPF} was developed. CMFs for the following road characteristics are used:

- Horizontal curve;
- Lane width;
- Inside shoulder width;
- Median width,
- Median barrier;
- High Volume;
- Lane change;
- Outside shoulder width;
- Outside clearance;
- Outside barrier

A reference group of sites with similar characteristics to the treated sites is used in order to estimate the calibration factor C. C is defined as the ratio of the total accidents observed on the reference sites in the before period to the total number of accidents predicted for all reference sites by the model. To adjust for time, the model (equation 2) is calibrated separately for each year of observation on the reference dataset (sections without the treatment).

For the evaluation of the expected crash frequencies for the high friction wearing courses CMF, the model recently developed within the SAVeRS project (La Torre et al., 2015) to estimate Run Off Road (ROR) crashes is adopted. The model takes the same form as equation (2) with the exception that N_{SPF} takes the following form:

$$N_{SPF} = \mathrm{L} \cdot e^{\beta_0 + \beta_1 \log(AADT)}$$

The CMFs used are:

- number of lanes;
- outside shoulder width (shoulder adjacent to the slow moving traffic);
- inside shoulder width (shoulder adjacent to the median);
- longitudinal gradient;
- shoulder rumble strips;
- lane width;
- horizontal curvature.

The expected number of accidents that would have occurred in the after period in the absence of a treatment ($N_{EXP,A}$) can be estimated as

$$N_{EXP,A} = N_{EXP,B} \cdot \frac{N_{PRED,A}}{N_{PRED,B}}$$

$N_{PRED,A}$ is the predicted number of accidents for similar entities as the entities in the treatment group in the after period in the absence of the treatment. Then,

$$Var(N_{EXP,A}) = N_{EXP,A} \cdot \frac{N_{PRED,A}}{N_{PRED,B}} \cdot (1 - w)$$

And the CMF estimate for the treatment under consideration is given by:

$$\widehat{CMF} = \frac{\sum_i N_{OBS,A\,i} / \sum_i N_{EXP,A\,i}}{1 + \sum_i Var(N_{EXP,A\,i}) / (\sum_i N_{EXP,A\,i)}{}^2}$$

The variance of the CMF estimate is given by

$$Var(\widehat{CMF}) = \frac{\widehat{CMF}^2 \cdot [1/\sum_i N_{OBS,A\,i} + \sum_i Var(N_{EXP,A\,i}) / (\sum_i N_{EXP,A\,i)}{}^2}{1 + \sum_i Var(N_{EXP,A\,i}) / (\sum_i N_{EXP,A\,i)}{}^2}$$

95% confidence intervals can be derived as follows:

$$\widehat{CMF} \pm 1.96 \cdot \sqrt{Var(\widehat{CMF})}$$

2.2 Negative Binomial Models

We model accident occurrence as the number of accidents y occurring in a road segment i using a negative binomial model. The Negative Binomial model is preferred over the Poisson model because it allows for overdispersion (i.e. $Var(y) > E(y)$) in contrast with the Poisson model which restricts $Var(y) = E(y)$. This latter assumption is a rather restrictive assumption as real data tend to be overdispersed. The Negative Binomial distribution is widely used for modelling accidents rates.

The probability of y accidents occurring in a year given a parameter μ is defined as

$$P(y/\mu) = \frac{e^{-\mu} \mu^y}{y!}$$

Where the expected number of accidents μ is given by

$$\mu = E(y/\boldsymbol{x}, \alpha) = \exp[\beta_0 + \ln(L) + \beta_{AADT}\ln(AADT) + \sum x_i\,\beta_i + t]v,$$

and

v is an iid[2] random variable that follows a $Gamma(\frac{1}{a}, a)$ distribution, where a is a parameter to be estimated $(a > 0)$[3]

L is segment length (metres)

AADT is traffic volume (Average Annual Daily Traffic – veh/day)

$\boldsymbol{x} = (x_1, \dots, x_i, ..)$ is a vector of factors affecting y such as lane width, the % of heavy vehicles, horizontal curvature and vertical gradient

t is a variable representing time trend if data for multiple years are available for estimation

and β_0, β_L, β_{AADT} and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_i, ..)$ are a set of parameters to be estimated

Then,

$$Var(y/\boldsymbol{x}, \alpha) = \mu + \mu\alpha^2$$

CMFs derived from a negative binomial model can be expressed as a function of the variable of interest. For the model described above, the CMFunction for variable x_i, the ith component of the vector $\boldsymbol{x}$, is given by

$$e^{\beta_i\,\Delta x_i}$$

where $e^{\Delta x_i}$ is the change in x_i.

2.2.1 Model for England

A negative binomial model is used to estimate CMFs for two-way two-lane rural roads. In addition to annual average daily traffic ($AADT$) and segment length (L), the model includes the following explanatory variables:

- Average horizontal radius of curvature HC defined as

$$HC = \frac{\sum_i HC_i L_i}{L}$$

where

HC_i is the horizontal radius of curvature of element i of the road segment [m]

and L_i is the length of element i of the road segment [m]

---

[2] iii = independently identically distributed

[3] For a Poisson model v = 1

- Vertical curvature VC [%] defines as

$$VC = \frac{\sum |s_i| * L_i}{L}$$

where

s_i is the % gradient of element i of the road segment

L_i is the length of element i of the road segment [m]

L is the length of the road section [m]

- % of traffic that are Heavy Goods Vehicles (HGV)
- % of traffic that is two-wheeled (i.e. cycles and motorcycles)

Binary dummy variables are also included in the model to represent different years as a panel dataset is used for estimation[4]. We also considered mean road segments length, but as the available data are unreliable the variable is excluded from the final model.

2.2.2 Model for Germany

A negative binomial model is used to estimate CMFs for two-way two-lane rural roads. In addition to annual average daily traffic ($AADT$) and segment length (L), the model includes the following explanatory variables:

- Road width
- Horizontal curvature
- Vertical gradient
- % of heavy vehicles

The model is estimated based on cross-sectional data, so no variables representing time are required. Variable construction is described below.

Road width

A comprehensive database on cross-section information can be used for this road design element. The road information database contains new entries for road width when road width changes by as little as 5 cm. For CMF development the mean road width for a road section is calculated as a weighted average (considering the length of subdivided road segments) for the different road widths of every segment. If large changes in width are identified for a section, the section is not included in the dataset for CMF development.

Horizontal curvature

For horizontal road design the parameters for all horizontal elements (straights, curves, clothoids) are available with their radii and respective length, so that curvature can be estimated. The horizontal curvature is defined as the sum of the angles of changes in directions per road section divided by the length of the road section (see also Maier et al, 2013). The formula is provided below.

[4] Dummy binary variables are included for years 2013, 2012 and 2011. The reference year is 2010.

$$HC = \frac{\sum \propto_i}{L}$$

where

HC = Horizontal curvature [gon/km]

α = Angle of changes in direction per element i [gon]

L = Length of road section [km]

Vertical gradient

The information on vertical road design is based on spot heights so that vertical gradient can be estimated. The vertical gradient is defined as the mean of absolute grades of a road section (regardless of the driving direction) (Maier et al, 2013). The respective formula is provided below.

$$VC = \frac{\sum |s_i| * L_i}{L}$$

where

VC = Vertical gradient [%]

s = Grade of element i [%]

L_i = Length of element i [km]

L = Length of road section [km]

In addition to the variables described above, several other variables can potentially have an effect on rural road safety; however, the necessary data for their estimation is not available from the road authorities. Such variables include sight distance, sight obstacles and traffic regulations (e.g. speed limits, restrictions on overtaking). Unfortunately, estimating these variables requires a comprehensive data survey and data preparation, something which is not possible during the PRACT project. Such surveys are typically limited to specialised research projects with a focus on the above mentioned rural road design characteristics.

3 Data

3.1 Italy

3.1.1 Data on road network characteristics

The study used data from the Italian motorway network managed by Autostrade per l'Italia (ASPI) to build a comprehensive dataset. The company manages about 3000 km of motorways throughout Italy.

The motorway segments database contains details about the roadway characteristics of about 2100 km of motorways' carriageways (each segment has two carriageways) such as horizontal curve, lane width, gradient, inside and outside shoulder width, median width, median and outside barrier. Some of the data had to be collected within the PRACT project as the official database does not contain the required data (e.g. lane and shoulder widths).

Friction and roughness data are also available for all the ASPI network but these will not be used in this project. For "high friction" wearing course the CMF is based on the type of surface and not on the actual measured friction value.

3.1.2 Accident data

The accidents database contains details about approximately 105,000 crashes occurred on the motorway segments from January 1, 2007 through December 31, 2012.

For each accident several details are provided such as date, hour, localization on the motorway segment, pavement and weather conditions, number of casualties, severity of casualties, number and type of vehicles involved (passenger car/heavy vehicle), weather conditions and road surface conditions (e.g. dry, wet), presence of a work zone.

3.1.3 Traffic data

Traffic data are given per each homogeneous segment (between two interchanges) in terms of daily traffic for each day of the year for each vehicle class.

Section data are also available for approximately 60 measuring stations. In these stations traffic counts and speed distribution are given per each hour of the year. This information is used in PRACT only to define the CMFs for high traffic volumes.

3.1.4 Work zones data

More than 30,000 stationary work zones were installed on the motorway network from January 1, 2007 through December 31, 2012. For each work zone, details about the exact positioning on the motorway, starting and ending date, the signalling and further details about which lane is closed to the traffic flow are provided. However each row of the database may not correspond necessarily to a single work zone, but very often refers to different working phases. For this reason it was necessary to group the rows referred to a single work zone in order to define the actual number of work zones.

Each work zone is associated to one of the stationary layout schemes, defined according to the Italian Ministerial Decree 10 July 2002 (Ministero Infrastrutture e Trasporti, 2002). A description of the work zones layout relevant for the PRACT project is given in Table 3.1.

Table 3.1: Work zone layouts considered

Code name	Work zone layout description
TWO-LANE CARRIEGEWAY	
Emergency2	Closure of emergency lane (outside paved shoulder)
Slow2	Closure of slow lane with traffic diverted to overtaking lane
Fast2	Closure of overtaking lane with traffic diverted to slow lane
Fast2(2)	Closure of overtaking lane with traffic diverted to slow & emergency lanes
Cross2(0+1)	Closure of slow lane with traffic diverted to overtaking lane; closure of overtaking lane and total diversion of traffic to the opposite carriageway through a single-lane crossover
Cross2(1+1)	Closure of slow lane with traffic diverted to overtaking lane; partial diversion of traffic to the opposite carriageway through a single-lane crossover (the driver is allowed to choose whether to stay on the overtaking lane or move to the opposite carriageway)
THREE-LANE CARRIEGEWAY	
Emergency3	Closure of emergency lane (outside paved shoulder)
Slow3	Closure of slow lane with traffic diverted to middle lane
Slow&Middle3	Closure of slow lane with traffic diverted to middle lane; closure of middle lane with traffic diverted to overtaking lane
Fast3	Closure of overtaking lane with traffic diverted to middle lane
Middle&Fast3	Closure of overtaking lane with traffic diverted to middle lane; closure of middle lane with traffic diverted to slow lane
Fast3(3)	Closure of overtaking lane with traffic diverted to middle, slow & emergency lanes
Middle&Fast3(2)	Closure of overtaking lane with traffic diverted to middle lane; closure of middle lane with traffic diverted to slow and emergency lanes
Cross3(1+1)	Closure of overtaking with traffic diverted to middle lane; closure of middle lane and partial diversion of traffic to slow lane & opposite carriageway through a single-lane crossover (driver can choose between slow lane and opposite carriageway)
Cross3(0+1)	Closure of slow lane with traffic diverted to middle lane; closure of middle lane with traffic diverted to overtaking lane and total diversion of traffic to the opposite carriageway through a single-lane crossover
Cross3(0+2)	Closure of slow lane with traffic diverted to middle lane; closure of carriageway and total diversion of traffic to the opposite side through a dual-lane crossover

3.2 England

Models are estimated using a panel dataset of 480 observations. The data are for 120 road segments observed over 4 years (2010-2013).

3.2.1 Data on road network characteristics

Data on road characteristics for roads maintained by Highways England[5] are available from the Highways Agency Pavement Management System (HAPMS) administered by Highways England. Highways England is a government owned company that is responsible for operating, maintaining and improving England's motorways and major A roads – a total of 4,300 miles of road carrying approximately a third of England's traffic by mileage. Models are estimated using a sample of 120 two-way two-lane rural road segments covering a total of approximately 661 km. The sample contains road segments from the main carriageway only (i.e. slip roads and intersections are excluded). Road segments vary in length from 68m to 2.8 km. Mean length is 1.4 km. The sample used for estimation contained 480 observations.

3.2.2 Accident data

Accident data are obtained from the STATS19 dataset. The dataset is based on personal injury accidents occurring on public roads that are reported to the police, and hence do not include data on property damage only accidents. The data includes accident location in terms of Easting and Northing coordinates as well as a number of parameters relating to each accident such date, time, number of casualties, severity of casualties, number of vehicles involved, weather conditions and road surface conditions (e.g. dry, wet). Data are available from 1979 until 2013. For the PRACT project, data from 2010 to 2013 were used. The final sample used in estimation contained 722 accidents.

3.2.3 Traffic data

Annual street level traffic counts for the years 2010 to 2013 are available from the UK Department from Transport. The data are available online[6]. Annual Average Daily Traffic (AADT) is available for every junction to junction link on the major (motorways and A-class roads) road network of England. AADT is the mean value over an entire year of the number of vehicles passing a fixed location on the road network every day. The exact location of traffic counters is available in terms of Easting and Northing coordinates so their precise location can be mapped and linked to HAPMS road segments. Traffic counts figures include separate counts for specific type of road users: cycles, motorbikes, buses &coaches, cars & taxis, Light Goods Vehicles and different categories of Heavy Good Vehicles. Traffic levels for the road segments used in estimation ranged from 5,138 vehicles per day to 36,334 vehicles per day.

---

[5] Formerly known as Highway Agency

[6] http://data.gov.uk/dataset/gb-road-traffic-counts (accessed 12/06/2015)

3.3 *Germany*

3.3.1 Data on road network characteristics and traffic data

Germany is subdivided in 16 federal states; thereof 3 city states (Berlin, Bremen, Hamburg which were not relevant as urban areas for PRACT project) and 13 area states (see Figure 3.1). All federal states are encouraged by the German Transport Ministry to gather, store and update road data in a digital road information database (*Straßeninformationsbank – SIB*). However, due to the cost of data surveys, data availability levels between federal states vary. As a result, the project team chose to concentrate on one federal state – Brandenburg- with comprehensive data availability for the present CMF development. The geographic coverage of the data used in the project is highlighted in Figure 3.1. Data on road network characteristics and accident numbers were obtained from Brandenburg state's road authority (*Landesbetrieb Straßenwesen Brandenburg*). The road network data were available with qualifying date December 2014.

The road information data are stored for the classified road network based on a node system together with the road chainage. Road sections are limited by the both bordering road intersections and precisely defined by their bijective node numbering.

The Brandenburg data generally have different data sources. For newer roads digital planning data are integrated in the digital road information database (*Straßeninformationsbank – SIB*); for older existing roads the data comes from data surveys and subsequent roadside inspections.

Data are generally available both for motorways and the rural road network. For PRACT, CMFs are developed for 2-lane 2-way rural roads.

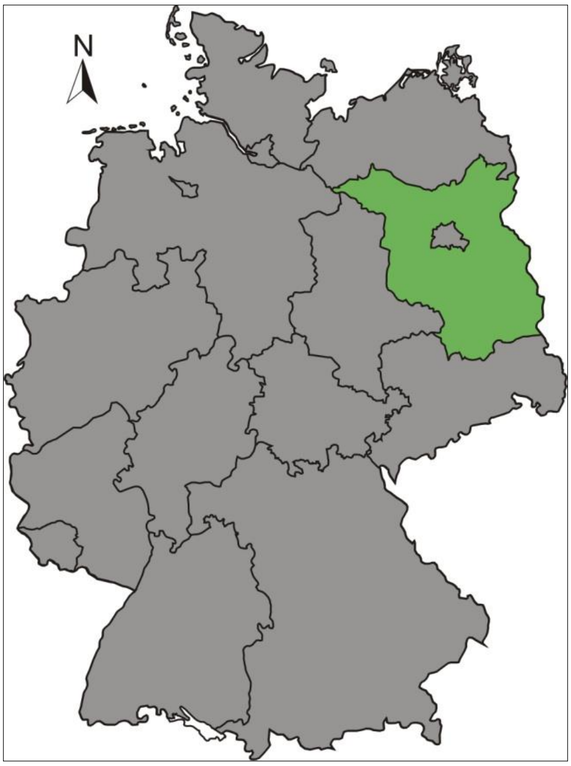

Figure 3.1: Area of data origin in Germany

The road network examined in the project contains every two-way two-lane rural road section in Brandenburg state. The initial number of available sites is 4,311 with an overall length of 8,277 kilometres (before data preparation) and includes all road categories (*Bundesstraßen, Landesstraßen, Kreisstraßen*). As the presence of intersections can affect accident occurrence and driver behaviour, road segments near intersections are excluded from the final sample used in estimation. In particular, road sections within 300 metres of a junction are not considered in the analysis as suggested by Maier et al (2013). The 300 metres comprise a 50 metres junction area and a 250 metres approach area. Several road sections are excluded from the final sample due to incomplete data for some of the explanatory variables included in the analysis. Moreover, due to the potentially distortionary effect of short road sections in accident analysis (minimum section length in the Brandenburg accident database is approximately 54 metres), road sections that are shorter than 500m are also excluded from the final sample (see Vieten et.al., 2010). The final sample used for CMF development contains 1,259 road sections that cover a total length of 3,175 kilometres. The length of the road sections in the sample ranges from 503 metres to 16.6 km. Annual average daily traffic ranges from 172 vehicles/day to 20,732 vehicles/day.

3.3.2 Accident data

The accident data required was requested from Brandenburg Police. The data covers the timeframe from 2010 to 2014. In this period, a total of approximately 411,000 were recorded in the entire Brandenburg area (urban roads, rural roads and freeways). The final sample used for CMF development includes 3,810 accidents involving a fatality or injury.

In general, accident databases in Germany contain accident information for all accident severity levels (excluding accidents not reported to the police). Six distinct accident categories are used (determined by the most serious consequence of an accident):

1. accidents with fatalities (death of at least 1 person within 30 days of the accident)
2. accidents with serious injury (minimum 24 h stationary medical care for at least 1 person involved)
3. accidents with minor injury (temporary ambulant medical care for at least 1 person involved)
4. accidents with serious material damage (minimum 1 vehicle not in working condition)
5. accidents with material damage, driver not intoxicated (all vehicles in working condition)
6. accidents with material damage, driver intoxicated (all vehicles are in working condition)

Generally the validity of accident data decreases with decreasing accident severity, so most accident studies focus on the most serious accident categories (1-3 or 1-4). For PRACT, the project team chose to consider accidents including a fatality or injury (i.e. categories 1 – 3 above). Finally, the following accident data are available for CMF development:

- period of time available: 2010 - 2014
- period of time used in estimation: 2010 - 2014
- number of accidents on rural roads: 108,777 (all accident categories)
  10,668 (fatal and injury accidents)

- number of accident in final sample: 41,770 (all accident categories)
  3,810 (fatal and injury accidents)

The final road network sample used for estimation iss obtained using the following criteria:

- focus on two-way two-lane rural roads (*Bundesstraßen, Landesstraßen, Kreisstraßen)*
- use of road sections with complete road design information
- non-consideration of intersection areas (300 m around the axis crossing)
- definition of a minimum road segment length

Because of road sections with no accidents involving fatalities or injuries, the final sample used in estimation includes 949 road sections. Nevertheless, the descriptive statistics regarding the road network that were provided in section 3.3.1 remain valid.

4 Results

Sections 4.1, 4.2 and 4.3 present and discuss the results for CMFs estimated for Italy, England and Germany respectively.

4.1 Italy

Table 4.1 summarises results for work zones CMFs for motorways in Italy. Significant estimates (5% level) are shown in bold. The layouts considered are defined in chapter 3. CMFs are for all fatal and injury crashes. Table 4.2 summarises results for speed enforcement (section control) CMFs. Significant estimates (5% level) are shown in bold. Separate results have been estimated for different traffic levels. Table 4.3 summarises results for high friction wearing course CMFs. Significant estimates (5% level) are shown in bold. The estimated CMF is for Run off Road crashes.

Table 4.1.1: *Estimated work zone CMFs*

Layout	CMF	Std. Dev. (CMF)	95% Confidence Interval (lower limit)	95% Confidence Interval (upper limit)
Cross2(1+1)	**3.11**	0.56	2.01	4.22
Cross3(1+1)	**2.80**	0.52	1.79	3.81
Cross3(0+1)	**2.15**	0.52	1.13	3.16
Cross2(0+1)	**2.08**	0.09	1.90	2.27
Slow&Middle3	1.91	0.56	0.81	3.01
Middle&Fast3	1.90	0.71	0.52	3.29
Fast2(2)	**1.64**	0.08	1.49	1.80
Slow2	**1.62**	0.12	1.39	1.85
Fast3(3)	**1.51**	0.18	1.16	1.87
Fast3	**1.49**	0.10	1.29	1.70
Emergency2	**1.27**	0.04	1.20	1.35
Cross3(0+2)	**1.25**	0.10	1.05	1.44
Fast2	1.08	0.06	0.95	1.20
Slow3	1.03	0.05	0.93	1.13
Emergency3	1.00	0.04	0.92	1.07
Middle&Fast3(2)	**0.84**	0.07	0.70	0.99
All	**1.33**	**0.02**	**1.30**	**1.37**

Table 4.1.2: *Estimated speed section control CMFs (PDO=Property Damage Only)*

Traffic level	CMF	Std. Dev. (CMF)	95% Confidence Interval (lower limit)	95% Confidence Interval (upper limit)	Type of crash + injury
ALL AADTs	0.98	0.05	0.88	1.08	Single vehicle Fatal+Injury
	0.92	0.03	0.86	0.97	Single vehicle PDO
	0.87	0.03	0.80	0.93	Multiple vehicle Fatal+Injury
	0.91	0.03	0.85	0.96	Multiple vehicle PDO
AADT < 25000	0.99	0.10	0.79	1.18	Single vehicle Fatal+Injury
	1.06	0.05	0.95	1.17	Single vehicle PDO
	1.12	0.10	0.93	1.325	Multiple vehicle Fatal+Injury
	1.55	0.10	1.36	1.74	Multiple vehicle PDO
25000 ≤ AADT < 40000	1.02	0.08	0.86	1.17	Single vehicle Fatal+Injury
	0.86	0.04	0.77	0.93	Single vehicle PDO
	0.88	0.05	0.78	0.98	Multiple vehicle Fatal+Injury
	0.88	0.04	0.8	0.96	Multiple vehicle PDO
40000 ≤ AADT < 55000	0.95	0.11	0.73	1.17	Single vehicle Fatal+Injury
	0.85	0.05	0.75	0.96	Single vehicle PDO
	0.81	0.06	0.67	0.93	Multiple vehicle Fatal+Injury
	0.75	0.05	0.66	0.84	Multiple vehicle PDO
AADT ≥ 55000	0.76	0.19	0.38	1.14	Single vehicle Fatal+Injury
	0.89	0.11	0.67	1.11	Single vehicle PDO
	0.60	0.09	0.43	0.78	Multiple vehicle Fatal+Injury
	0.52	0.07	0.38	0.66	Multiple vehicle PDO

__Table 4.1.3:__ Estimated high friction wearing course CMFs

CMF	Standard Deviation	95% Confidence Interval (lower limit)	95% Confidence Interval (upper limit)
0.27	0.10	0.08	0.47

4.2 England

The final accident prediction model for all personal injury accidents on two-way two-lane rural roads in England is given by:

$$AF = L * AADT^{0.46} * e^{-10.68} * e^{0.09*VC} * e^{-7.58*HGV}$$

where

$AADT$ is annual average daily traffic [veh/day]

VC is horizontal gradient [%] (see chapter 2 for variable definition)

and HGV is the % of traffic that are Heavy Goods Vehicles

The model includes only variables that were significant at the 5% level. Table 4.4 presents parameter estimates for all variables, including standard errors and significance levels.

Table 4.2.1: Model results (England)

Variable	Parameter	Standard error	p-value (5% sig. level)
Constant	**-10.68**	1.35	0.000
AADT (logarithm)	**0.46**	0.13	0.000
Horizontal curvature	-0.0001	0.00015	0.595
Vertical gradient	**0.09**	0.044	0.044
% HGV	**-7.58**	1.96	0.000
% two-wheel traffic	4.05	14.70	0.783
Year 2013	-0.06	0.13	0.637
Year 2012	0.13	0.13	0.297
Year 2011	-0.09	0.13	0.503

Model results suggest that only AADT, vertical gradient and the % of HGVs in traffic affect accident frequency. Surprisingly, horizontal curvature is found not to have an effect on accident rates. This could be because few sharp curves are included in the dataset used in estimation. The proportion of two-wheel vehicles in traffic also appears not to affect accidents. However, this is not surprising, especially since two-wheel vehicles are not common on English roads. The proportion of two-wheel vehicles in traffic on the road segments included in the dataset used in model estimation is on average 0.8%; the proportion does not exceed 2% for any road segment included in the dataset.

Time trend variables are also insignificant in the model, suggesting changes in traffic and traffic composition adequately capture time trends.

The following CMF function can be derived from the model for vertical curvature:

$$CMF = e^{0.09*\Delta VC}$$

(ΔVCdenotes the change in vertical curvature).

The function is illustrated graphically in figure 4.1. A road segment with a vertical gradient of 5% is found to have more than 50% more accidents compare to a road segment that is completely flat. The coefficient for vertical curvature in the estimated CMF function for England (1.905) is comparable to results from existing literature. The review conducted in WP2 of the PRACT found coefficients to range from 0.05 to 1.9 (Yannis et al, 2015).

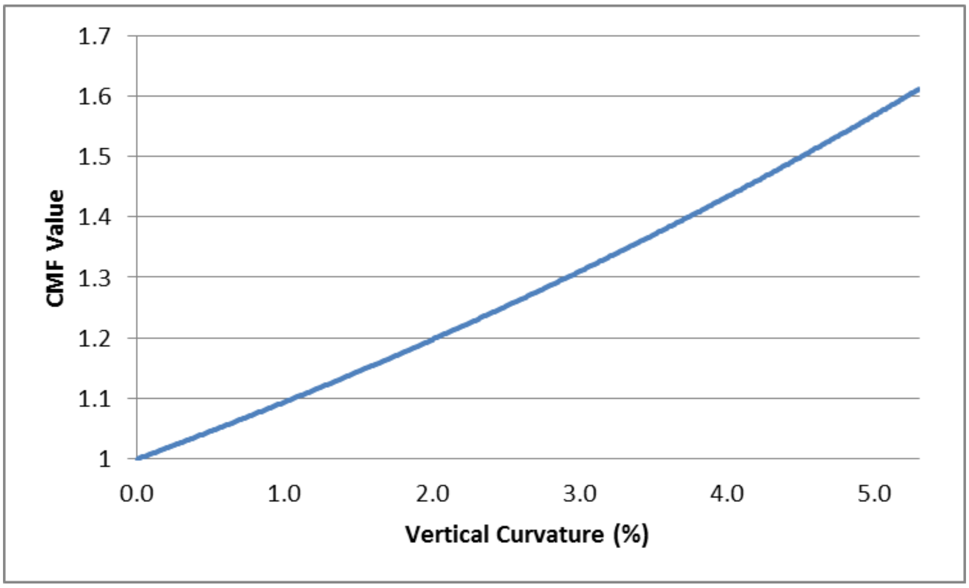

Figure 4.1: *CMF function for vertical curvature based on a base condition of 0% vertical gradient*

The following CMF function can be derived from the model for the % of HGV:

$$CMF = e^{-7.58*\Delta HGV}$$

(ΔHGV denotes the change in the proportion of HGVs).

The function is illustrated graphically in figure 4.2. The presence of HGV appears to considerably reduce accident frequency. Moreover, the effect increases significantly as their presence in traffic rises. This could be because of lower traffic speeds in roads with a high number of HGVs.

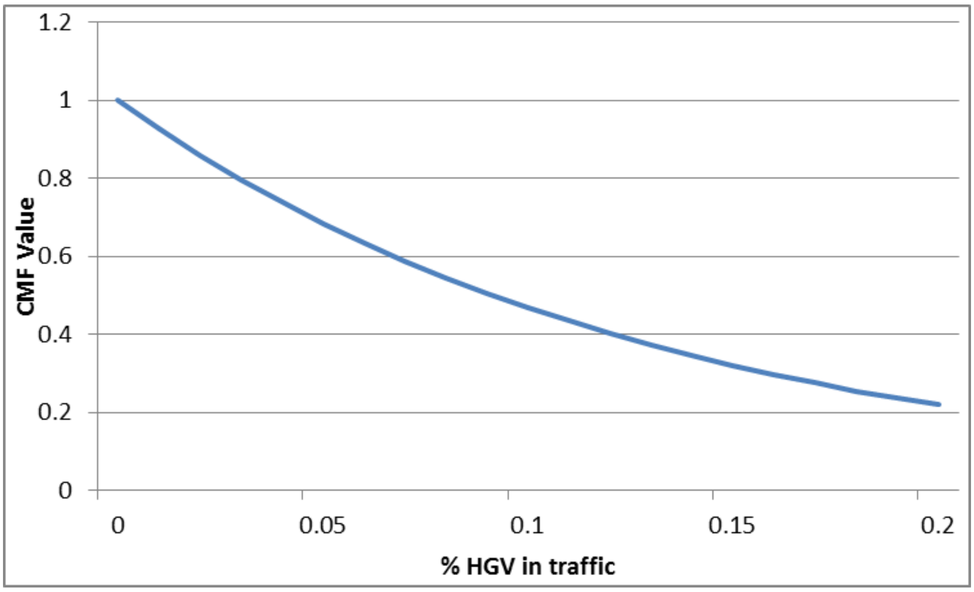

Figure 4.2: *CMF function for % of HGV in traffic based on a base condition of no HGV vehicles in traffic*

4.3 Germany

The final accident prediction model for all personal injury accidents on two-way two-lane rural roads in Germany is given by:

$$AF = L * AADT^{0.607} * e^{-5.146} * e^{-0.171*RW} * e^{0.003*CU}$$

where

AF is accident frequency [number of accidents per year]

L is segment length [km]

$AADT$ is annual average daily traffic [veh/24h]

RW is road width [m]

CU is horizontal curvature [gon/km]

The model includes only variables that are significant at the 10% level. Table 4.4 presents parameter estimates together with their standard errors and significance levels. Vertical curvature and the number of heavy vehicles were also tested in the model, but were found to be insignificant. This could suggest that the variables have no effect on accidents, i.e. that the corresponding CMFs are equal to 1. In the case of vertical curvature, however, the lack of significance could be due to insufficient variation in the explanatory variable as Brandenburg, the area used in the study, is generally flat.

Table 4.3.1: *Model results (Germany)*

Variable	Parameter	Standard error	p-value (5% sig. level)

Constant	-5.15	0.47	0.000
AADT	0.61	0.07	0.000
Road Width	-0.17	0.09	0.050
Horizontal curvature	0.00	0.00	0.064

Because of the high number of variables in the model the procedure of 'stepwise selection' is chosen, which is also useful for identifying the effects of correlated variables (Taylor et al, 2000). More details on the modelling process can be found in the appendix to the report. The correlation between any pair of explanatory variables is below 0.45, suggesting that no collinearity issues are present. Cook's distance, leverage and Q-Q plots are used to test for outliers and subsequently 28 road segments are removed from the sample used in estimation (more details can be found in appendix).

The following CMF function can be derived from the model for road width:

$$CMF = e^{-0.171*\Delta RW}$$

The following CMF function can be derived from the model for horizontal curvature:

$$CMF = e^{0.003*\Delta CU}$$

The CMF functions for lane and horizontal curvature are plotted in figures 4.3 and 4.4 respectively.

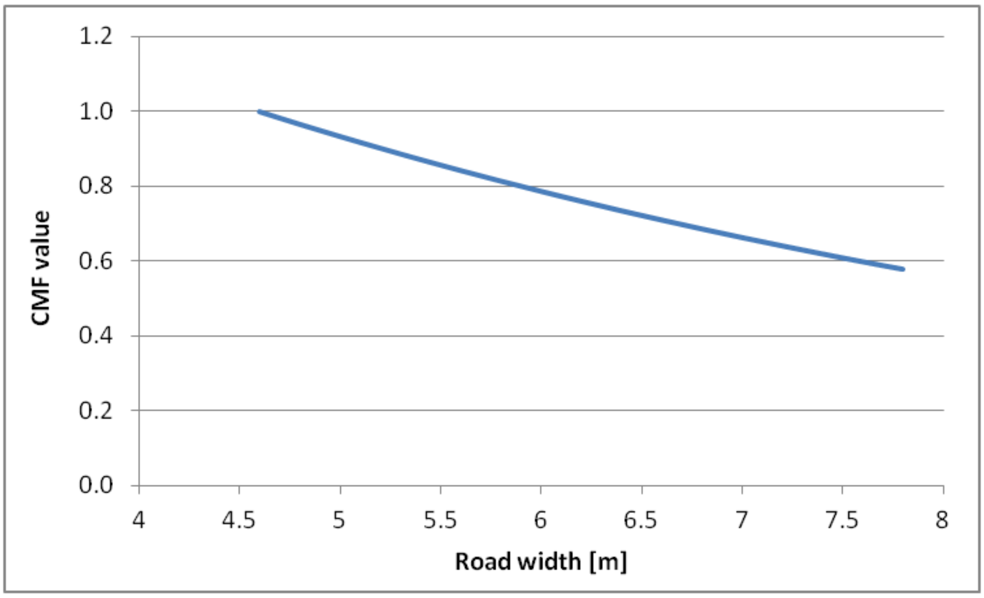

Figure 4.3: *CMF function for road width based on a base condition of 4.6 m road width (2 lanes)*

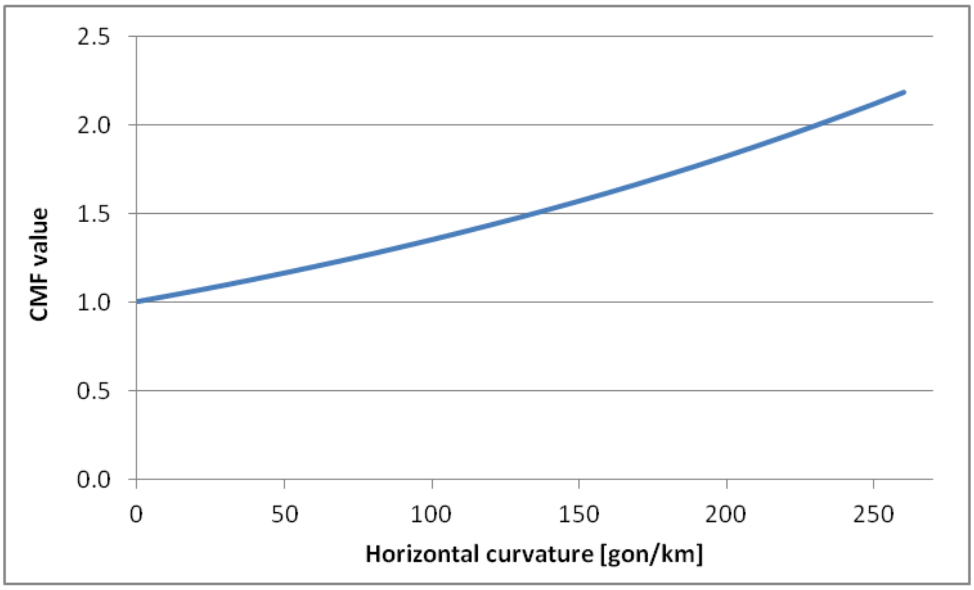

Figure 4.4: *CMF function for horizontal curvature based on a base condition of 0 gon/km*

It is recognizable, that the accident occurrence decreases with increasing road width (this is similar to safety improvements). Moreover the accident occurrence increases with increasing horizontal curvature (this is similar to a safety declining). These tendencies are generally comparable with the results of Leutzbach and Zoellmer (1989), Lamm et al (1999) and Maier at al (2013).

4.3.1 Discussion of potential errors and biases

Despite the quality of the models and their parameter estimation and the expected results of the considered CMFs, a discussion of potential errors and biases is necessary to demonstrate the potential for further improvements. Below, some limitations of our approach and data used are discussed.

A more differentiated consideration of potential accident types and their different road geometric reasons could improve the modeling. Generally every accident can have other reasons and also different overlapping reasons, which are all considered together in the approach used. This can cause additional variance and more complex coherences to the considered road design. For example the effects of road width on different accident types can also vary. Reduced road width may reduce driver tendency to overtake, which may turn in reducing the number of overtaking accidents, whereas. the accident occurrence of accidents with oncoming traffic (without overtaking) can increase with decreasing road width because of the decreasing space for manoeuvring, which is exactly the opposite tendency.

Moreover the omission of differentiated road features can lead to biases. In this approach the whole road network of Brandenburg was considered without focus on a special road type/category or detailed road information, for which surveys would be necessary. For example the different configurations of traffic regulations (e.g. speed limits or restrictions on overtaking) have different effects on driving behaviour and also accident occurrence. Whereas the restrictions on overtaking mostly affect overtaking manoeuvres/accidents, speed limits can affect the accident occurrence of different accident types. Such missing road information and differentiations of road design features can lead to unexplained variance in a model and can therefore lead to rejected variables (because of insignificant model effects) or imprecise parameter estimation.

Finally the effect of individual curves and especially the sequence of different curve radii on a road section (radii relations) can affect road safety primarily. Due to consideration of such defined variables the coherences between horizontal curvature and road safety maybe would be stronger than the horizontal curvature of a whole road section like in this approach.

5 Conclusion

Deliverable D2 reports on the CMF development undertaken as part of WP2 of the PRACT project. Development of CMFs followed a critical review of existing literature on 92 countermeasures or road characteristics. CMF estimates were developed for England, Germany and Italy.

An important finding of the project is that although gaps exist in the CMF literature, these are difficult to fill due to a lack of suitable data. Data availability issues also limit methodological choices for CMF estimations. Although the use of Empirical Bayes analyses was preferred for the PRACT project, this was not always possible due to lack of data.

CMF values estimated using Empirical Bayes Before-After analysis are summarised in Table 5.1. Results that are significant at the 5% level are shown in bold. CMF functions obtained by estimating negative binomial models for accident predictions are summarized in table 5.2. It should be noted that results from negative binomial models should be treated with caution, as discussed in Chapter 2.

Results in table 5.2 are obtained from two similarly specified negative binomial models that were estimated on data from England and Germany. It can be seen that, with the exception of AADT, results obtained from the two models are not comparable. For instance, model results suggest that horizontal curvature has a positive impact on accident rates in Germany but not in England. On the other hand, the presence of HGVs in traffic appears to substantially decrease accident rates in England but to have no effect in Germany. These differences could be due to slight differences in model specification, including differences in variable construction and definition, or to differences in the data used in estimation. However, they may also illustrate that CMFs may not be transferable between countries.

Table 5.1: *CMF values estimated for the PRACT project*

[F=Fatal, I=Injury, PDO=Property Damage Only, SV= Single Vehicle, MV=Multiple Vehicle, ROR=Run-off-Road – Significant estimates shown in bold]

Countermeasure description	Country	Value	Standard error	Crash severity	Crash types
Work zone layout Cross2(1+1)	Italy	**3.11**	0.56	F+I	ALL
Work zone layout Cross3(1+1)	Italy	**2.8**	0.52	F+I	ALL
Work zone layout Cross3(0+1)	Italy	**2.15**	0.52	F+I	ALL
Work zone layout Cross2(0+1)	Italy	**2.08**	0.09	F+I	ALL
Work zone layout Slow&Middle3	Italy	1.91	0.56	F+I	ALL
Work zone layout Middle&Fast3	Italy	1.9	0.71	F+I	ALL
Work zone layout Fast2(2)	Italy	**1.64**	0.08	F+I	ALL
Work zone layout Slow2	Italy	**1.62**	0.12	F+I	ALL
Work zone layout Fast3(3)	Italy	**1.51**	0.18	F+I	ALL
Work zone layout Fast3	Italy	**1.49**	0.1	F+I	ALL
Work zone layout Emergency2	Italy	**1.27**	0.04	F+I	ALL
Work zone layout Cross3(0+2)	Italy	**1.25**	0.1	F+I	ALL
Work zone layout Fast2	Italy	1.08	0.06	F+I	ALL
Work zone layout Slow3	Italy	1.03	0.05	F+I	ALL
Work zone layout Emergency3	Italy	1	0.04	F+I	ALL
Work zone layout Middle&Fast3(2)	Italy	0.84	0.07	F+I	ALL
Presence of a work zone (any layout)	Italy	**1.33**	0.02	F+I	ALL
Speed enforcement - section control	Italy	0.99	0.10	F+I	SV
Speed enforcement - section control	Italy	1.06	0.05	PDO	SV
Speed enforcement - section control	Italy	1.12	0.10	F+I	MV
Speed enforcement - section control	Italy	**1.55**	0.10	PDO	MV
Speed enforcement - section control	Italy	1.02	0.08	F+I	SV
Speed enforcement - section control	Italy	**0.86**	0.04	PDO	SV
Speed enforcement - section control	Italy	**0.88**	0.05	F+I	MV
Speed enforcement - section control	Italy	**0.88**	0.04	PDO	MV
Speed enforcement - section control	Italy	0.95	0.11	F+I	SV
Speed enforcement - section control	Italy	**0.85**	0.05	PDO	SV
Speed enforcement - section control	Italy	**0.81**	0.06	F+I	MV
Speed enforcement - section control	Italy	**0.75**	0.05	PDO	MV
Speed enforcement - section control	Italy	0.76	0.19	F+I	SV
Speed enforcement - section control	Italy	0.89	0.11	PDO	SV
Speed enforcement - section control	Italy	**0.6**	0.09	F+I	MV
Speed enforcement - section control	Italy	**0.52**	0.07	PDO	MV
High friction wearing course	Italy	**0.27**	0.10	F+I	ROR, wet pavement

Table 5.2: *CMF functions estimated for the PRACT project*

[‘ denotes value following a change, Δ denotes change (i.e. before – after)]

Value	Country	Function	Injury Type	Crash Type
Horizontal curvature (HC)	England	no significant effect	F+I	ALL
Vertical curvature (VC)	England	$CMF = e^{0.09*\Delta VC}$	F+I	ALL
% HGV (HGV)	England	$CMF = e^{-7.58*\Delta HGV}$	F+I	ALL
% two wheel traffic	England	no significant effect	F+I	ALL
Road width	Germany	$CMF = e^{-0.17*\Delta RW}$	F+I	ALL
Horizontal curvature	Germany	$CMF = e^{0.003*\Delta CU}$	F+I	ALL
Vertical curvature (VC)	Germany	no significant effect	F+I	ALL
% HGV (HGV)	Germany	no significant effect	F+I	ALL

6 References

1. AASHTO (2014). *Highway Safety Manual (supplement)*, First Edition, American Association of State and Highway Transportation Officials.

2. AASHTO (2010). *Highway Safety Manual*, First Edition, American Association of State and Highway Transportation Officials.

3. Bonneson, J. A., Geedipally, S., Pratt, M. P., Lord, D. (2012). *Safety Prediction Methodology and Analysis Tool for Freeways and Interchanges*. NCHRP Project 17-45, Final Report.

4. Bonneson, J. , Pratt, M. (2008). Procedure for Developing Accident Modification Factors from Cross-Sectional Data. *Transportation Research Record*. Journal of the Transportation Research Board 12/2008; 2083(2083):40-48. DOI: 10.3141/2083-05

5. Dinu, R. R., Veeraragavan, A. (2011). Random parameter models for accident prediction on two-lane undivided highways in India. *Journal of Safety Research* 42, pp. 39-42.

6. Eenink R., Reurings M., Elvik R., Cardoso J., Wichert S., Stefan, C. (2007): Accident Prediction Models and Road safety Impact Assessment: Results of the pilot studies. RIPCORD-ISEREST-Projekt, Internal Report No. 2.4

7. Elvik R., Høye A., Vaa T., Sørensen M. (2009) "*The Handbook of Road Safety Measures - second edition*" Ed. Emerald Group Publishing Limited - ISBN: 978-1-84855-250-0

8. Fahrmeir L., Kneib, T., Lang, S. (2009): Regression - Modelle, Methoden und Anwendungen. 2. Auflage, Heidelberg u.a.

9. Federal Highway Administration, Everyday Counts (2015), *"Frequently Asked Questions about High Friction Surface Treatments (HFST)" XXXX*

10. Federal Highway Administration (2015) *Facts and Statistics - Work Zone. Work Zone Injuries and Fatalities.* Available at http://ops.fhwa.dot.gov/wz/resources/facts_stats/injuries_fatalities.htm (Accessed 5 February 2015).

11. Gan, A. e J. Shen, A. Rodriguez (2005), "Update of Florida Crash Reduction Factors and Countermeasures to Improve the Development of District Safety Improvement Project", Florida Department of Transportation, Tallahassee, Florida

12. Harkey, D.L., Srinivasan, R., Baek, J., Council, F., Eccles, K., Lefler, N., Gross, F., Persaud, B., Lyon, C., Hauer, E., Bonneson, J. (2008). *Accident Modification Factors for Traffic Engineering and ITS Improvements*. NCHRP Report 617, NCHRP, Transportation Research Board, Washington DC.

13. Hauer, E. (1997). *Observational Before-After studies in road safety*. Pergamon Press, Elsevier Science Ltd., Oxford, England.

14. Hutcheson G.D., Sofroniou, N. (1999): The Multivariate Social Scientist - Introductory Statistics Using Generalized Linear Models. London/Thousand Oaks/New Dehli.

15. Kennedy J.V., Hall R.D., Barnard S.R. (1998): Accidents at urban mini-roundabouts. Transport Research Laboratory (TRL Report 281), Crowthorne.

16. Khan M., Abdel-Rahim A., Williams C.J. (2015). Potential crash reduction benefits of shoulder rumble strips in two-lane rural highways*. Accident Analysis and Prevention* 75, pp. 35–42.

17. Khattak, A.J., Council F.M. (2002). Effects of work zone presence on injury and non-injury crashes. *Accident Analysis and Prevention*, vol. 34, pp. 19-29.

18. Lamm R., Beck A., Zumkeller K. (1999). *Analyse von Zusammenhängen zwischen Verkehrssicherheit und Straßenentwurf auf Außerortsstraßen - Analysis of relations between traffic safety and road design on rural roads*. Straßen- und Tiefbau Jg. 53, H. 12, S. 6-12.

19. La Torre F., Fanfani F., Matteini F., (2015). *Accident analysis of Italian motorway sections with a high fiction surface treatment*. Bachelor Thesis presented in March 2015.

20. Leutzbach W., Zoellmer J. (1989). *Zusammenhang zwischen der Verkehrssicherheit und den Elementen des Straßenentwurfs - Relationship between traffic safety and road design elements*. Bundesminister für Verkehr, Abt. Straßenbau.

21. Lyon C, Persaud B. (2008). *Safety Effects of a Targeted Skid Resistance Improvement Program - Valid for angle crashes.*

22. Maier R., Berger R., Schüller H. (2013). *Bewertungsmodell für die Verkehrssicherheit von Landstraßen* - Evaluation model for traffic safety on rural roads, Federal Highway Research Institute (BASt).

23. Ministero delle Infrastrutture e dei Trasporti (2002). Decreto 10 luglio 2002: *Disciplinare tecnico relativo agli schemi segnaletici, differenziati per categoria di strada, da adottare per il segnalamento temporaneo*. MIT, Via G. Caraci 36, Roma.

24. Park, B.-J., Fitzpatrick, K. Brewer, M. (2012). *Safety Effectiveness of Super 2 Highways in Texas.* Transportation Research Record 2280, pp. 38-50.

25. Patel R.B., Council F.M., Griffith M.S. (2007). *Estimating Safety Benefits of Shoulder Rumble Strips on Two-Lane Rural Highways in Minnesota: Empirical Bayes Observational Before-and-After Study.* Transportation Research Record 2019, pp. 205-211.

26. Persaud B., Retting R., Lyon C., (2004). *Crash reductions following installation of centre-line rumble strips.* Accident Analysis and Prevention 36, pp. 1073–1079.

27. Persaud, B., Lyon, C., Bagdade, J., Ceifetz, A.H. (2012). *Evaluation of Safety Performance of Passing Relief Lanes.* Transportation Research Record 2348, pp. 58-63.

28. Saleh, P., Aleksa, M., Etl, F., Stütz, R., La Torre, F., Nocentini, A. (2013). *Experience of Speed Management in Practise*. Deliverable D3.1, Appropriate Speed saves All People, CEDR in the Transnational Road Research Programme Call 2012: Safety.

29. Stefan, C. (2006). "*Section control – Automatic speed enforcement in the Kaisermühlen tunnel (Vienna, A22 motorway)*". Wien: Kuratorium für Verkehrssicherheit.

30. Stevens J.P. (2002). Applied multivariate statistics for the social sciences - SAS – SPSS. 4th edition, Mahwah.

31. Taylor M.C., Lynam D.A., Baruya A. (2000). *The effect of driver's speed on the frequency of road accident*s. Transport Research Laboratory (TRL Report 421), Crowthorne.

32. Vieten M., Dohmen R., Dürhager U., Legge K. (2010). *Quantifizierung der Sicherheitswirkungen verschiedener Bau-, Gestaltungs- und Betriebsformen auf Landstraßen* - Quantification of safety effects through different construction, design and operation measures on rural roads, Federal Highway Research Institute (BASt).

33. Vogt, A., Bared, J., (1998). Accident models for two-lane rural segments and intersections. *Transportation Research Record* 1635, pp. 18–29.

34. Yannis, G., Dragomanovits, A., Laiou, A., Richter,T., Ruhl, S., Graham, G., Karathodorou, N, La Torre, F., Domenichini, L., Fanfani, F. (2015). I*nventory and Critical Review of existing APMs and CMFs and related Data Sources*. Deliverable D4, PRACT Project.

Appendix A: Modelling procedure for Negative Binomial models for Germany

Because of the high number of variables in the model the procedure of ‘stepwise selection’ was chosen, which is also important to identify the effects of correlating variables (Taylor et al, 2008). Therefore the individual variable significances were tested within the ‘null’ model (see also Kennedy et al, 1998) with the likelihood-ratio on a significance level of 95 %. Here the coherences between accident occurrence and the individual variables can be evaluated. The result of the individual variable tests is shown in Table A.1. Variables are listed beginning with the most important variables with the highest likelihood-ratio.

Table A.1*: Individual variable significance*

Variable	Likelihood ratio	Significance level
AADT	133.11	0.0000
Road width	34.15	0.0000
Amount of heavy vehicles	13.55	0.0002
Horizontal curvature	3.32	0.0684
Vertical curvature	0.29	0.5895

Based on these results the horizontal and vertical curvature showed no individual significant effects on accident occurrence. That’s why it can be assumed, that these both variables also don’t have an effect in the further accident modelling process. Nevertheless it is noticeable that horizontal curvature is just slight below the boundary value of significance. Finally all the variables were tested in the model during the modelling procedure step by step, starting with the variables with the most important effect on accident occurrence. Which coherences for the variables are assessable is described below.

The full modelling procedure starts with the variable AADT and road width, which have the highest individual effects on safety. Both variables result in a high significance level of model effects and also the parameter estimations are significant on a high level. Also the calculated parameter and their effects on safety are comprehensible.

After putting the amount of heavy vehicles in this step of the modelling, the results of road width turns into insignificance and also the safety effects are not logical. Notwithstanding that there was no high correlation between the percentage of heavy vehicles and the other variables this result have to be considered in depth.

Generally it can be assumed that the overall amount of vehicles on a road segment influences the accident occurrence and also road safety, not just the percentage of heavy vehicles. If there are for example 2 identical road segments with the same percentage of heavy vehicles (e.g. 20 percent), but one road have an overall vehicle amount of 500 vehicles per day (result in 100 heavy vehicles) and the other road has 20,000 vehicles per day (result in 4,000 heavy vehicles), then also the accident occurrence will be very different, despite of the similar percentage of heavy vehicles. This overall amount of vehicles on a road segment is already modelled with the exposition (AADT) and was considered within the model on a high level of significance. That's why it was decided to reject the variable 'amount of heavy vehicles' in the model because of the mentioned reasons. Here the most relevant variable is already involved by AADT in the model, the effect of percentage of heavy vehicles is seen as not relevant.

In a further step the horizontal curvature was tested on model effects. It turns out, that beside the slight individual variable effects on accident occurrence there are strong model effects. That's why also the horizontal curvature was involved in the model. Here just the significance level had to be modified from 5 % to 10 % as inclusion criteria for this variable.

We also tested the leverage values and Cook Distances of the considered road segments within this model to identify road segments in the database, which can be seen as outliers and also can have a distortionary effect on the accident prediction model and the modeling procedure. This road segments can possibly have a negative effect on the goodness of fit of the model to the database and should be examined.

Therefore the road sections with leverage values above the following boundary value can be seen as critical and should be investigated and potentially rejected from the database because of distortionary effects (Stevens, 2002):

$$LV = \frac{3 * (k + 1)}{n}$$

with:

LV = *Leverage value*

k = *Number of parameters*

n = *Number of cases (sample size/road segments)*

Moreover road sections with Cook-distances above 0.5 are seen as conspicuous cases in a database, cases with Cook-distances above 1 should be investigated in depth (Fahrmeier et al, 2009). For this study the stricter boundary values of Hutcheson and Sofroniou (1999) were used with the following calculation:

$$CD = \frac{4}{(n - k)}$$

with:

CD = *Cook-Distance*

k = *Number of parameters*

n = *Number of cases (sample size/road segments)*

For this specific accident prediction model the boundary values for the leverage values is 0.0158 and for the Cook-distances is given by 0.0042. It was noticeable that the road segments with leverage values and Cook-distances above the calculated boundary values have very special condition. Here the AADT often is very high or extreme low and also the road segments are disproportionately long. The comparison with the appertaining accident occurrence and the condition of the other road segments confirmed their status as an outlier. That's why the road segments with exceedance of the boundary values are rejected from the modelling database to improve model and parameter quality. All in all 28 road segments were rejected from the database for the final accident prediction model with better parameter estimation.

After this step the goodness of fit of the model was tested by residual analysis. Thought usage of negative binomial distribution the residuals should have a normal distribution (Eenink et.al., 2007). Therefore the Q-Q-plots of the standardized deviance residuals (see also Maier et al, 2013) were considered and are shown in Figure A.1 before the rejection of the above mentioned 28 road sections and also for the model after their rejection.

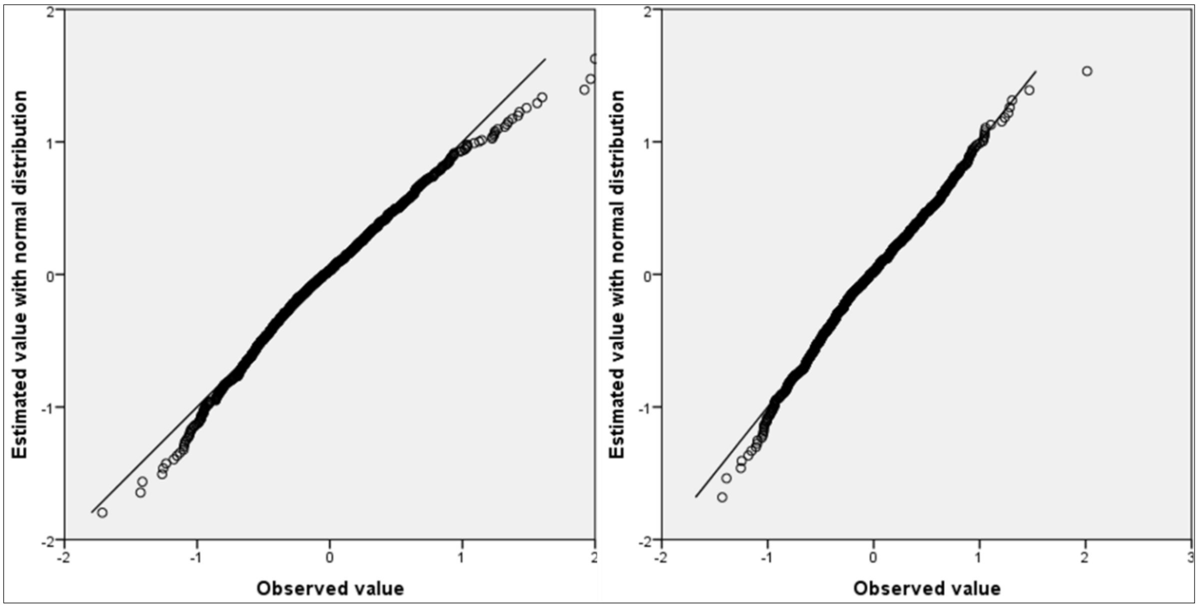

Figure A.1: *Q-Q-plot of model before (left) and after (right) the rejection of segments with high leverage values and cook-distances*

Before the rejection of the 28 road segments it is recognisable, that these outliers have a distortionary effect on the goodness of fit of the model, because there are some differences between the observed values of standardized deviance residuals and those of the standardized deviance residuals estimated by a normal distribution. After their rejections it is noticeable that the distortionary effects were regulated. This will also have positive effects on the model and its estimated parameters.

The histogram for the standardized deviance residuals is depicted in Figure A. and generally shows a normal distribution of standardized deviance residuals with dispersion around ‘zero’. So finally a good description of the database with the model can be assumed.

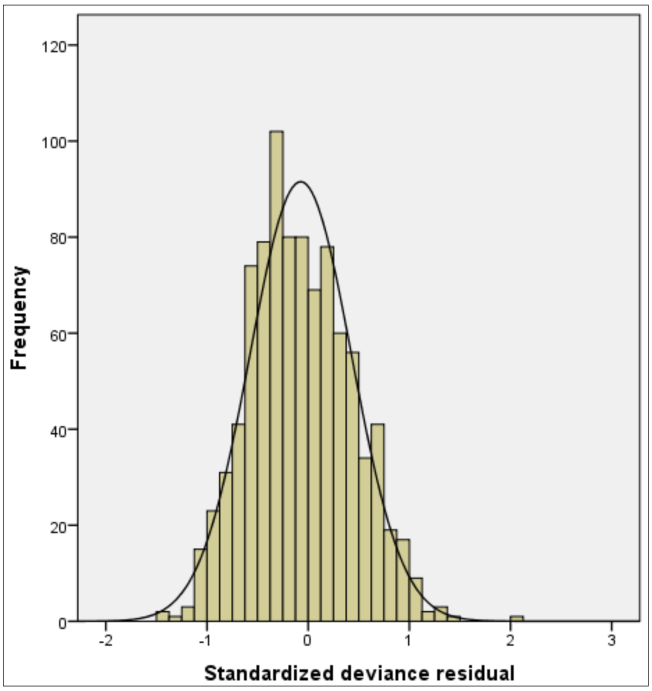

Figure A.2: *Residual plot of final model*