

Research Article

Predicting and Visualizing the Uncertainty Propagations in Traffic Assignments Model Using Monte Carlo Simulation Method

Mundher Seger  and Lajos Kisgyörgy

Highway and Railway Department, Budapest University of Technology and Economics, 1111 Budapest, Hungary

Correspondence should be addressed to Mundher Seger; mundher.seger@gmail.com

Received 20 July 2018; Accepted 10 September 2018; Published 1 October 2018

Guest Editor: Lele Zhang

Copyright © 2018 Mundher Seger and Lajos Kisgyörgy. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Uncertainty can be found at all stages of travel demand model, where the error is passing from one stage to another and propagating over the whole model. Therefore, studying the uncertainty in the last stage is more important because it represents the result of uncertainty in the travel demand model. The objective of this paper is to assist transport modellers in perceiving uncertainty in traffic assignment in the transport network, by building a new methodology to predict the traffic flow and compare predicted values to the real values or values calculated in analytical methods. This methodology was built using Monte Carlo simulation method to quantify uncertainty in traffic flows on a transport network. The values of OD matrix were considered as stochastic variables following a specific probability distribution. And, the results of the simulation process represent the predicted traffic flows in each link on the transport network. Consequently, these predicted results are classified into four cases according to variability and bias. Finally, the results are drawn into figures to visualize the uncertainty in traffic assignments. This methodology was applied to a case study using different scenarios. These scenarios are varying according to inputs parameters used in MC simulation. The simulation results for the scenarios gave different bias for each link separately according to the physical feature of the transport network and original OD matrix, but in general, there is a direct relationship between the input parameter of standard deviation with the bias and variability of the predicted traffic flow for all scenarios.

1. Introduction

Forecasting of travel demand represents the fundamental step of planning and management of transportation facilities [1]. These forecasts are subjected to various sources of errors including error in the measurement of input data, error in the estimated value of model parameters, and error in the specification of the underlying models themselves; also the model itself may be stochastic, and the scenarios adopted for model forecasting may not necessarily be compatible with the real evolution of the transport system [2, 3]. Studying of the uncertainty of travel demand not only is aimed at stripping the values of traffic flow from errors but mainly aims to know the effect of these errors regarding the nature of the variability and bias on traffic flow and the likelihood of occurrences; this will be discussed in this study.

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 [1, 4]. However, it is impossible to give an exact prediction; no model can be constructed to provide 100% accurate predictions of the future behaviour 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 [5, 6]. As a result, any method used for prediction should include an assessment of the uncertainty in the predicted values. Uncertainty in traffic forecasts is inherent significantly in travel demand models. Accordingly, a solid understanding of the degree of uncertainty is helpful not only

for the model designation process but also for the sampling process and the value of precise model results for policy scenario development and strategy evaluation [3, 7, 8].

Although many researchers have studied uncertainty in the travel demand models, only a few studies analysed the impact of the error propagation in whole four-stage sequential transport model frameworks. For example, Zhao and Kockelman [9] found that the uncertainty increased throughout the first three model steps and declined in the traffic assignment step. They argued that this reduction might be due to the network congestion effects on the trip assignment equilibrium system, meaning that capacity limitations might reduce the variability of the outcomes at the link traffic flows. However, they also found out that the reduction of uncertainty in the traffic assignment might also be the result of the accumulation on the same links of independent trips related to various OD pairs. De Jong et al. [6, 10] examined the uncertainty of the Dutch national model and used the standard deviations and correlations of 20-year moving averages of some input data to obtain values from a multivariate normal probability distribution function; they found that congestion reduced final model output uncertainty but only to a minor degree. Ziems et al. [11] analysed the effects of the randomness on different traffic characteristics. They observed that the variability of traffic flow is larger during the congested period of the day and there is a greater variability in the individual corridor rather than aggregated subareas. Manzo et al. [12, 13] stated that the congestion in transportation network does not show a high uncertainty effect on the final output of transport model. Hence, the final uncertainty of traffic flow for links with higher volume/capacity ratio showed a lower dispersion around the base uncertainty value. Rasouli and Timmermans [14, 15] investigated uncertainty of OD matrix using the Dutch national transport model and found that higher levels of traffic flow result in lower levels of uncertainty for different model output. Thus, the researcher emphasized that the degree of uncertainty grows higher if the focus of attention shifts from aggregate system performance indicators through OD matrix to disaggregated space-time sequences and performance indicators.

Future year travel demand forecasting is not an exact science, and there are complicated underlying mechanisms that inherently generate uncertainty in the forecasts. Modelling these complicated mechanisms requires numerous variables and behavioural components whose variability may be poorly determined or simply ignored. In this case, it is illogical to take a single view of the future without considering the uncertainty in travel demand modelling. Thus, to provide more efficient and reliable transport solutions for future, transport analysts and planners have to observe and predict uncertainty in transport systems [3, 8, 16].

Uncertainty becomes relevant in transportation modelling not only in case of diverging views such as if risks are very high if the policy is controversial and if there are concerns about model limitations, but also in case of certain views: several points estimated based on different scenarios are given to account for uncertainty [17]. 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 [8]. 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 modelling process extremely challenging [18, 19]. 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 [20].

This study presents a new methodology for exploring nature of uncertainty propagation deriving from input OD matrix in a four-stage transport model using Monte Carlo simulation method. MC method was used to generate data for OD matrix for three types of probability distributions (i.e., normal, lognormal, and extreme value). Then, the generated OD matrices were entered into the VISUM software to predict traffic assignment attributes. Moreover, the final step was analysing the visualizing the variability and bias of predicted for traffic flow in the transport network links.

The current study contributes to the present literature of uncertainty in transportation planning, primarily by (i) developing a methodology to predict the uncertainty in transport network depending on the variability of input OD matrix, (ii) examining the uncertainty impact on transport model by using different probability distributions in the input data, (iii) adopting a new method to visualize the uncertainty according to a probability of occurrences, and (iv) investigating the probability distributions of output traffic flow on transport network depending on the probability distributions in input data.

2. Methodology

A new methodology was developed 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 calculated by traffic assignment models. This methodology consists of five stages: (i) input stage; (ii) MC simulation process stage; (iii) analysis of predicted traffic flow stage; (iv) predictive uncertainty stage; (v) uncertainty visualization stage. The relationships connecting these stages of the methodology are presented in Figure 1, and the mathematical and logical computations of this methodology are illustrated in an algorithm (see Algorithm 1).

2.1. Input Stage. The principal task in predictive modelling is to estimate the behaviour of a modelling function, in this case, traffic assignment function $f_{TO \rightarrow D}$ defined to calculate the traffic flow between Oregon zone (O) to Destination zone (D). This work addresses the case where $f_{TO \rightarrow D}$ can be calculated at a finite set of iterations (N_i). Monte Carlo simulation method was used to generate the input data with the parameters: standard deviation (σ) and mean (μ) to produce finite OD matrices. The required data consist of three parts: the first part is setting physical features of the transport network (TN) and traffic analysis zones (TAZ)

| | |
|---|---|
| Require: $Z_{i=1,n}; \forall Z_i \in \mathbb{N}$ | ▷ Zone definition. |
| $O_{i=1,n}; O_i \subset Z$ | ▷ Origin definition. |
| $D_{i=1,n}; D_i \subset Z$ | ▷ Destination definition. |
| $n; \forall n \in \mathbb{N}$ | ▷ number of zones. |
| $OD_{(n,n)}; \forall OD_{(n,n)}$ | ▷ Matrix of Origin Destination (OD). |
| $OD_{(i,i)}^{obs}; \forall OD_{(i,i)}^{obs} \in \mathbb{R}$ | ▷ Observed value of OD pair. |
| $f_{T \rightarrow D}$ | ▷ Traffic Assignment function. |
| $x_{m=1,M}; \forall x_m \in \mathbb{R}$ | ▷ Observed values of traffic assignment function. |
| $M; \forall M \in \mathbb{N}$ | ▷ Number of Links of the transport network. |
| PD | ▷ Define the Probability Distribution type. |
| $\mu, \sigma; \forall \mu, \sigma \in \mathbb{R}$ | ▷ Parameters range for the probability distribution. |
| $N_s; \forall N_s \in \mathbb{N}$ | ▷ Number of iterations. |
| (01) $\{\varphi_k\}_{k=1}^{N_s} \leftarrow MC[PD(\mu, \sigma)]$ | ▷ Get samples by the defined probability distribution and parameters using Monte Carlo procedure. |
| (02) For $k = 1$ to N_s do | |
| (03) For $i = 1$ to n do | |
| (04) For $j = 1$ to n do | |
| (05) $OD_{(i,i)}^k \leftarrow OD_{(i,i)}^{obs} * \{\varphi_k\}$ | ▷Get OD matrix randomly. |
| (06) end for | |
| (07) end for | |
| (08) end for | |
| (09) For $k = 1$ to N_s do | ▷Run Visum. |
| (10) $\chi_m^k \leftarrow f_{T \rightarrow D}(OD_{(i,i)}^k); \forall \chi_m^k \in \mathbb{R}$ | ▷Get result attribute from traffic assignment function. |
| (11) end for | |
| (12) For $m = 1$ to M do | |
| (13) For $k = 1$ to N_s do | |
| (14) $GEH_m^k \leftarrow \sqrt{\frac{2(\chi_m^k - x_m)^2}{(\chi_m^k + x_m)}}$ | ▷Get the value of GEH statistic for each iteration. |
| (15) end for | |
| (16) $GEH_m \leftarrow \frac{\sum_{k=1}^{N_s} GEH_m^k}{N_s}; \forall GEH_m \in \mathbb{R}$ | ▷Get the average value of GEH statistic for each link. |
| (17) $\bar{\chi}_m \leftarrow \frac{\sum_{k=1}^{N_s} \chi_m^k}{N_s}$ | ▷Get the average for each link. |
| (18) $\eta_m \leftarrow (\bar{\chi}_m - x_m); \forall \eta_m \in \mathbb{R}$ | ▷Get the total bias for each link. |
| (19) $\sigma_m \leftarrow \sqrt{\frac{1}{N_s} \sum_{k=1}^{N_s} (\chi_m^k - \bar{\chi}_m)^2}; \forall \sigma_m \in \mathbb{R}$ | ▷Get the standard deviation for each link. |
| (20) $c_l \leftarrow x_m + 6.25 - \frac{\sqrt{100x_m + 156.25}}{2}$ | ▷Get the bias lower limit for each link. |
| (21) $c_u \leftarrow x_m + 6.25 + \frac{\sqrt{100x_m + 156.25}}{2}$ | ▷Get the bias upper limit for each link. |
| (22) end for | |
| (23) For $m = 1$ to M do | |
| (24) If $GEH_m < 5$ | Then end for |
| (25) For $k = 1$ to N_s do | ▷Check the error and the bias for each iteration. |
| (26) If $\bar{\chi}_m - \sigma_m \leq \chi_m^k \leq \bar{\chi}_m + \sigma_m$ | Then low variety |
| | Else high variety |
| (27) If $c_l \leq \chi_m^k \leq c_u$ | Then low bias |
| | Else high bias |
| (28) end for | |
| (29) end for | |
| (30) End | |

ALGORITHM 1: Prediction of uncertainty in traffic assignment using Monte Carlo simulation.

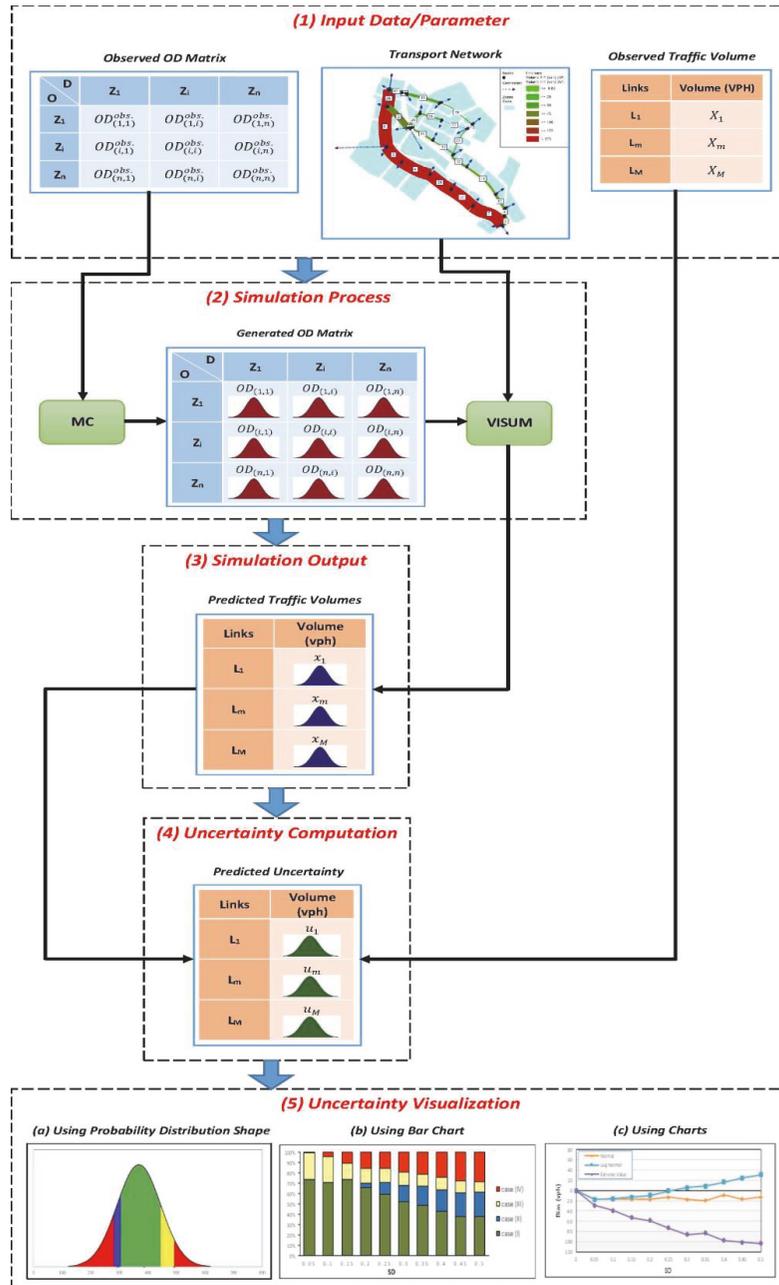


FIGURE 1: The methodology of uncertainty prediction.

by VISUM. The second part is defining the observed OD matrix. Finally, the third part is finding the observed traffic flow attribute either by counting the real value of traffic flow or by running VISUM to get it. Monte Carlo simulation method is used to generate the OD matrices. Therefore, the required parameters for this process are standard deviation (σ), mean (μ), and the number of iterations (N_c). Moreover, the type of probability distribution (PD) must be chosen at the beginning. The required input data and parameters are explained in Algorithm 1.

2.2. Monte Carlo Simulation Stage. Monte Carlo (MC) method or random sampling method is a division of computational mathematics. It is created from the mathematics

concepts for “the frequency approximates the probability.” When the solution for a problem is the occurrence probability of a certain event or an expected value of any variants, a testing method is used to obtain the occurrence frequency of an event or the average value for these variants. MC method is based on the probability model and according to the described process by this model. The results of the simulation test are approximate solutions [21]. MC method plays a fundamental role in characterization and quantification of uncertainty. When the accurate calculation of output uncertainties is 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 [22].

In this stage, the simulation code has been written using both Matlab software and Component Object Model (COM) of VISUM software in purpose of predicting traffic flows on the transport network. The simulation code involves two processes: the first is the generation of OD matrix (OD^k) by Monte Carlo method for a given type of probability distribution via observed OD matrix (OD^{obs}).

$$\{\varphi_k\}_{k=1}^{N_s} = MC [PD(\mu, \sigma)] \quad (1)$$

$$OD^k = OD^{obs} * \{\varphi_k\} \quad (2)$$

where $\{\varphi_k\}$ is sample value using Monte Carlo procedure, MC is Monte Carlo procedure, k is iteration number, N_s is number of iterations, PD is probability distribution, μ is mean parameter of PD , σ is standard deviation parameter of PD , OD^k is generated OD matrix, and OD^{obs} is observed OD matrix.

The second process is the process of entering generated OD matrix and the transport network (TN) in the VISUM to produce the traffic flow attribute for the road network from traffic assignment function.

$$\chi_{m=1,M}^k = f_{TO \rightarrow D}(OD^k) \quad (3)$$

where $f_{TO \rightarrow D}$ is traffic assignment function, $\chi_{m=1,M}^k$ is traffic flow attribute, m is link number, and M is number of links.

This simulation is repeated for a finite number of iterations N_s specified by the researcher. With a note, increasing number of iterations N_s leads to improving the accuracy of results but extending simulation time.

2.3. Analysis of Predicted Traffic Flow Stage. One of the most important aspects of a simulation study is an analysis of simulation experiments. In this stage, the results of the simulation process are analysed for both statistical characterization and examining the fitting for the distribution (Goodness-of-Fit Test). The outcomes of the simulation process are traffic flow attributes for all links on the transport network. The number of obtained attributes is equal to the number of iterations N_s used in the simulation.

The statistical characteristics of predictive traffic flow are defined by two parameters: the average value results (μ) and the variability (σ). In this case, the average traffic flow for all links in transport network is shown in (4). Finally, standard deviation represents the variability of results in (5).

$$\bar{\chi}_m = \frac{\sum_{k=1}^{N_s} \chi_m}{N_s} \quad (4)$$

$$\sigma_m = \sqrt{\frac{1}{N_s} \sum (\chi_m - \bar{\chi}_m)^2} \quad (5)$$

where χ_m is traffic flow for the link (m), $\bar{\chi}_m$ is average traffic flow for the link (m), and σ_m is the standard deviation in traffic flow for the link (m).

2.4. Predictive Uncertainty. Meaningful quantification of data and structural uncertainties in conceptual travel demand modelling is a significant challenge in the transportation system. Heterogeneous, insufficient, unstable, and unsteady characterize data used to build OD matrix because of the difficulty in measurements, and the nature of the individual behaviour leads to variation of the parameter of OD matrix in one side. Besides limitations and lack of information in mathematical concepts and the structure in four-stage sequential transport model frameworks lead to variation in traffic flow in transport networks.

The 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 methodology, the uncertainty has been predicted by joining the traffic flow attribute from input stage (first stage) and the traffic flow attributes obtained by simulation processes (fourth stage).

The mathematical operations in this stage include both calculating bias in traffic flow as (6) and determining the limitations of allowed biases in the transport network.

$$\eta_m = (\chi_m - x_m) \quad (6)$$

where η_m is bias in traffic volume for the link (m), χ_m is predicted traffic flow for the link (m), and x_m is observed traffic flow for the link (m).

In any uncertainty quantification process, setting limits for the predictive uncertainty is required to increase understanding of the researchers to models behaviour in both bias and predictive variability. The GEH statistic has been used as a limitation of the bias in this study. The GEH statistic is a form of Chi-squared statistic that can be used to compare observed and modelled counts as (7) [1]. It is helpful for these comparisons because it is sophisticated for relative and absolute errors. And, as the estimation of the standard error, the standard deviation statistic was adopted as a limitation for the variability in traffic volumes.

$$GEH = \sqrt{\frac{(M - C)^2}{(M + C)/2}} \quad (7)$$

where M is the modelled traffic flow and C is the observed traffic flow.

GEH statistic bands less than 5 [1] are used to explain bias limit for each link.

$$GEH = \sqrt{\frac{(\chi_m - x_m)^2}{(\chi_m + x_m)/2}} = 5 \quad (8)$$

By solving (8), the upper and the lower bias limit in predicted traffic flow are as follows:

$$c_m = x_m + 6.25 - \frac{\sqrt{100x_m + 156.25}}{2} \quad (9)$$

$$c_{u_m} = x_m + 6.25 + \frac{\sqrt{100x_m + 156.25}}{2} \quad (10)$$

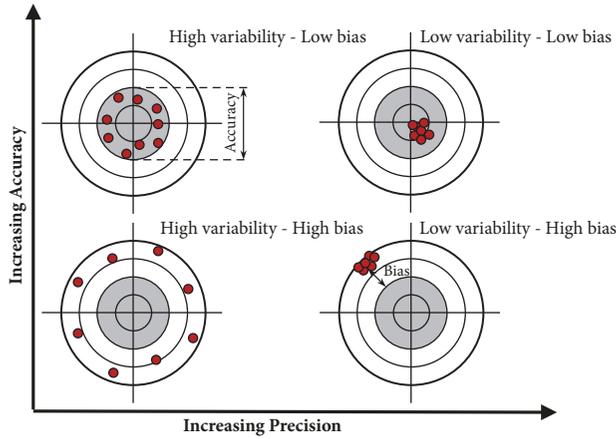


FIGURE 2: The relationship between precision and accuracy.

where χ_m is predicted traffic flow for the link (m), x_m is observed traffic flow for the link (m), c_l is lower accurate limit for the link (m), and c_u is upper accurate limit for the link (m).

The allowed accuracy and variability are presented in (11) and (12), respectively.

$$\text{Accuracy allowable for the link } (m) = c_u - c_l \quad (11)$$

$$\text{Variability allowable for the link } (m) = \mp \sigma \quad (12)$$

If the predicted traffic flow values lie within the standard deviation, then these are precise. Otherwise, these predicted values are imprecise (see (13) and (14)).

$$\begin{aligned} \text{If } \bar{\chi}_m - \sigma_m \leq \chi_m^k \\ \leq \bar{\chi}_m + \sigma_m, \text{ then low variety (precise)} \end{aligned} \quad (13)$$

$$\begin{aligned} \text{If } \bar{\chi}_m - \sigma_m > \chi_m^k \vee \chi_m^k \\ > \bar{\chi}_m \\ + \sigma_m, \text{ then high variety (imprecise)} \end{aligned} \quad (14)$$

In the same way, if the predicted traffic flow values lie within upper and lower bias limits, then these are accurate. Otherwise, these predicted values are biased (see (15) and (16)).

$$\text{If } c_l \leq \chi_m^k \leq c_u, \text{ then low bias (accurate)} \quad (15)$$

$$\text{If } c_l > \chi_m^k \vee \chi_m^k \leq c_u, \text{ then high bias (inaccurate)} \quad (16)$$

To interpret the concept of precision (i.e., the variability of values) and the accuracy (i.e., the bias of values), see Figure 2.

2.5. Visualizing Predictive Uncertainty. The last stage of this methodology is uncertainty visualization. Uncertainty visualization is endeavouring to display data together with additional uncertainty information. These visualizations present a more complete and accurate interpretation of data for

researchers to analyse [23]. Thus, visualization is a useful method for addressing many forms of information uncertainty, and it is a helpful approach to the investigation and communicating of large data sets [24].

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 [25]. Consequently, visualizing the uncertainty is essential for risk analysis and decision-making tasks. However, it is still a challenge, because describing the uncertainty is a complex 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 [25, 26]. 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 [26, 27].

There are different methods used to visualize uncertainty: statistical and probability-based visualization, point and global visualization, used colours, financial visualization, icons, ontology, lexicon, etc. [20]. In this methodology, 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 with a bar chart, probability distributions, and traditional charts represented by random variables. This method demonstrates the central tendency, dispersion, skewness, and modal characteristics of a random variable. In this methodology, two steps are used to visualize the uncertainty of the predicted traffic flow.

The first step classifies the predicted traffic flow values into four cases according to bias and variability.

Case I (accurate and precise (low predictive uncertainty)). This case occurred when the predicted traffic flows (χ_m) are close to the mean prediction ($\bar{\chi}_m$), i.e., within the standard deviation ($\pm\sigma$). And predicted traffic flows are within the accuracy limits (c_u and c_l). This means that the results are of low variability and low bias.

Case II (accurate and imprecise (moderate predictive uncertainty)). This case occurred when the predicted traffic flows (χ_m) are close to the mean prediction ($\bar{\chi}_m$), i.e., within the standard deviation ($\pm\sigma$). And predicted traffic flows are out of the accuracy limits (c_u and c_l). This means that the results are of low variability and high bias.

Case III (inaccurate and precise (high predictive uncertainty with ensemble agreement)). This case occurred when the predicted traffic flows (χ_m) are far from the mean prediction ($\bar{\chi}_m$), i.e., out the standard deviation ($\pm\sigma$). And predicted traffic flows are within the accuracy limits (c_u and c_l). This means that the results are of high variability and low bias.

TABLE 1: This table shows the classification of the four cases of uncertainties.

| | | Variability limitations (Precision) | |
|-----------------------------|------|---|---|
| | | Low | High |
| Bias limitations (Accuracy) | Low | Case I Low variability - Low bias (Low predictive uncertainty) | Case III Low variability - High bias (High predictive uncertainty with ensemble estimates) |
| | High | Case II Low variability - High bias (Moderate predictive uncertainty) | Case IV High variability - High bias (High predictive uncertainty with divergent estimates) |

Case IV (inaccurate and imprecise (high predictive uncertainty with divergent estimates)). This case occurred when the predicted traffic flows (χ_m) are far from the mean prediction ($\bar{\chi}_m$), i.e., out the standard deviation ($\pm\sigma$). And predicted traffic flows are out of the accuracy limits (c_u and c_l). This means that the results are of high variability and high bias.

The second step of the visualization process is giving specific colour for each case of uncertainty. Table 1 illustrates the colours, characterizations, and limitations of the four uncertainty cases. Figure 3 demonstrates the probability density curves of predictive uncertainty cases. In practice, all the four cases can appear with different likelihood proportions. Figure 4 shows an example of a probability density curve for predicted traffic flow for a link (m). The coloured areas under the curve represent the predictive uncertainty cases.

The major role of this visualization of uncertainty is to give information about the level of uncertainty to the decision-makers. Based on these cases they can see how reliable the predictions of the model are and if they can accept the risks relating to the given predictive uncertainty. If predictions that are more precise are needed, then sensitivity analysis helps to identify the input data most dominantly influencing the predictive uncertainty. Then these input data should be measured more precisely.

3. Case Study

The developed methodology has been applied in a small city, Ajka, located in Hungary. The results of case study were presented earlier in [27] for demonstrating the effects of variability of input variables on the model results. Here we used the case study and its results to describe the implementation of the methodology. Figure 5 presents the transport network (NT) and traffic analysis zones (TAZ) for the study area. This case study applied the developed methodology step by step, as follows.

3.1. Input Data and Parameters. The required data include information about the physical feature of the study area, where the number of TAZ=25 and the number of links (M) = 50. Also collecting 625 OD pairs represents the travel distribution between zones (i.e., OD^{obs}). Besides, in this case,

the travel assignment attribute has been gotten by executing the VISUM (i.e., $x_{m=1,M}$).

In this case, three types of probability distributions (PD) have been used: (1) normal distribution, (2) lognormal distribution, and (3) extreme value distribution. Ten values of standard deviations (σ) have been examined for each distribution; these values are ranged from 0.05 to 0.50, while the observed values of the OD matrix have been used as mean values (μ).

3.2. Monte Carlo Simulation. In this simulation, 30 scenarios have been experimented. These scenarios have been grouped into three groups according to probability distribution (PD) type. And for each group, the parameter of standard deviation (σ) has been changed for each scenario. Besides, 1000 iterations have been used per each scenario. Thus, the total number of iterations used in the whole simulation process is as follows: (30 scenarios) x (1000 iterations for each scenario) equal to (30000 iterations). Therefore, this simulation process is expensive; the time spent to complete this simulation is around 300 hours, using a computer with the following specifications: CPU from Intel 8th generation, Core i7 8700K, 6 Cores, 3.70GHz, RAM 32GB, DDR 4. The simulation time of this methodology depends on (1) the number of links on the TN, (2) the number of TAZ, and (3) the number of iterations N_s .

3.3. Analysis of Predicted Traffic Flow. The results of the simulation process are represented by attributes of traffic flow for the links of the transport network. The number of these attributes is equal to the number of simulation iterations. Consequently, the outcomes for each link have been studied separately, by finding the statistical parameters for predictive traffic flows: both of the average value ($\bar{\chi}_m$) and the variability (σ_m). Because there are 50 links on the transport network in this study case, only one link will be presented (we chose link No. 6). For example, Figures 6, 7, and 8 represent three scenarios having a different probability distribution and different standard deviations for the link No. 6.

3.4. Predictive Uncertainty. The obtained data from the previous stage has been processed by Excel-sheet to find the uncertainty in traffic flows. The mathematical calculations

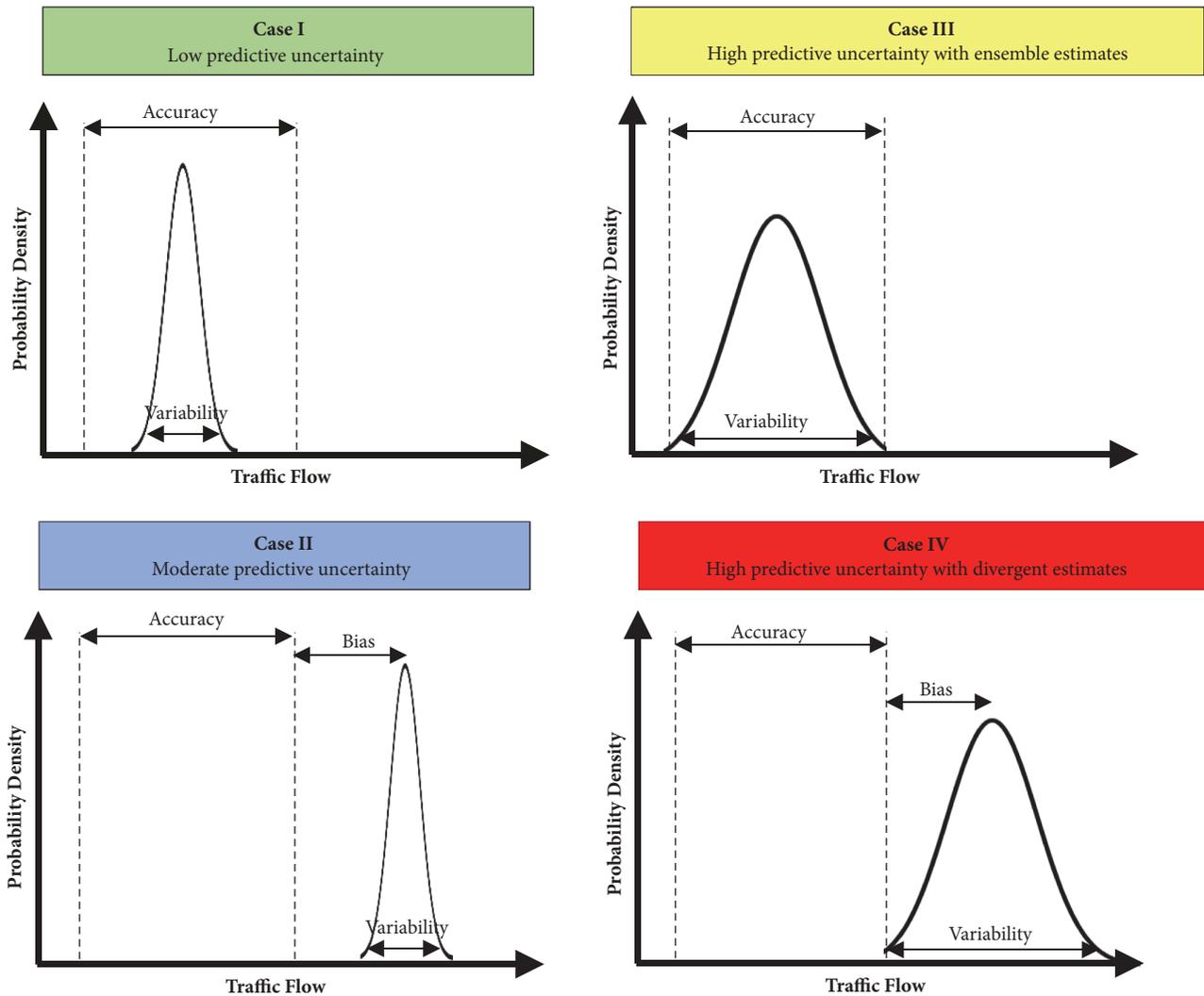


FIGURE 3: The probability density curves of the predicted traffic flow for the four cases.

determine the bias and variability for the predicted traffic flows, as well as the accurate limits for each link separately. For the link No. 6, the lower accurate limit (c_l) = 295 vph and the upper accurate limit (c_u) = 492 vph. Figure 6 shows the percentage of bias in the predicted traffic flows associated with input standard deviations. It can be observed that increasing the standard deviation of inputs leads to increasing bias for the predicted traffic flows, for normal distribution scenarios and lognormal distribution scenarios; there are no or few biases for standard deviations less than 0.10, and the biases increased to reach about 42% in standard deviation equal to 0.50, while for the extreme value distribution scenarios the biases increased in all standard deviations to reach about 54% in standard deviation equal to 0.50. Figure 7 shows the percentage of converging degree for predicted traffic flows joined with input standard deviations. It can be noted that the variability of the predicted traffic flows for normal distribution scenarios ranged between 67.1% and 70.3% for standard deviations less than or equal to 0.40 and decreased after that to reach 57.6% in the standard deviation

equal to 0.50. And, the variability of the predicted traffic flows for lognormal distribution scenarios ranged between 67.9% and 70.3% for standard deviations less than or equal to 0.40 and increased after that to reach 76.4% in the standard deviation equal to 0.50, while the variability of the predicted traffic flows for extreme value distribution scenarios ranged between 67.2% and 72.0% for standard deviations less than or equal to 0.35 and decreased after that to reach 61.4% in the standard deviation equal to 0.50.

3.5. Visualizing Predictive Uncertainty. Finally, the probability of the uncertainty cases for the predicted traffic flows in transport network has been visualized into bar-charts by merging Figures 6 and 7. This visualization provides transport planners and engineers with the possibility to monitor and identify which of the links suffers from bias and unexpected change in traffic volumes in the event of a change in the conditions of traffic parameters and experiment different scenarios on the transport network.

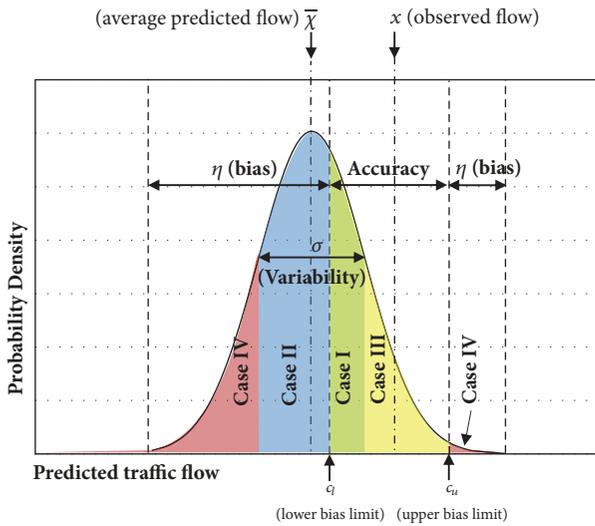


FIGURE 4: The probability density of predicted traffic flows addressed with the uncertainty cases.

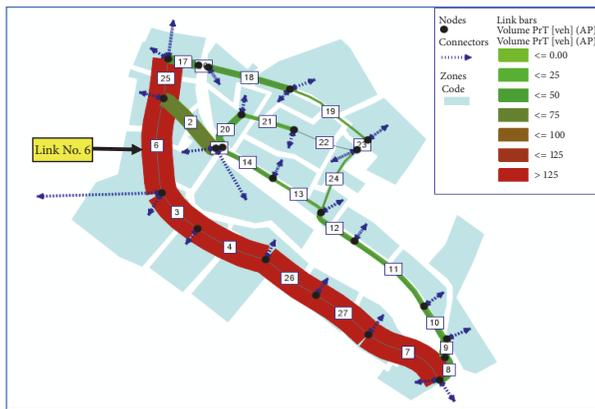


FIGURE 5: Study area (Ajka, Hungary), showing observed traffic flows on the transport network [27].

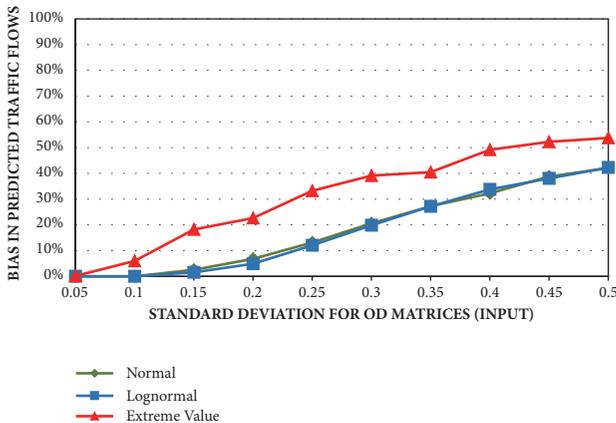


FIGURE 6: The relationship between the standard deviation for OD matrices (input) and the bias in the predicted traffic flows for link No. 6.

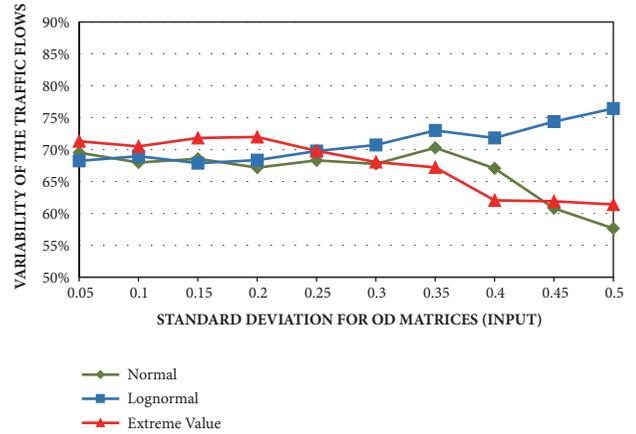


FIGURE 7: The relationship between the standard deviation for OD matrices (input) and the variability of the predicted traffic flows (output) for link No. 6.

The coloured bar-charts give a whole idea about the probability ratios of the uncertainty cases. In general, the percentage of green colour (Case I) means the probability of predicted traffic flow will lie within the allowed limits of accuracy and variability. And the percentage of yellow colour (Case III) means the probability of predicted traffic flow will lie still within the allowed accuracy but outside the allowed variability. Both percentages of green colour (Case I) and yellow colour (Case III) are located within the accuracy range, while the percentage of blue colour (Case II) means the probability of predicted traffic flow will lie outside the allowed accuracy but still within the allowed variability. And, the percentage of red colour (Case IV) means the probability of predicted traffic flow will lie outside the allowed limits of accuracy and variability. Both percentages of blue colour (Case II) and red colour (Case IV) are located outside the accuracy range.

Although the purpose of the research is visualization of the uncertainty, it is good to give more details by text. In the same example, for the link No. 6, concerning normal distribution scenarios as shown in Figure 8, we can see that Case I is decreased from 69.5% in SD=0.05 to 54.7% in SD=0.50; Case II started appearing from SD=0.30 and increased to 5.3% in SD=0.35 and decreased after that to reach 2.9% in SD=0.50; Case III is decreased from 32% in SD=0.05 to 3.2% in SD=0.50; and Case IV started appearing from SD=0.15 and increased to reach 39.1% in SD=0.50.

Similarly, we can interpret Figure 9 for lognormal scenarios; we can see that Case I is decreased from 68.2% in SD=0.05 to 57.7% in SD=0.50; Case II started appearing from SD=0.35 and increased to 18.7% in SD=0.50; Case III is decreased from 31.8% in SD=0.05 to 0.6% in SD=0.35 and disappeared after that; and Case IV started appearing from SD=0.15 and increased to reach 23.6% in SD=0.50.

Likewise, we can interpret Figure 10 for extreme value scenarios; we can see that Case I is decreased from 71.3% in SD=0.05 to 33.3% in SD=0.50; Case II started appearing from SD=0.15 and increased to 28.1% in SD=0.50; Case III is decreased from 28.6% in SD=0.05 to 12.6% in SD=0.50; and

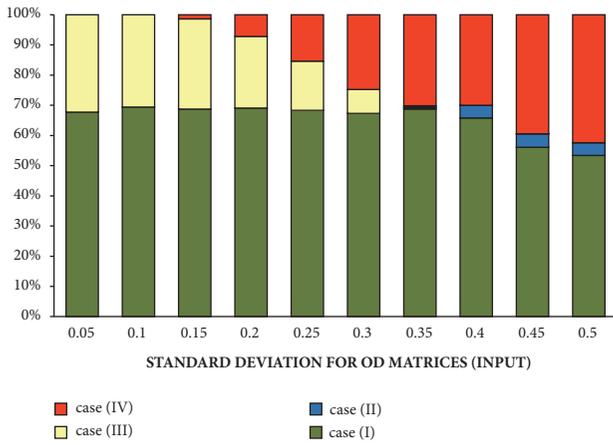


FIGURE 8: The predicted uncertainty for the link No. 6, applying normal distribution scenarios [27].

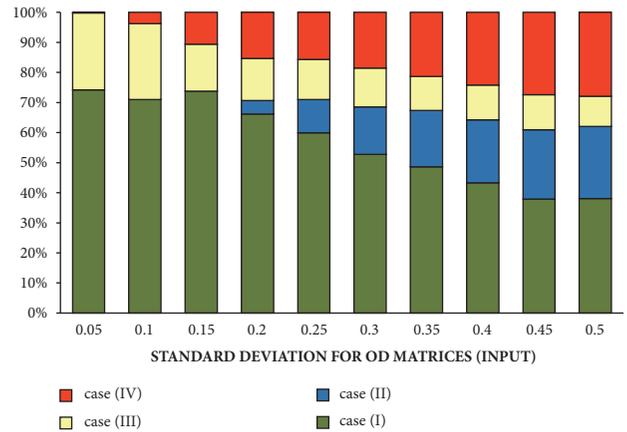


FIGURE 10: The predicted uncertainty for the link No. 6, applying extreme value distribution scenarios [27].

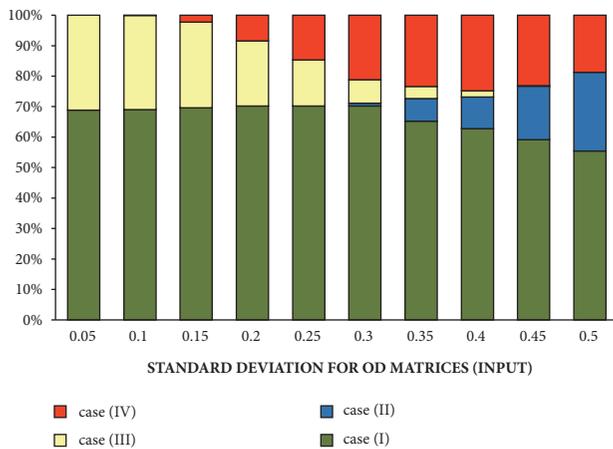


FIGURE 9: The predicted uncertainty for the link No. 6, applying lognormal distribution scenarios [27].

Case IV started increasing from 0.6% in SD=0.05 to 25.7% in SD=0.50.

Figures 8, 9, and 10 show that the predictive uncertainty strongly depends on the uncertainties of the input data. This makes the application of stochastic approaches in the practice of transport modelling highly advisable. Quantifying the uncertainty levels of predictions indicates the acceptability of the predictions and helps to identify the relating risks.

4. Conclusions

Visualization of predictive uncertainties helps to understand the stochastic nature of predictions. To be able to make decisions based on model predictions, decision-makers should have information about the accuracy and precision of the predictions. If the predictive uncertainty of the model is not acceptable, then sensitivity analysis helps to identify the input data most dominantly influencing the predictive uncertainty.

In this paper, a new methodology has been presented to predict traffic flow and visualize the uncertainty in those

predicted values. This methodology enables applying various scenarios showing the variation in traffic flow on transport network by supposing that the input values of OD matrix are varying according to a specific probability distribution. The importance of this methodology is permitting transport planners and decision-makers to monitor and identify which of the links suffers from bias and unexpected change in traffic volumes in the event of a change in the conditions of inputs OD matrix.

The algorithm of this methodology consists of two parts: the first part has been built on Monte Carlo simulation method to generate numerous OD matrices, and VISUM software for getting the traffic assignment on a transport network. The results of this part represent predicted traffic flows on each link of the transport network. These predicted traffic flows suffer from uncertainty in both a bias from the observed value and variability from the average predicted value, while the second part of the algorithm was designed to categorize the uncertainty of the predicted traffic flows into four cases according to variability and bias: Case I (low variability, low bias), Case II (low variability, high bias), Case III (high variability, low bias), and Case IV (high variability, high bias). Finally, the percentages of these cases have been visualized in coloured bar-charts. The percentage of each case represents the likelihood of occurring (i.e., the likelihood of the predicted traffic flow to biasing or varying depends on the percentages of these cases).

Finally, the methodology has been tested in a small study area using three main-scenarios; each of them has (i) different probability distribution (normal, lognormal, and extreme value) and (ii) 10 subscenarios different according to standard deviation parameter graded from 0.05 to 0.50. The obtained results of this study area showed that uncertainty in traffic flow is found on all links of the transport network but in different degrees, depending on the scenario's parameters and the observed traffic flow.

The current case study shows that the effect of applying scenarios had the same simulation parameters for all zones. Future research will consider applying different simulation

parameters in the same scenario according to land use characteristic of each zone and how the accuracy and precision of the predicted traffic flows can be improved once the case of uncertainty of the predicted traffic flow is known.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request. The software used in this study can be requested from the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] J. d. Ortúzar and L. G. Willumsen, *Modelling Transport*, John Wiley and Sons Ltd Publication, Chichester, UK, 4th edition, 2011.
- [2] H. S. Mahmassani, "Uncertainty in transportation systems evaluation: Issues and approaches," *Transportation Planning and Technology*, vol. 9, no. 1, pp. 1–12, 1984.
- [3] M. Ottomanelli and C. K. Wong, "Modelling uncertainty in traffic and transportation systems," *Transportmetrica*, vol. 7, no. 1, pp. 1–3, 2011.
- [4] Cambridge Systematics Inc., Vanasse Hangen Brustlin Inc., Gallop Corporation, C. R. Bhat, Shapiro Transportation Consulting LLC, and Martin Alexiou Bryson PLLC, *Travel Demand Forecasting: Parameters and Techniques*, 2012.
- [5] W. Briggs, *Uncertainty: The Soul of Modeling, Probability and Statistics*, Springer International Publishing, Switzerland, 2016.
- [6] G. de Jong, A. Daly, M. Pieters, S. Miller, R. Plasmeijer, and F. Hofman, "Uncertainty in traffic forecasts: Literature review and new results for The Netherlands," *Transportation*, vol. 34, no. 4, pp. 375–395, 2007.
- [7] G. J. De Jong, A. Daly, M. Pieters, S. Miller, and F. Hofman, "Uncertainty in traffic forecasts: literature review and new results for the Netherlands," *European Journal of Transport and Infrastructure Research (EJTIR)*, pp. 1–23, 2005.
- [8] A. Matas, J.-L. Raymond, and A. Ruiz, "Traffic forecasts under uncertainty and capacity constraints," *Transportation*, vol. 39, no. 1, pp. 1–17, 2012.
- [9] Y. Zhao and K. M. Kockelman, "The propagation of uncertainty through travel demand models: an exploratory analysis," *Annals of Regional Science*, vol. 36, no. 1, pp. 145–163, 2002.
- [10] D. J. Van Der Laan, E. De Jonge, and J. Solcer, "Effect of displaying uncertainty in line and bar charts: Presentation and interpretation," in *Proceedings of the 6th International Conference on Information Visualization Theory and Applications, IVAPP 2015*, pp. 225–232, Germany, March 2015.
- [11] S. E. Ziems, B. Sana, J. Plotz, and R. M. Pendyala, "Stochastic variability in microsimulation modeling results and convergence of corridor-level characteristics," in *Proceedings of the Transportation Research Board 90th Annual Meeting*, vol. 11-3560, 2010.
- [12] S. Manzo, O. A. Nielsen, and C. G. Prato, "Investigating uncertainty in transport modelling: a four-stage model case study," *Latsis Symposium*, 2012.
- [13] S. Manzo, O. A. Nielsen, and C. G. Prato, "How uncertainty in input and parameters influences transport model: Output A four-stage model case-study," *Transport Policy*, vol. 38, pp. 64–72, 2015.
- [14] S. Rasouli and H. Timmermans, "Uncertainty in travel demand forecasting models: Literature review and research agenda," *Transportation Letters*, vol. 4, no. 1, pp. 55–73, 2012.
- [15] S. Rasouli and H. Timmermans, "Applications of theories and models of choice and decision-making under conditions of uncertainty in travel behavior research," *Travel Behaviour and Society*, vol. 1, no. 3, pp. 79–90, 2014.
- [16] A. Pradhan and K. M. Kockelman, "Uncertainty propagation in an integrated land use-transportation modeling framework: Output variation via UrbanSim," *Transportation Research Record*, no. 1805, pp. 128–135, 2002.
- [17] W. J. Rider, J. R. Kamm, and V. G. Weirs, "Verification, validation and uncertainty quantification workflow in CASL," *Sandia National Laboratories*, no. SAND2010234P, pp. 1–27, 2010.
- [18] W. Walker, P. Harremoës, J. Rotmans et al., "Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support," *Integrated Assessment*, vol. 4, no. 1, pp. 5–17, 2003.
- [19] S. I. Gass and M. C. Fu, *Encyclopedia of Operations Research and Management Science*, Springer US, Boston, MA, USA, 2013.
- [20] B. M. Ayyub and G. J. Klir, *Uncertainty Modeling and Analysis in Engineering and the Sciences*, Chapman and Hall/CRC, Boca Raton, FL, USA, 2006.
- [21] Y. Du, Y. Geng, and L. Sun, "Simulation model based on monte carlo method for traffic assignment in local area road network," *Frontiers of Architecture and Civil Engineering in China*, vol. 3, no. 2, pp. 195–203, 2009.
- [22] H. Janssen, "Monte-Carlo based uncertainty analysis: sampling efficiency and sampling convergence," *Reliability Engineering & System Safety*, vol. 109, pp. 123–132, 2013.
- [23] A. T. Pang, C. M. Wittenbrink, and S. K. Lodha, "Approaches to uncertainty visualization," *The Visual Computer*, vol. 13, no. 8, pp. 370–390, 1997.
- [24] H. Griethe and H. Schumann, "The visualization of uncertain data: methods and problems," in *Proceedings of the SimVis '06*, vol. vi, pp. 143–156, SCS Publishing House, 2006.
- [25] L. Gosink, K. Bensema, T. Pulsipher et al., "Characterizing and visualizing predictive uncertainty in numerical ensembles through bayesian model averaging," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2703–2712, 2013.
- [26] J. Chang-Sung and A. Pang, "Reconfigurable disc trees for visualizing large hierarchical information space," in *Proceedings of the IEEE Symposium on Information Visualization*, pp. 19–25, IEEE, Research Triangle, Calif, USA, 1998.
- [27] M. Seger and L. Kisgyorgy, "Visualize the effect of input variability on model output in traffic assignment," in *Proceedings of the Fifth International Conference on Road and Rail Infrastructure*, pp. 1233–1240, 2018.



Hindawi

Submit your manuscripts at
www.hindawi.com

