



Doctoral Thesis

Roadway accident risk prediction based on Bayesian probabilistic networks

Author(s):

Deublein, Markus K.

Publication Date:

2013

Permanent Link:

<https://doi.org/10.3929/ethz-a-009795915> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

DISS.ETH Nr. 21093

**ROADWAY ACCIDENT RISK PREDICTION
BASED ON BAYESIAN PROBABILISTIC NETWORKS**

A dissertation submitted to
ETH ZURICH
for the degree of
DOCTOR OF SCIENCES

presented by

MARKUS KLAUS DEUBLEIN

Dipl.-Ing. Silv. Univ., Technische Universität München

born 15.03.1979

citizen of
Germany

accepted on the recommendation of

Prof. Dr. Bryan Tyrone Adey, examiner
Prof. Dr. Michael Havbro Faber, co-examiner
Finn Harald Amundsen, co-examiner
Dr. Matthias Ernst Schubert, co-examiner

2013

Abstract

Transportation networks are an important part of a nation's entire infrastructure asset. They substantially contribute to the economic and social development of that nation but do also entail considerable accident risks for the users. On a world-wide scale, of all transportation infrastructures that people have to deal with every day, the public roads transportation system is the most dangerous and the resultant injuries are a major public health issue.

Only one third of the countries being investigated in the WHO World Report on Road Traffic Injury Prevention (Peden et al., 2004) have a national road safety strategy that is endorsed by their government with specific targets and intervention strategies to reduce the accident risk and that is funded appropriately for its implementation in practice. Although an affirmative trend has been observable in the last years towards lower death and injury rates incurred due to road accidents, traffic injuries remain an important cause of death, injury and disability (NHTSA, 2010, Peden et al., 2009). Road safety is the interest and responsibility of society as a whole and further serious efforts need to be made to reduce the number of road accident victims. There is an urgent need to recognize the human impacts of traffic injuries and deaths as well as their large economic costs to the society (ETSC, 2007). Appropriate steps have to be taken for increasing the sensitivity and attention given to road traffic injury prevention. Significant research is needed to generate more insights into the nature of road traffic problems and into actions which could lead to successful prevention. Models have to be provided that will enable decision makers to understand better how road infrastructure should be developed in order to improve road safety. Such models are relevant for road infrastructure risk-based decision making and can serve as instruments for road infrastructure safety management (European Parliament, 2008).

The aim of this thesis is to establish a methodology that allows developing accurate accident prediction models. The aim is furthermore to support on optimized allocation of available budgets, e.g. for the implementation of accident risk reducing interventions. If the methodology presented in this thesis is used appropriately, it will contribute to improve accident risk analysis and, based on this, to a reduction of the general societal burden of injuries and fatalities incurred by road accidents. Accident risk is herewith the core element. It is considered as the product of three major dimensions: the *exposure*, the *occurrence probabilities* of accident and injury events and the resultant *consequences*. This differentiation helps to deal with the complex issue of accident risk assessment, this differentiation also helps to structure the content of this thesis:

First, an indicator-based generic methodology for the assessment of accident and injury *occurrence probabilities* on road networks is developed. The proposed methodology is designed to be used by accident analysts and road engineers. The methodology is generic in the sense that it is formulated in terms of observable risk indicating variables. Thus, it can be adapted to other infrastructure systems, given appropriate data is accessible. The *exposure* of

road users is taken into account by assessing the occurrence rates (e.g. accidents per million vehicle kilometre) for the different investigated accident events instead of their absolute frequencies (e.g. accidents per year). Bayesian probabilistic inference calculations are used to account for the randomness of accident events and the corresponding uncertainties. A multivariate regression analysis is additionally used to provide empirical prior parameter estimates of the linear relationships between risk indicating and model response variables. These estimates help to establish the structure of a Bayesian Probabilistic Network. When observed accident data becomes available, the causal relationships in the Bayesian Probabilistic Network can be updated by means of parameter learning algorithms. Posterior joint probability distributions are determined for the occurrence rates of the accident and injury events.

Second, two Bayesian methods are compared, which can be used to develop accident prediction models. One of these methods is the Bayesian Probabilistic Network method (developed in this thesis) and the other one is the Empirical Bayes method. The latter is currently the state-of-the-art methodology for accident risk assessment and the evaluation of countermeasures. The differences between the two methods are discussed in terms of their predictive precisions and their applicability as supportive instruments for road infrastructure safety management.

Third, the *consequences* of accident and injury events in terms of economic costs are addressed. The proposed methodology for the development of accident prediction models is imbedded into a framework for road user impact assessment. Such a probabilistic modelling approach in combination with the use of Bayesian Probabilistic Networks is new in the context of road infrastructure decision making. Impact models for different stakeholder groups affected by road networks require various model assumptions. The different types of uncertainties connected to these model assumptions can be considered using the features of the Bayesian Probabilistic Networks. The predicted impacts are provided in terms of joint posterior distribution functions of the expected costs.

The scope for the model development and its application is set on a probabilistic investigation of road design and traffic related risk promoting factors and their influence on the occurrence probabilities of injury accidents and different levels of injuries. The developed models are applied to road segments of a rural motorway network and implemented into impact models for road users.

Zusammenfassung

Verkehrsnetze sind ein wichtiger Teil der gesamten Infrastruktur einer Nation. Sie tragen wesentlich zur wirtschaftlichen und sozialen Entwicklung des Landes bei, sie verursachen aber auch bedeutsame Risiken für die Nutzer. Von allen Verkehrssystemen, die die tägliche Mobilität der Gesellschaft gewährleisten, ist das öffentliche Strassensystem weltweit eines der gefährlichsten Verkehrssysteme. Verletzungen und Todesfälle durch Verkehrsunfälle auf Strassen sind ein beachtliches Gesundheitsproblem für die Gesellschaft.

Nur etwa ein Drittel der Länder, die im WHO World Report on Road Traffic Injury Prevention (Peden et al., 2004) aufgeführt werden, kann ein nationales Konzept für Strassensicherheit vorweisen, das von den jeweiligen Regierungen verabschiedet ist, spezifische Ziele und Massnahmen zur Risikoreduktion beinhaltet und mit angemessenen finanziellen Mitteln ausgestattet ist, um in der Praxis eingesetzt werden zu können. Obwohl während der letzten Jahre ein positiver Trend hin zu geringeren Todes- und Verletztenraten beobachtbar war, bleiben Unfälle im Strassenverkehr eine bedeutende Ursache für Tod, Verletzung und körperlicher Behinderung (Peden et al., 2009, NHTSA, 2010). Die Strassensicherheit liegt im Interesse und Verantwortungsbereich der gesamten Gesellschaft und es müssen anhaltend ernsthafte Bemühungen unternommen werden, um die Anzahl an Verkehrsoptern weiterhin zu senken. Dabei ist es wichtig, die Auswirkungen von Verletzungen und Todesfällen auf die Menschen zu verstehen und die ökonomischen Konsequenzen für die Gesellschaft zu berücksichtigen (ETSC, 2007). Darüber hinaus müssen geeignete Strategien entwickelt werden, um eine grössere öffentliche und politische Aufmerksamkeit und Sensibilität für das Problem von Verkehrsunfällen zu erreichen. Massgebliche Forschung ist notwendig, um mehr Einblicke in die Natur des Ursachen-Wirkungs-Geflechts von Verkehrsunfällen zu erlangen und um Handlungsalternativen zu identifizieren, die wirkungsvoll das Unfallrisiko auf Strassen mindern können. Modelle werden benötigt, die den Entscheidungsträgern zu einem besseren Verständnis verhelfen, wie sich die Strasseninfrastruktur entwickeln sollte, damit die Verkehrssicherheit auf den Strassen verbessert werden kann. Die Modelle können als Hilfsmittel im Rahmen einer risikobasierten Entscheidungsfindung eingesetzt werden oder als Baustein für Infrastruktur-Sicherheitsinstrumente (European Parliament, 2008) dienen.

Das Ziel dieser Arbeit ist die Ausarbeitung einer Methode, die es ermöglicht, genaue Vorhersagemodelle für das Unfallrisiko von Strassennutzern zu entwickeln. Das Unfallrisiko wird als das Produkt dreier Hauptdimensionen verstanden: die *Exposition* der Strassennutzer, die *Wahrscheinlichkeit*, dass Unfall- und Verletzungsereignisse eintreten, sowie die daraus resultierenden *Konsequenzen* für die unterschiedlichen Interessensvertreter. Diese Differenzierung hilft, mit dem komplexen Sachverhalt von Unfallrisiken umzugehen, diese Differenzierung hilft auch, um den thematischen Inhalt der vorliegenden Arbeit zu strukturieren:

Den ersten Schritt der Untersuchungen bildet die Entwicklung einer indikatorbasierten, generischen Methode zur Ermittlung der *Wahrscheinlichkeiten* von Unfall- und Verletzungsereignissen auf Strassennetzen. Die vorgeschlagene Methode ist so gestaltet, dass sie schnell und einfach von Unfallgutachtern und Strassen-Ingenieuren angewendet werden kann. Die Methode ist generisch, da sie auf Basis von beobachtbaren Indikatorvariablen formuliert wird und daher auch auf unterschiedliche Infrastruktursysteme übertragbar ist, sofern die erforderlichen Daten zur Verfügung stehen. Die *Exposition* der Strassennutzer wird berücksichtigt, indem Ereignisraten (z.B. Verletzte pro Millionen Fahrzeugkilometer) und nicht absolute Häufigkeiten (z.B. Verletzte pro Jahr) betrachtet werden. Um der natürlichen Zufälligkeit von Unfall- und Verletzungsereignissen Rechnung zu tragen, wird ein probabilistischer Modellansatz gewählt, der auf der Bayes'schen Wahrscheinlichkeitstheorie beruht. Solch ein Ansatz ermöglicht die Berücksichtigung von Unsicherheiten, die aufgrund von Modellannahmen oder durch die zur Verfügung stehende Datenmenge im Modell vorhanden sind. Die im Modell enthaltenen Variablen stehen in bestimmten funktionalen und kausalen Zusammenhängen zueinander. Eine multivariate Regressionsanalyse wird angewandt, um zuerst empirisch die Parameter dieser Zusammenhänge zu ermitteln. Auf Basis dieser Schätzwerte wird ein Bayes'sches Wahrscheinlichkeitsnetz entwickelt. Die Parameter im Bayes'schen Wahrscheinlichkeitsnetz werden anhand zusätzlicher Daten und unter Verwendung von Lern-Algorithmen aktualisiert. Die gemeinsamen, voneinander abhängigen Wahrscheinlichkeitsverteilungen für die Ereignisraten von Unfällen und Verletzten werden simultan ermittelt.

In einem zweiten Schritt werden zwei Methoden verglichen, die beide auf der Bayes'schen Wahrscheinlichkeitstheorie basieren und verwendet werden können, um Unfallvorhersagemodelle zu entwickeln. Eine dieser Methoden ist die in dieser Arbeit vorgeschlagene Methode unter Verwendung von Bayes'schen Wahrscheinlichkeitsnetzen. Die zweite Methode ist die sogenannte Empirical Bayes Methode. Diese wird derzeit in der Analyse von Strassenverkehrsrisiken und bei der Evaluation von geeigneten Gegenmassnahmen am häufigsten verwendet. Die Unterschiede zwischen diesen beiden Methoden werden diskutiert, ihre Vorhersagegenauigkeiten werden evaluiert und ihre Möglichkeiten zur Implementierung in Infrastruktur-Sicherheitsinstrumente werden erörtert.

Als dritter Schritt werden die *Konsequenzen* von Verletztenunfällen betrachtet. Die vorher in dieser Arbeit vorgeschlagene Methode zur Entwicklung von Unfallvorhersagemodellen wird in ein übergeordnetes Modell für die Abschätzung von multiplen Auswirkungen von Strassen-Infrastruktursystemen auf deren Nutzer implementiert. Dieser probabilistische Modellansatz stellt eine neuartige Methode im Zusammenhang mit einer risiko-basierten Entscheidungsfindung im Strassen-Infrastrukturmanagement dar. Für die Entwicklung von Modellen, die die Auswirkungen von Strassennetzen auf unterschiedliche Interessensvertreter abbilden, müssen zahlreiche Annahmen getroffen werden. Diesen Modellannahmen haften unterschiedliche Formen von Unsicherheiten an, die durch einen probabilistischen

Modellansatz im allgemeinen und durch die Verwendung von Bayes'schen Wahrscheinlichkeitsnetzen in der vorliegenden Arbeit berücksichtigt werden. Auf dieser Basis werden die vorhergesagten Auswirkungen anhand von gemeinsamen Wahrscheinlichkeitsverteilungen der zu erwarteten Kosten abgebildet.

Die Untersuchungen dieser Arbeit beschränken sich auf eine probabilistische Analyse von Strassenparametern und Verkehrscharakteristiken und deren Einfluss auf das Unfall- und Verletzungsrisiko im Strassenverkehr. Die entwickelten Modelle werden in einem Fallbeispiel für Strassenabschnitte von Autobahnnetzen ausserhalb von Ortschaften angewendet. Sie werden zusätzlich in Modelle zur Bestimmung von multiplen Auswirkungen auf Strassenutzer implementiert.