CETRA 2018

5th International Conference on Road and Rail Infrastructure
17–19 May 2018, Zadar, Croatia

Road and Rail Infrastructure V

Stjepan Lakušić – EDITOR

IRF GLOBAL
FEHRL

Organizer
University of Zagreb
Faculty of Civil Engineering
Department of Transportation

CETRA 2018

5th International Conference on Road and Rail Infrastructure

17–19 May 2018, Zadar, Croatia

TITLE

Road and Rail Infrastructure V, Proceedings of the Conference CETRA 2018

EDITED BY

Stjepan Lakušić

ISSN

1848-9850

ISBN

978-953-8168-25-3

DOI

10.5592/CO/CETRA.2018

PUBLISHED BY

Department of Transportation
Faculty of Civil Engineering
University of Zagreb
Kačićeva 26, 10000 Zagreb, Croatia

DESIGN, LAYOUT & COVER PAGE

minimum d.o.o.
Marko Uremović · Matej Korlaet

PRINTED IN ZAGREB, CROATIA BY

“Tiskara Zelina”, May 2018

COPIES

500

Zagreb, May 2018.

Proceedings of the
5th International Conference on Road and Rail Infrastructures – CETRA 2018
17–19 May 2018, Zadar, Croatia

Road and Rail Infrastructure V

EDITOR
Stjepan Lakušić
Department of Transportation
Faculty of Civil Engineering
University of Zagreb
Zagreb, Croatia

ORGANISATION

CHAIRMEN

Prof. **Stjepan Lakušić**, University of Zagreb, Faculty of Civil Engineering
Prof. emer. **Željko Korlaet**, University of Zagreb, Faculty of Civil Engineering

ORGANIZING COMMITTEE

Prof. Stjepan Lakušić
Prof. emer. Željko Korlaet
Prof. Vesna Dragčević
Prof. Tatjana Rukavina
Assist. Prof. Ivica Stančerić
Assist. Prof. Maja Ahac
Assist. Prof. Saša Ahac
Assist. Prof. Ivo Haladin
Assist. Prof. Josipa Domitrović
Tamara Džambas
Viktorija Grgić
Šime Bezina
Katarina Vranešić
Željko Stepan
Prof. Rudolf Eger
Prof. Kenneth Gavin
Prof. Janusz Madejski
Prof. Nencho Nenov
Prof. Andrei Petriaev
Prof. Otto Plašek
Assist. Prof. Andreas Schoebel
Prof. Adam Szeląg
Brendan Halleman

INTERNATIONAL ACADEMIC SCIENTIFIC COMMITTEE

Stjepan Lakušić, University of Zagreb, president
Borna Abramović, University of Zagreb
Maja Ahac, University of Zagreb
Saša Ahac, University of Zagreb
Darko Babić, University of Zagreb
Danijela Barić, University of Zagreb
Davor Brčić, University of Zagreb
Domagoj Damjanović, University of Zagreb
Sanja Dimter, J. J. Strossmayer University of Osijek
Aleksandra Deluka Tibljaš, University of Rijeka
Josipa Domitrović, University of Zagreb
Vesna Dragčević, University of Zagreb
Rudolf Eger, RheinMain Univ. of App. Sciences, Wiesbaden
Adelino Ferreira, University of Coimbra
Makoto Fujiu, Kanazawa University
Laszlo Gaspar, Széchenyi István University in Győr
Kenneth Gavin, Delft University of Technology
Nenad Gucunski, Rutgers University
Ivo Haladin, University of Zagreb
Staša Jovanović, University of Novi Sad
Lajos Kisgyörgy, Budapest Univ. of Tech. and Economics
Anastasia Konon, St. Petersburg State Transport Univ.
Željko Korlaet, University of Zagreb
Meho Saša Kovačević, University of Zagreb
Zoran Krakutovski, Ss. Cyril and Methodius Univ. in Skopje
Dirk Lauwers, Ghent University
Janusz Madejski, Silesian University of Technology
Goran Mladenović, University of Belgrade
Tomislav Josip Mlinarić, University of Zagreb
Nencho Nenov, University of Transport in Sofia
Mladen Nikšić, University of Zagreb
Andrei Petriaev, St. Petersburg State Transport University
Otto Plašek, Brno University of Technology
Mauricio Pradena, University of Concepcion
Carmen Racanel, Tech. Univ. of Civil Eng. Bucharest
Tatjana Rukavina, University of Zagreb
Andreas Schoebel, Vienna University of Technology
Ivica Stančerić, University of Zagreb
Adam Szeląg, Warsaw University of Technology
Marjan Tušar, National Institute of Chemistry, Ljubljana
Audrius Vaitkus, Vilnius Gediminas Technical University
Andrei Zaitsev, Russian University of transport, Moscow

CETRA[2018] 17–19 May 2018, Zadar, Croatia
5th International Conference on Road and Rail Infrastructure

VISUALIZE THE EFFECT OF INPUT VARIABILITY ON MODEL OUTPUT IN TRAFFIC ASSIGNMENT

Mundher Ali Seger, Lajos Kisgyorgy
Budapest University of Technology and Economics, Hungary

Abstract

Uncertainty can be found at every stage of travel demand model, where passed from each stage to another and propagated over the whole model. Thus, studying the uncertainty in the last stage (traffic assignment) is more important because it represents the result of uncertainty in the travel demand model. The purpose of this paper is to assist transportation modelers and decision makers, to have a fresh look at the uncertainty in traffic assignments of transportation models. By building a new methodology to predict the likelihood of traffic assignment probability distribution and compare predicted values to real values or another prediction methods, the paper shows the uncertainty in traffic volumes, and the amounts of errors and biases in the results as well. The methodology quantifies the uncertainty in modeling by Monte Carlo simulation. A probability distribution is assigned to all cells of the OD matrix, considering them as stochastic input variables. The distributions of the output values of traffic assignment are classified and into four cases according to errors and bias. Finally, the results are drawn into figures to visual the uncertainty in traffic assignments. The paper constructs three types of probability distributions to the input data. For each type of distributions different parameter assignments, such as different variation values; are analyzed. For each of these parameter assignments, one thousand Monte Carlo samples were made, with the classification and visualization of the results.

Keywords: uncertainty quantification, uncertainty visualization, Monte Carlo simulation, traffic assignment

1 Introduction

Visualization is a useful method for addressing many forms of information uncertainty. Applications that use visual graphs and comparative figures to indicate information variability or draw levels of confidence in data values help analysts better understand and cope with uncertain information better than using digital tables and metadata [1]. Consequently, visualizing the uncertainty is essential for risk analysis and decision-making tasks. But, it is still a challenge, because of describing the uncertainty is a complex the concept with many interactions, definitions, and interpretations in transportation models.

Uncertainty can be introduced into information visualizations as the data is collected, transformed and integrated into information. In the absence of combined presentation of data and its associated uncertainty, the analysis of the information visualization is incomplete at best and may lead to inaccurate or incorrect conclusions. Therefore, there is a need to display information together with their uncertainty for accurate interpretation and precise decision making [2].

There are different methods used to visualize uncertainty; statistical and probability-based visualization, point and global visualization, used colors, financial visualization, icons, on-

DOI: https://doi.org/10.5592/CO/CETRA.2018.683

tology, lexicon, etc. [3]. In this paper, statistical and probability-based visualization method is used to visualize uncertainty. This method is one of the most powerful methods to address conceptual model uncertainty is with a traditional histogram, and probability distributions represented by random variables. This method demonstrates the central tendency, dispersion, skewness, and modal characteristics of a random variable.

2 Uncertainty analysis in transportation modeling

The traffic forecasts produced by transport models are subject to some sources of uncertainty including errors in the measurement of input data, errors in the estimated value of model parameters and errors in the specification of the underlying models themselves. Ideally, analysts would wish to understand the separate and collective impact of these errors on the uncertainty of model forecasts, to be able to attach credible confidence intervals to model forecasts and optimize the allocation of study resources. However, in large model systems, the interaction between each of these sources of error can be very complicated, making the analysis of propagation of uncertainty through the modeling process extremely challenging. Nevertheless, the increased participation in recent years of the private sector in the delivery of transport infrastructure projects has raised the requirement for accurate traffic demand forecasts and led to renewed interest in the analysis of model uncertainty [4]. Uncertainty becomes relevant in transportation modeling only in case of diverging views if risks are very high if the policy is controversial and if there are concerns about model limitations. In certain cases, several points estimate based on different scenarios are given to account for uncertainty [5]. The main goal of travel demand model is traffic forecasting in different stages; generation, distribution, and assignment are to determine future values of the model output variables that are associated with a specific combination of input variables. However, it is impossible to give an exact prediction; no model can be constructed to provide 100 % accurate predictions of the future behavior of a system. A prediction should handle uncertainties by treating output variables stochastically. Without the additional information provided by probability analysis, there is no solid evidence for comparing the predicted value to real value or another prediction. As a result, any method used for prediction should include an assessment of the uncertainty in the predicted values.

3 Methodology

This paper introduces a new methodology for quantifying and characterizing predictive uncertainty in traffic assignment models. The structure of this work directly supports a visual segmentation of uncertainty for transport network to present error and bias in traffic volumes that calculated by traffic assignment models. We start by presenting the formalisms required to review our method (Section 3.1). We then introduce the simulation method used (Section 3.2). We then explain the process used to predict the uncertainty (Section 3.3). We then define a strategy for the uncertainty tolerance (Section 3.4). Finally, we present the method used to classify and visualize the uncertainty (Section 3.5).

3.1 Formalism

The principal task in predictive modeling is to estimate the behavior of a modeling function, in this case, traffic assignment function $f_{T_{O \to D}}$ defined to calculate the traffic volume between Oregon zone (O) to Destination zone (D). This work addresses the case where $f_{T_{O \to D}}$ can be calculated at a finite set of samples N_s. Monte Carlo simulation method was used to generate the input data with standard deviation σ and average value μ, to produce OD trip distribution for three types of probability distribution. Visum software was used in this work to calculate the traffic assignment function. The required data to find the predicted traffic assignment are:

$Z; \forall Z \in N$ – Zones definition;
$O; O \subset Z$ – Origin definition;
$D; D \subset Z$ – Destination definition;
$OD; \forall OD \in R$ – matrix of Origin Destination (OD);
$m; \forall m \in N$ – any link in the network;
$M; \forall M \in N$ – number of links;
$x; \forall x \in R$ – observed values of traffic volume for the links;
$\mu, \sigma; \forall \mu, \sigma \in R$ – parameters range of MC simulation;
μ – observed OD values;
σ – standard deviation;
$N_s; \forall N_s \in N$ – number of iterations.

3.2 Monte Carlo simulation

Monte Carlo (MC) methods play a fundamental role in characterization and quantification of uncertainty. When the accurate calculation of output uncertainties needed then Monte Carlo based analysis is a reliable technique, and it is widely applicable. As a result, its application can be found in virtually all engineering fields. Monte Carlo simulation was usually utilized to observe how errors or variability of a system can propagate to the final result.
In this work, a Matlab software and Visum-COM programming are used to build a Matlab code to calculate the uncertainty in traffic volume for the transport network links. And using 1000 iterations in this simulation. The application of this simulation was done in four steps: (1) generating a random data for all parameter OD matrix, (2) running Visum to find the traffic assignments, (3) finding uncertainty of all traffic links, and (4) analyses of the process output.

3.3 Predictive uncertainty

Predictive uncertainty is defined by joint consideration of the mean predictive error (i.e., statistical bias) and the predictive variability (i.e., statistical standard deviation). In this case, the uncertainty has been predicted by getting the traffic volume attribute for links from traffic assignment function (Eq. 1) using Visum; this equation has been applied for 1000 iterations. And then find the average traffic volume for all links in transport network (Eq. 2). Statistical bias in traffic volume represents the difference between the average calculated value and observed value (Eq. 3). Finally, standard deviation represents the variability of results (i.e., statistical error) (Eq. 4).

$$x_m = f_{TO \to D}(OD) \tag{1}$$

$$\overline{x}_m = \sum x_m / N_s \tag{2}$$

$$n_m = \left(\overline{x}_m - x_m\right) \tag{3}$$

$$\sigma_m = \sqrt{\frac{1}{N_s} \sum \left(x_m - \overline{x}_m\right)^2} \tag{4}$$

Where:
χ_m – traffic volume of traffic assignment function for the link (m);
$f_{T_{O \to D}}$ – traffic assignment function;
$\overline{\chi}_m$ – average traffic volume of traffic assignment function for the link (m);
η_m – bias in traffic volume for the link (m);
σ_m – the andard deviation in traffic volume for the link (m).

3.4 Uncertainty limitation

In any uncertainty quantification process, setting limits for the predictive uncertainty required to increase understanding the researchers to models behavior in both bias and predictive variability. The GEH statistic has used as a limitation of the bias in this study (Eq. 5). The GEH statistic is a form of Chi-squared statistic that can be used to compare observed and modeled counts [6]. It is helpful for these comparisons because it is sophisticated of relative and absolute errors. And, the standard deviation statistic was adopted as a limitation for the variability in traffic volumes.

$$GEH = \sqrt{\frac{(M-C)^2}{(M+C)/2}} \tag{5}$$

Where:
M – the modeled flow;
C – the observed flow.

GEH statistic bands less than 5, is used to explain bias limit for each link (Eq. 6).

$$GEH = \sqrt{\frac{(x_m - x_m)^2}{(x_m + x_m)/2}} = 5 \tag{6}$$

By solving (Eq. 6) the upper and the lower bias limit in predicted traffic volume as follow:

$$c_{l_m} = x_m + 6.25 - \frac{\sqrt{100\ x_m + 156.25}}{2} \tag{7}$$

$$c_{u_m} = x_m + 6.25 + \frac{\sqrt{100\ x_m + 156.25}}{2} \tag{8}$$

Where:
c_{l_m} – the lower bias limit for each link;
c_{u_m} – the upper bias limit for each link.

3.5 Visualizing predictive uncertainty

The last step of this methodology is uncertainty visualization. First, we characterize this uncertainty according to bias and variability for results into four cases:
Case I: Accurate and converged. This case occurred when the predicted traffic volumes (χ_m) are close to the mean prediction χ_m, and predicted traffic volumes are within the allowed bias limits (c_u and c_l). That is mean the results low variability and low bias.
Case II: Accurate and not convergent. This case occurred when the predicted traffic volumes (χ_m) are close to the mean prediction $\overline{\chi}_m$, and predicted traffic volumes are out of the allowed limits (c_u and c_l). That is mean the results low variability and high bias.
Case III: Inaccurate and converged. This case occurred when the predicted traffic volumes (χ_m) are far from the mean prediction $\overline{\chi}_m$, and predicted traffic volumes are within the allowed bias limits (c_u and c_l). That is mean the results high variability and low bias.
Case VI: Inaccurate and not convergent. This case occurred when the predicted traffic volumes (χ_m) are far from the mean prediction $\overline{\chi}_m$, and predicted traffic volumes are out of the allowed limits (c_u and c_l). That is mean the results high variability and high bias.
The visualization process of this methodology is done by giving specific colour for each case of uncertainty. Table 1. represent the colours, characterizations, and limitations of the four

uncertainty cases. In this paper, histograms were used to visualize the probability of uncertainty for all cases. Fig. 1 shows an example for probability distribution shape for predicted traffic volume for a link (m), the coloured areas under the curve represents the probability of occurring all uncertainty cases.

Table 1 This table shows the classification for the four cases of uncertainties

	$(c_l \leq \chi_m) \wedge (\chi_m \leq c_u)$	$(c_l > \chi_m) \vee (\chi_m > c_u)$
$(\chi_m - \sigma_m \leq \chi_m) \wedge (\chi_m \leq \overline{\chi}_m + \sigma_m)$	Case I Low variability – Low bias	Case II Low variability – High bias
$(\overline{\chi}_m - \sigma_m > \chi_m) \vee (\chi_m > \overline{\chi}_m + \sigma_m)$	Case III High variability – Low bias	Case IV High variability – High bias

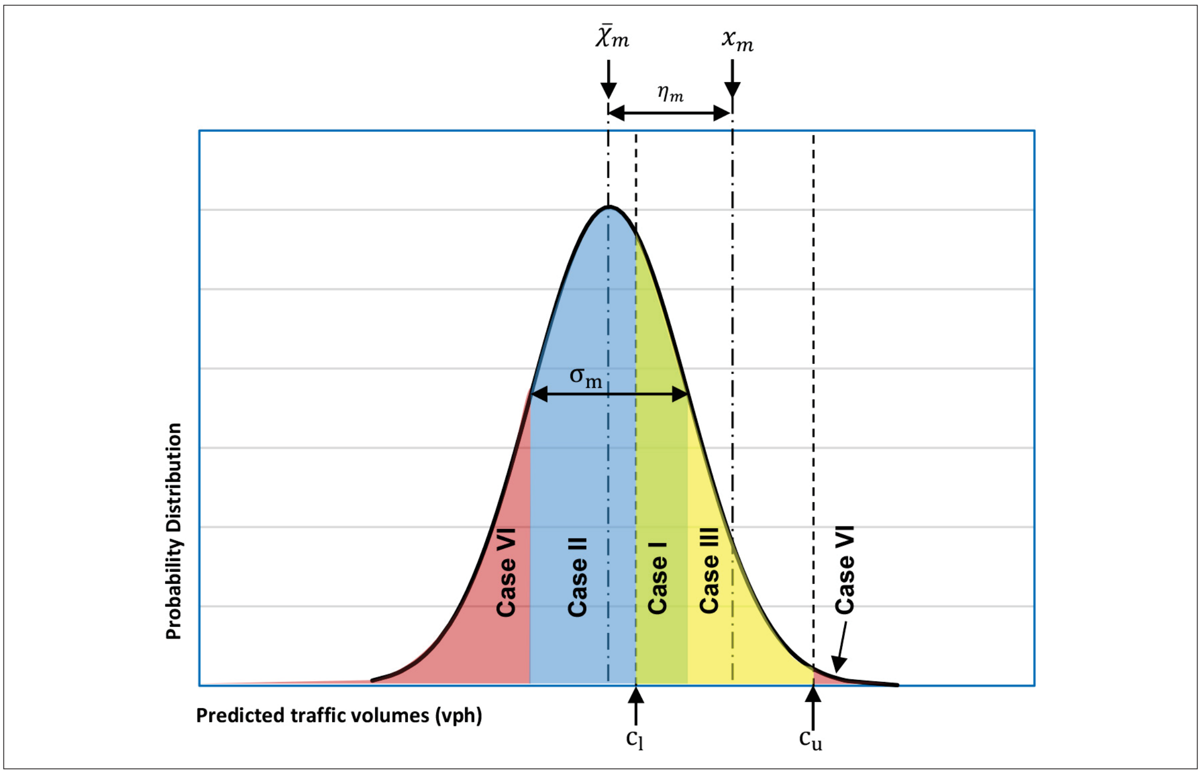

Figure 1 This figure shows the probability distribution of predicted traffic volumes addressed with the uncertainty cases.

4 Case study

4.1 Data

This methodology was applied in a small city located in Hungary (Ajka) as shown in Fig. 2. In the MC simulation processes; three types of input probability distributions applied on OD parameters; (1) Normal Distribution, (2) Lognormal Distribution, and (3) Extreme value distribution. And, for each probability distribution, 10 standard deviations were examined variated between 0.05 to 0.5. The total number of iterations used in this simulation; (3 probability distribution) x (10 standard deviations for each probability distribution) x (1000 iterations for each standard deviation) equal to (30000 iterations). Thus, this type of simulation is expensive; the time was spent to complete this simulation is around (300 hours), using a computer has a specification; 8th Generation Intel CPU Core i7-8700K – 6 Cores – 3.70GHz, RAM 32GB. The simulation time depends on; the number of links of the transport network, number of OD

parameters, and number of iterations. The simulation time depends on; the number of links of the transport network, number of OD parameters, number of iterations and numbers of scenarios that will be studied.
The output of the simulation process is a traffic volume attributes for links. The number of attributes equal to the number of simulation iterations. The other challenge of this methodology is a big output data. This data needs to organize and categorize according to several aspects: type of links, the direction of traffic movement, Probability distribution, GEH, etc.

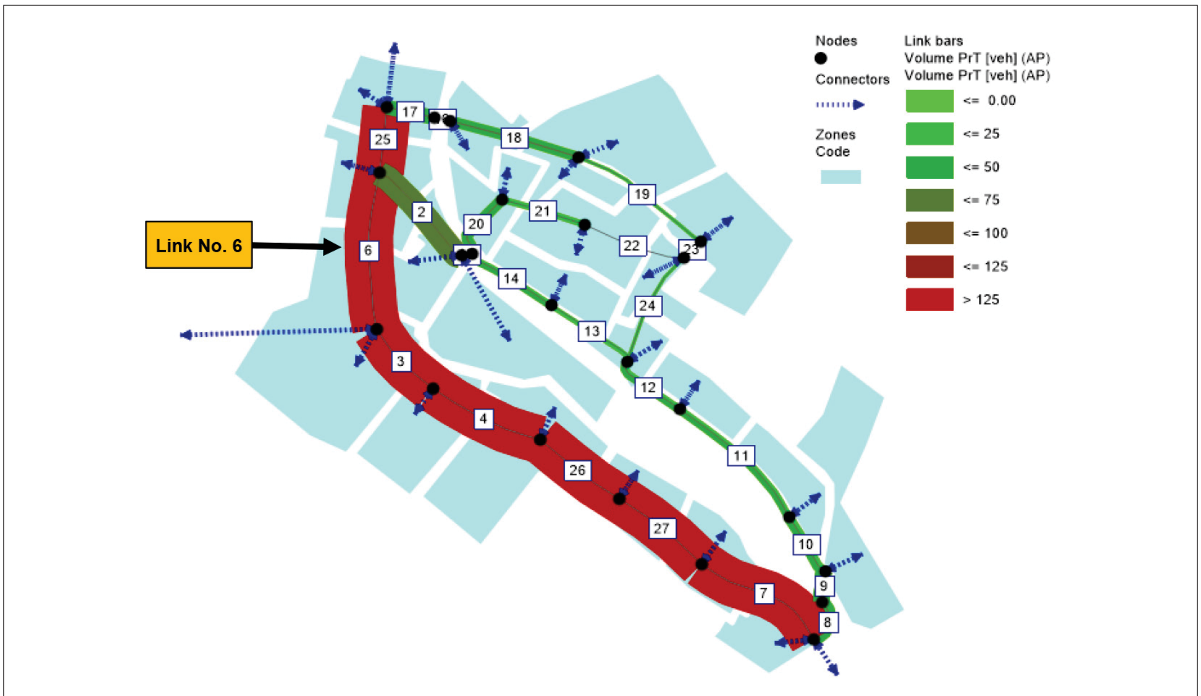

Figure 2 Represent the observed traffic volumes of the study area (Ajka, Hungary)

4.2 Visualizing of the predicted uncertainty

Three main scenarios were applied to this case study according to the probability distribution. Each main scenario was divided into ten sub-scenarios according to the standard deviation value. Then the developed methodology is applied to all sub-scenarios. The result of each sub-scenario presents a predictive uncertainty addressed by colours mirror the uncertainty cases for all links. For example, Fig. 3, Fig.4 and Fig. 5; displays the uncertainty status of the link No. 6. By this methodology, it is possible to monitor and identify which of the links that suffer from bias and unexpected change in traffic volumes in the event of a change in the conditions of traffic inputs i.e. different scenarios. In the same example: in normal distribution scenario as shown in Fig.3, we can see; (1) Case I (low variability – low bias) is decreased from 68 % in SD=0.05 to 50 % in SD=0.50. (2) Case II (low variability – high bias) start appearing from SD=0.35 and increased to 5 % in SD=0.40-0.50. (3) Case III (high variability – low bias) is decreased from 32 % in SD=0.05 to 0 % in SD=0.35-0.50. (4) Case III (high variability – high bias) start appearing from SD=0.15 to 44 % in SD=0.50. In a similar way, we can interpret Fig. 4 and Fig. 5. Fig. 6 represents the summary of the relationship between bias and variability of the three scenarios for link No.6. we can see; (1) In normal distribution scenario: a negative bias in traffic volume for all SD. (2) In lognormal distribution scenario: start from negative bias in SD=0.5-0.20 and changed to the positive bias after SD=0.25. (3) In extreme-value distribution scenario: a high drop negative bias in traffic volume increased directly with SD values.

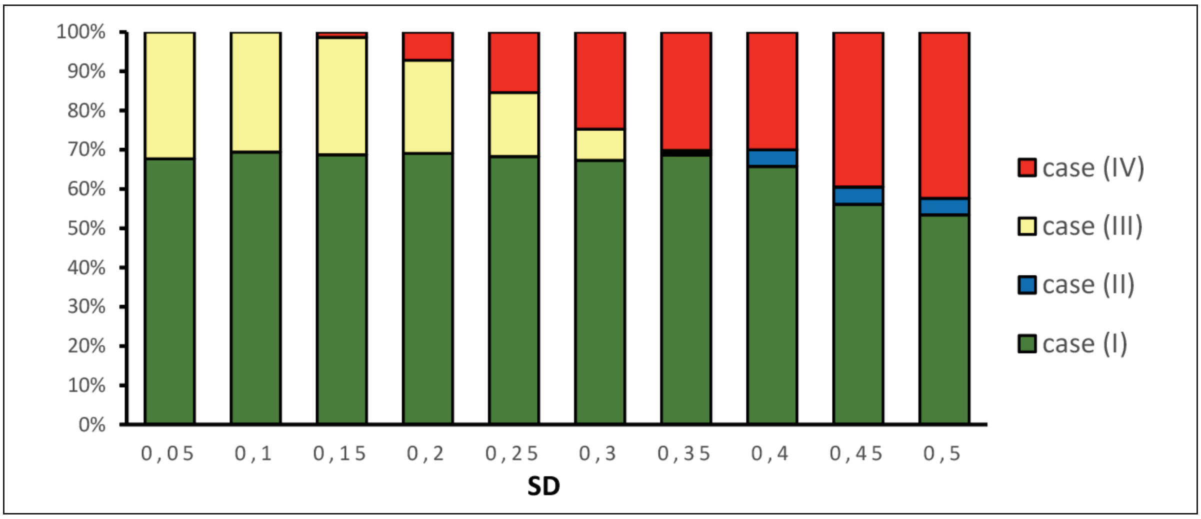

Figure 3 Represent the predicted uncertainty for the link No. 6, applying normal distribution

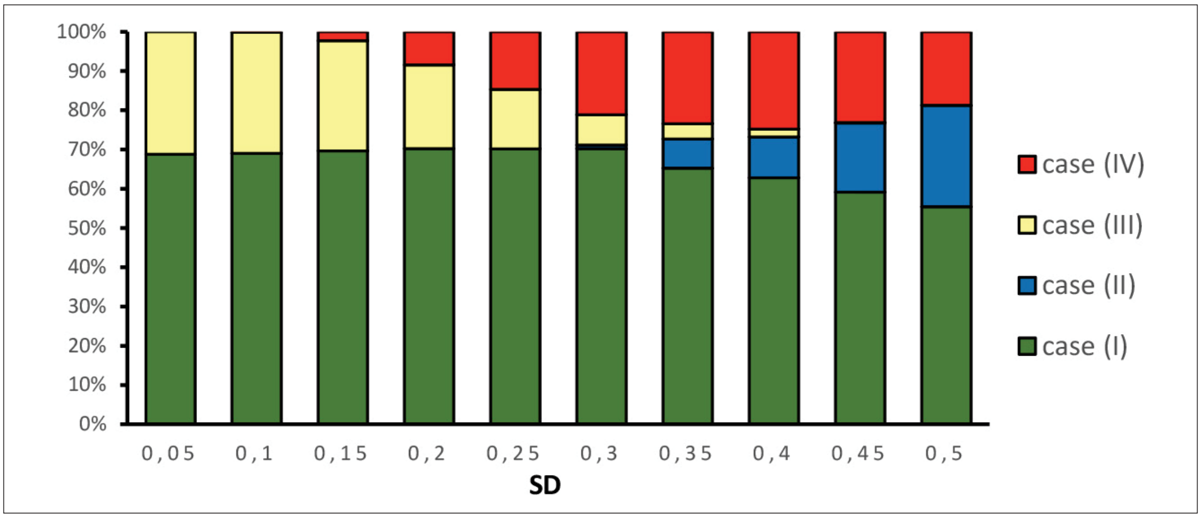

Figure 4 Represent the predicted uncertainty for the link No. 6, applying lognormal distribution

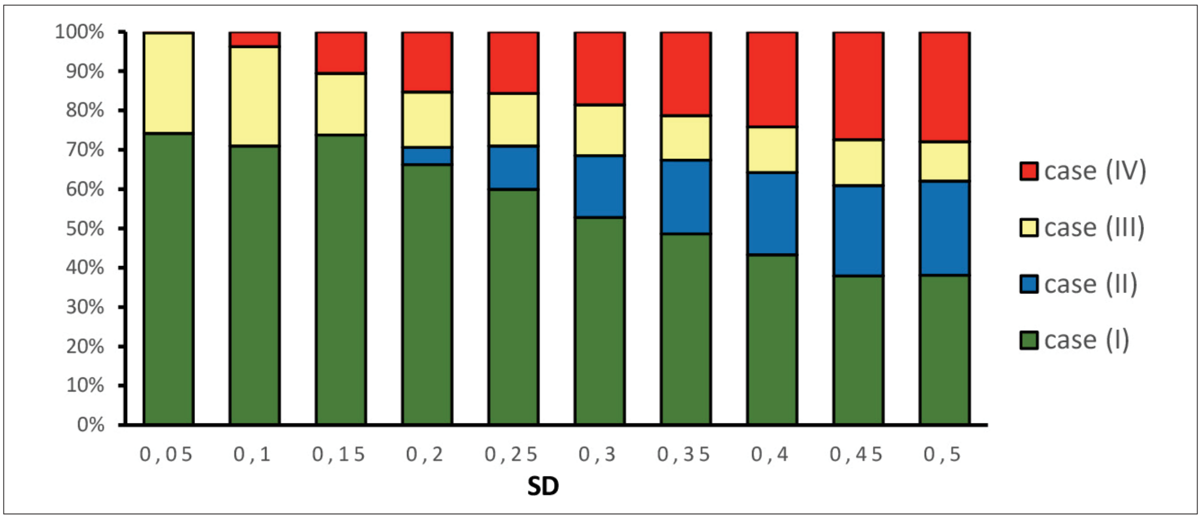

Figure 5 Represent the predicted uncertainty for the link No. 6, applying extreme value distribution

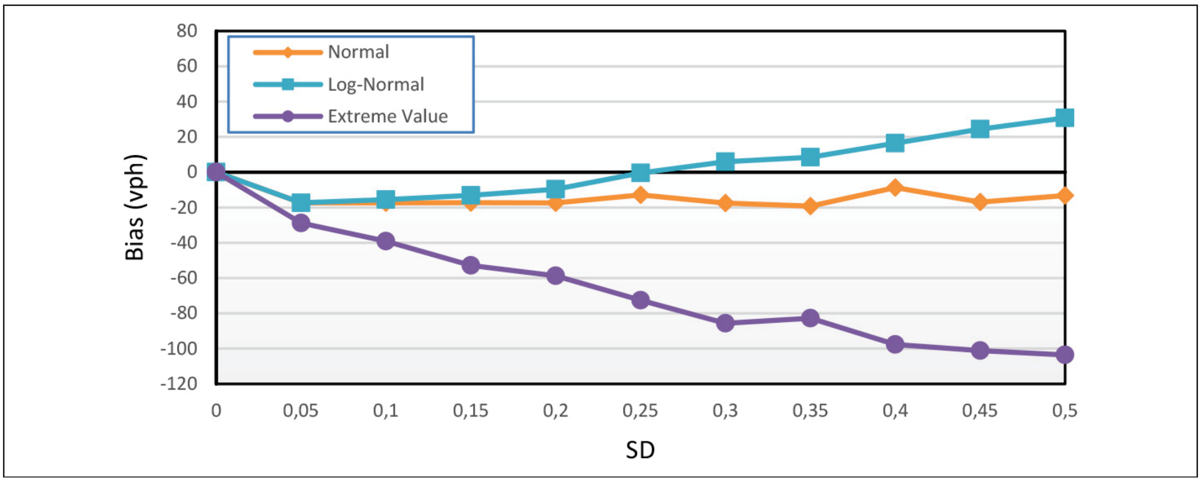

Figure 6 Represent the relationship between bias and input variability in deferent scenarios for the link No. 6

5 Conclusion

The aspect of information visualization is the examination of models that are not wholly understood, and visually characterizing the uncertainty helps in full understanding. In this paper, we introduce a new methodology to predict the uncertainty in traffic assignment and visualize the predicted uncertainty. This methodology was built on Monte Carlo simulation method, the result of this methodology is bias and variability of traffic volume comparing to the observed traffic volume for links of the transport network. This methodology is tested in a small study area using three type of probability distribution. The obtained results showed the applicability of this approach to predict the uncertainty in traffic assignment models. As well as, visualize the uncertainty in different scenarios varied according to the probability distribution types and the parameter's value for each probability distribution. The challenge of this method is time because this method needs a high number of iteration to get more precise results not less than 1000 iterations. This paper shows the importance of studying and visualizing the uncertainty in the input OD values of traffic assignment model. Increasing the variability of OD leads to decreasing on the reality of outputs of traffic volumes. And, the variability of lognormal distribution and extreme-value distribution scenarios gives a higher bias rather than the normal distribution scenario. Future research will consider applying different types of probability distribution in the same simulation according to types of zones. And investigate the posterior probability distributions in traffic volumes for transport network links.

References

[1] Gosink, L. et al.: "Characterizing and visualizing predictive uncertainty in numerical ensembles through Bayesian model averaging," IEEE Trans. Vis. Comput. Graph., vol. 19, no. 12, pp. 2703–2712, 2013.

[2] Jeong, C., Pang, A.: "Uncertainty Visualization of Hierarchical Information Structure Using Reconfigurable Disc Trees," 1997. ftp://ftp.soe.ucsc.edu/pub/reinas/papers/urdt.pdf.gz

[3] Ayyub, B., Klir, G.: Uncertainty Modeling, and Analysis in Engineering and the Sciences. 6000 Broken Sound Parkway NW: Chapman & Hall/CRC, 2006.

[4] Cheung, K., Polak, J.: "A Bayesian approach to modeling uncertainty in transport infrastructure project forecasts," Eur. Transp. Conf. 2009, pp. 1–22, 2009.

[5] Transport, E., et al.: "Verification, Validation and Uncertainty Quantification Workflow in CASL," Transport, vol. 5, no. SAND2010-234P, pp. 1–27, 2005.

[6] Departmant for Transport, UK, "Highway Assignment Modelling," London, 2014. https://www.gov.uk/guidance/transport-analysis-guidance-webtag.