

学 位 論 文 題 名

A Study on Traffic Accident Analysis using Parametric and Non-Parametric Approaches

(パラメトリック法とノンパラメトリック法を用いた
交通事故分析に関する研究)

学位論文内容の要旨

This study concentrates on road geometry related causes of traffic crashes. The goal is to find out how to renovate the existing road geometry in order to reduce the number of traffic crashes in an effective manner. Traffic crash data and road geometry data for the national highways in Hokkaido were used in this study. There are several reasons to focus on national highways in Hokkaido. First, vehicle traffic is a life line for the rural areas due to less frequent transit service. Traffic crashes in rural area have high effect on life of people in Hokkaido. Second, rural road accounted for the half of the traffic crashes occurred in Hokkaido. Also, rural crashes have high fatal crash ratios due to high speed traffic. Third, there are a few studies referred to the causal relation between number of crashes and road geometries clearly in Japan. When analyzing the effects of road geometry on traffic crashes, a study area with less compounding road factors such as cross roads and pedestrian crosses is preferable. Forth, digital database regarding road geometries, traffic crashes and traffic conditions for 16 years are available for the national highways in Hokkaido. Owing above mentioned factors, rural sections of the road are selected for this study. In addition to the road geometry, effects of winter season on traffic crashes have to be included. Numerous studies have concentrated on the effects of road geometry on traffic crashes. In such studies, reported effects of variables on traffic crashes differ study to study, and methodologies used are not always suitable. Due to filling needs and shortcoming listed above, this study is designed as a traffic crash data analysis, with the hypothesis that the traffic crashes in rural roads are affected by road related and weather variables. Objectives are to find the effects of road geometry on traffic crashes as a function of season, and to develop a better performing methodology to do so. Different road segment types or lengths, different regression model distributions, and combination of parametric with non-parametric models were used to find a better performing traffic crash prediction model.

This dissertation is divided into 7 chapters and the details are explained hereafter. Chapter 1 introduces the background, problem statement, objectives, outline, and scope of this study. This chapter highlights the current traffic crash situation in Hokkaido. Lay down arguments in favor of this study. Chapter ends with a description of the structure of this dissertation.

Chapter 2 summarizes the literatures related to this study. Reviewed literature included literature on traffic crashes occurred on rural roads, head-on crashes, and traffic crash analysis methods. Under traffic crash analysis methods, uses of parametric and non-parametric approaches in literature are reviewed. Nature previous studies their empirical settings and methodology were summarized.

Chapter 3 presents the use of parametric approach to model head-on crashes. Poisson regression model (PRM) as a parametric model was used to relate head-on crashes with road geometry variables. As a comparison study, single vehicle crashes were also modeled with PRM. Fixed-length road segments, of four different lengths, were used to create 4 PRMs. Results of developed PRMs identified several

critical variables affecting head-on crashes. The model proposed in this study is potentially capable of identifying the causal factors of head-on and single vehicle crashes.

Chapter 4 presents the combine use of parametric and non parametric approaches. Negative binomial regression model (NBRM) was used to relate traffic crashes. Road geometry and cross section variables were used as independent variables and crash data for 16 years was used as the dependent variable. In addition, this study developed a methodology to combine variables using non parametric approach. Combination variables for road geometry and cross-section variables were developed using the Chi-squared automatic interaction detection (CHAID) model. Homogeneous, and 1-km road segments were used in two separate NBRMs. Homogeneous road segments were divided based on the horizontal alignment of the road. Results indicated that the use of homogenous road segments would increase the model accuracy. Further the combination effects were found have a significant effect on traffic crashes. Chapter 5 presents a similar attempt as in chapter 4. However, in this chapter traffic crash data was limited to 5 years (1997-2002) and homogeneous road segments were computed in two methods; one based only on the horizontal alignment of the road and the other based both horizontal alignment and the vertical alignment of the road. In addition to the homogenous road segments a 1 km fixed length road segments was also used. Accordingly three NBRMs were developed, correlating the road geometry and cross section variables with traffic crashes. The NBRM with homogeneous road segments, which were divided based only on the horizontal alignment, had the largest value of log-likelihood ratio, largest value of the coefficient of determination (R^2), and the highest number of significant variables.

Chapter 6 presents the use of non parametric approaches to model non intersectional, rural traffic crashes in national highways. A non parametric approach called, hierarchical tree based regression model (HTBRM) was used in the study. CHAID algorithm was used for HTBRM. CHAID model can produce a non binary tree this means that CHAID model can capture the complex relationship between independent variables and traffic crashes. In addition CHAID model can process incomplete databases and produce predicted values for the road segments with missing variable data. Traffic crash data for 5 years (1997-2002) with two fixed length road segments (1 km, and 3 km) were used to create two sets of CHAID models (winter, non-winter). In contrast with the NBRM results in the previous chapters, CHAID model have identified several complex relationship between the number of traffic crashes variables such as slope, bendiness, and average shoulder width.

Chapter 7 concludes this dissertation, highlighting the main contributions of this study along with some recommendations on future research directions. This study contributed with developing several traffic crash prediction models. Since the developed traffic crash prediction models used a large crash database covering 6 national highways, the conclusions would be reliable. Further parametric, non parametric, and combined parametric models approaches were used to create a reliable traffic crash prediction model. Several causal factors to traffic crashes such as winter months, tunnels, slopes and total curved length were identified from the models. This study contributed with creation of combination variables. A methodology to use non parametric model called CHAID to compute ' combination variables' was developed. Such created combination variables were used in a parametric model, NBRM. When the combination variables were used in the NBRM, the combination variables were found to be significantly affecting the number of traffic crashes. This study contributed with road segment selection for the traffic crash analysis. In creating the traffic crash prediction models, various homogeneous and fixed length road segments were used. The use of homogenous road segments is desirable over the fixed length road segments, as it increased the model accuracy. This study contributed with use of CHAID model in traffic crash prediction model. CHAID models developed have identified the effects of slopes, tunnels bridges in winter. However the NBRM model developed with same data in chapter 5 did not find any significant effect with respect to tunnels or slopes. Future studies have to find suitable counter measures for the causal factors identified in this study.

学位論文審査の要旨

主 査 教 授 萩 原 亨

副 査 教 授 加賀屋 誠 一

副 査 教 授 中 辻 隆

副 査 准教授 浜 岡 秀 勝 (秋田大学)

学 位 論 文 題 名

A Study on Traffic Accident Analysis using Parametric and Non-Parametric Approaches

(パラメトリック法とノンパラメトリック法を用いた

交通事故分析に関する研究)

日本では死亡事故および重軽傷事故は、近年減少してきているが、自動車に依存する地域において、交通事故は生活を脅かすものであり、交通事故を減らすための継続的な施策は必要と言える。また、これから車社会となろうとしているアジア諸国にとって、自動車利用の増大による交通事故増加を抑制する技術は社会基盤整備政策にとって重要となる。

本論文では、道路線形が交通事故件数に与える影響を明らかにする事故予測モデルの構築を主に論じている。交通事故を抑制する効率的な道路線形改良を提案することを目指すために必要な技術となる。このような予測モデルは、交通工学の黎明期から最近に至るまで多数研究されてきている。しかし、事故データベースの不備などから、パラメトリックな手法を用いる理論的な基礎研究が先行する一方、実用的な予測モデルが少ない点が課題となっている。そこで、本論文では、(独)土木研究所・寒地土木研究所・寒地交通チームが長年に渡って蓄積してきた北海道の国道で発生した交通事故データベースを基礎に、事故予測モデルに関する新しい知見を得ることを目的としている。本論文は、7章から構成されている。

第1章は、序論であり、本研究の目的と意義が説明されている。ここでは、北海道の交通問題として郊外における事故が減っていないこと、冬期における事故が依然として多いことを指摘し、道路交通に依存する北海道において自動車事故の発生は生活に直結しており、解決すべき課題であることを記述している。これらの課題を道路側から解決するため、効率的かつ効果的な道路改良を提案するための事故予測モデルに関する検討の流れを記述している。

第2章では、道路線形が交通事故に与える影響を明らかにする手法についてレビューした結果についてまとめている。道路線形と交通事故発生の因果関係を論じる研究は、1950年代からあるが、ここでは、一般化線形モデルが適用されるようになってきた1990年頃からの事故件数予測モデルに着目し、その研究について追跡し、予測モデルの課題について言及している。

第3章では、北海道で大きな課題となっている正面衝突事故に焦点を当て、その事故件数予測モ

デルを構築している。正面衝突事故と比較するため単独事故についても同様の分析を行っている。パラメトリックな手法を用いたモデルとして、ポアソン回帰 (PRM) を適用し、道路線形と季節が正面衝突事故および単独事故件数に与える影響とその事故件数の予測を行っている。道路区間を 1km, 3km, 5km, 10km とし、それらの事故を増加させる道路線形を明らかにしている。また、正面衝突事故と単独事故の差異から、正面衝突事故の特徴を示している。しかし、説明変数の影響として解釈が困難なものがあるなど、予測モデルの課題が明らかになった。

第 4 章では、パラメトリックな手法とノンパラメトリックな手法とを組み合わせ、3 章で明らかとなった課題をクリアすべく、交通事故件数の予測を行う新しい手法にアプローチしている。事故件数の予測には、パラメトリックな手法である Negative binomial regression model (NBRM) を用いている。事故データベースとして、1991-2007 の 16 年間のものを用いた。道路線形データベースおよび交通量データベースは 1999 年のものを用いた。本章では、説明変数として道路線形を組み合わせた変数をノンパラメトリックな手法で生成し、予測モデルに利用している。また、事故区間として 1km というような固定した区間に加え、平面線形が均一な区間を設定し、事故予測精度の比較を行っている。NBRM による予測結果から、平面線形を均一な区間とすることが予測精度を高くすることを示し、かつ組み合わせ変数が説明変数として有意となり、事故件数に影響する具体的な道路線形を示し得ることを明らかにした。

第 5 章では、事故データベース (1997-2002) と道路線形データベースおよび交通量データベースの時期に一致させ、予測精度の更なる検討を行っている。道路線形が均一な区間として、平面線形に加えて縦断線形も考慮した均一区間を作成し、予測精度の結果を検討している。第 4 章と同様に NBRM を用いて、対数尤度比・決定係数・有意となった説明変数の種類から予測結果を比較検証し、平面線形を均一区間とした場合が最も高い精度となることを明らかにした。

第 6 章は、ノンパラメトリックなモデルを用いて事故件数の予測を行っている。ノンパラメトリックな手法として、回帰木とした階層ツリー構造 (階段折れ線でモデルを構成) を用いている。CHAID (Chi-squared automatic interaction detection) アルゴリズムを使い、変数の基準としてカイ 2 乗統計量を使って、階層構造を生成している。CHAID アルゴリズムの特徴は、多分岐が可能であり、説明変数と目的変数の複雑な関係を知ることができる点にある。事故分析のように多数の説明変数があるときには有効と言える。1997-2002 年の事故データから、1km および 3km の固定区間を用いて、道路線形と季節が事故件数に与える影響を分析している。5 章で示したパラメトリックな手法の結果と異なり、勾配や曲線および路肩幅員などの影響を示すことができ、ノンパラメトリックな手法のメリットを示すことに成功している。

第 7 章は、以上の検討結果を踏まえ、ここで開発した 4 種類の事故予測モデルの結果を示し、事故予測モデル間の差異について取りまとめている。

これを要するに、著者は、ポアソン回帰、負の二項分布回帰などの一般化線形モデルによる手法およびデータマイニングを適用し、道路線形が交通事故に与える影響に関する新しい予測モデルを提案し、自動車社会における負の側面である交通事故を軽減するための具体的な道路構造の評価手法において新知見を得たものであり、社会基盤計画学および交通工学において、貢献すること大なるものがある。よって、著者は北海道大学博士 (工学) の学位を授与される資格あるものと認める。