

An Efficient Sensitivity Analysis Approach for Computationally Expensive Microscopic Traffic Simulation Models

Qiao Ge¹ and Monica Menendez²

Institute for Transport Planning and Systems, ETH Zurich, 8093 Zurich, Switzerland
¹*qiao.ge@ivt.baug.ethz.ch;* ²*monica.menendez@ivt.baug.ethz.ch*

Abstract

Microscopic traffic simulators are useful tools for designing, evaluating and optimizing transportation systems. In order for a simulator to accurately describe reality, the corresponding traffic model must be properly calibrated. However, the calibration can be rather difficult when the model is computationally expensive and has many parameters. To overcome these difficulties, Sensitivity Analysis (SA) can be applied as an essential instrument for supporting model calibration. Through SA the practitioners can obtain valuable information about the relationship between model inputs and outputs, and hence focus on the proper set of most influential parameters for the calibration. Notice, however, that many quantitative SA techniques may also fail with computationally expensive models.

To address the above issues, in this paper we developed the quasi-OTEE method, an efficient and qualitative SA approach based on the Elementary Effects (EE) method. The quasi-OTEE approach is able to screen the most influential parameters of a complex model through computing and comparing their Sensitivity Indexes. With the improved sampling strategy, this approach is much more efficient than the original EE method. The approach is validated through a numerical analysis, and a case study of a small synthetic network in Aimsun. A more detailed case study of the Zurich network in VISSIM is then provided to illustrate its application. The results demonstrate that the proposed approach is an effective SA tool for the computationally expensive microscopic traffic models, as well as other complex models in the general scientific community.

Keywords: *sensitivity analysis, quasi-OTEE, microscopic traffic simulation, computationally expensive model, calibration.*

1. Introduction

Microscopic traffic simulators are useful tools for designing, evaluating and optimizing transportation systems. In order for a simulator to accurately describe reality, it must be supported by a valid traffic model, and it must be properly calibrated. Calibration implies that the input parameters (*e.g.*, driving behavior, desired speed) allow the model to recreate the specific network under certain circumstances (*i.e.*, replicate observations, field measurements, and other empirical data). This step is vital, but can be rather complex. Such complexity is especially severe when the model contains a large number of parameters (many of them might be unobservable in the field and/or hard to measure), and/or when the simulation model lacks transparency (*e.g.*, the algorithm used is not accessible to the model user, the input data is processed in a black box). In addition, when the modeled network is complex, even running the simulation itself might be very time consuming. Due to the limitation of available time and other resources, when dealing with a computationally expensive model, one feasible way is to calibrate only the most influential parameters (*i.e.*, the parameters whose variations are

expected to have significant impacts on the model output). Through fine tuning these parameters, it is expected that the model outputs can be efficiently adjusted towards the target values 0.

To select the correct set of influential parameters from a complex model, one widely adopted instrument in the general scientific community is the Sensitivity Analysis (SA) [1]. SA can be defined as the study of how the variations in the model outputs can be attributed to the variations in the model inputs [2]. Through providing the modeler qualitative and/or quantitative information regarding the effects of parameters and their variations on the simulation results, SA can help to efficiently identify the most critical parameters for the subsequent calibration.

Unfortunately, to the authors' knowledge, there are few examples in which the SA of microscopic traffic models is performed: the authors of [3] and [4] used the One-At-a-Time (OAT) method to evaluate the variance of the VISSIM output. The variance-based SA approach was employed in [5] for calibrating VISSIM; in [6], [7], and [8] for Aimsun; and in [9] and [10] for Paramics. In [1], Kriging metamodelling was employed to approximate and simplify the original complex model in order to investigate the effects of different optimization algorithms. Despite these examples, there appears to be no previous research suggesting a general method for the sensitivity analysis of traffic simulation models. Furthermore, when the traffic model is complex and has many parameters, many traditional SA techniques (*e.g.*, Monte Carlo approaches, for details see [2]) are unfeasible because of the high computational demand of the traffic model itself.

This paper aims at providing a practical and efficient SA approach, which is especially useful for computationally expensive microscopic traffic models. The proposed approach is developed based on the Elementary Effects (EE) method [11], but represents a significant improvement in terms of efficiency. The paper is organized as follows. Section 2 presents a brief review of the EE Method and other modifications made until now. Section 3 explains the methodology for the proposed SA approach, and demonstrates two validation experiments. Section 4 illustrates a case study based on a real calibration project with the City of Zurich. Section 5 presents the conclusions and some recommendations for future research.

2. Review of Elementary Effects Method

2.1. Introduction of Elementary Effects Method

The Elementary Effects (EE) method is a qualitative and stochastic approach developed by Morris [11] for screening the influential parameters from a complex model [2]. It has been successfully applied in different fields such as environmental engineering [12] and chemistry [13], but never in traffic engineering. The main feature of the EE method is described below.

The EE method employs the OAT design in the data sampling. Suppose a model Y has k independent parameters. For any possible value of these k parameters in the input space, *i.e.*, $\mathbf{X} = [X_1, X_2, \dots, X_k]$, the corresponding model output $Y(\mathbf{X})$ is then defined as $Y(X_1, X_2, \dots, X_k)$. If only the i -th parameter in \mathbf{X} is changed by a certain value, *e.g.*, to $X_i + \Delta$, while all other parameters remain unchanged, the output will consequently be $Y(X_1, X_2, \dots, X_{i-1}, X_i + \Delta, X_{i+1}, \dots, X_k)$. The Elementary Effect of the i -th parameter (*i.e.*, EE_i) is then defined as:

$$EE_i = \frac{Y(X_1, \dots, X_{i-1}, X_i + \Delta, X_{i+1}, \dots, X_k) - Y(X_1, \dots, X_{i-1}, X_i, X_{i+1}, \dots, X_k)}{\Delta} \quad (1)$$

As the Elementary Effect itself is indeed a local sensitivity measure, to evaluate the global sensitivity of the model, the above calculation should be performed throughout the whole

input space based on random samples. By randomly generating a certain number (e.g., m) of \mathbf{X} in the input space, and each time only changing the i -th parameter by Δ , then m EEs of the i -th parameter can be obtained according to the definition above. The mean μ , the standard deviation σ , and the absolute mean μ^* of these m EEs, which are called the Sensitivity Indexes (SI), can be derived as:

$$\mu_i = \frac{1}{m} \sum_{j=1}^m EE_i^j \quad (2)$$

$$\sigma_i = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (EE_i^j - \mu_i)^2} \quad (3)$$

$$\mu_i^* = \frac{1}{m} \sum_{j=1}^m |EE_i^j| \quad (4)$$

where EE_i^j is the Elementary Effect of the i -th parameter computed by using the j -th random sample of \mathbf{X} .

According to [11] and [13], the SI can be used to classify the parameters as:

- 1) negligible parameters (μ^* is low);
- 2) parameters with linear and additive effects but no interactions with others (μ^* is high but σ is low); and
- 3) parameters with non-linear effects and/or strong interactions with others (μ^* and σ are both high).

In addition, when μ is low but μ^* is high, the parameter will have oscillating effects (i.e., both positive and negative effects) depending on the value of other parameters.

Furthermore, it is possible that some model parameters can have strong correlations during the SA, and they may cause different impacts on the model outputs when they are changed simultaneously versus individually. In such case, a possible solution is performing the above analysis using the EE method for each individual parameter first, and calculating the μ^* and σ respectively. Then in the second round of SA, we put the parameters with relatively similar μ^* and σ into a same group, and redo the sampling in the input space based on groups (i.e., in the group sampling, unlike the original OAT design, all parameters in a group are changed at the same time, for details see [2]). Finally the μ^* of different groups are calculated and used to further derive the sensitivity information (interested readers may refer to the example of group SA for the Oakley-O'Hagan function in [2]).

According to the definition above, two model runs are required to calculate one EE for one model parameter: first with the initial inputs [$X_1, X_2, \dots, X_{i-1}, X_i, X_{i+1}, \dots, X_k$], and then with the varied inputs [$X_1, X_2, \dots, X_{i-1}, X_i + \Delta, X_{i+1}, \dots, X_k$]. Suppose in a k -parameter model, m EEs are needed for calculating the global SI of each parameter, then $2mk$ runs of the model are required, i.e., the computational cost of the basic EE method is $2mk$.

2.2. Improvement by Tamplng with Morris Trajectories

To improve the efficiency of the EE method, we can sample the parameter input space by using the Morris trajectories [11]. Figure 1 gives an example of one Morris trajectory for a model with two parameters (named X_1 and X_2).

The first point P^0 is randomly picked in the input space with the coordinates [X_1^0, X_2^0]. The second point P^1 is generated by randomly increasing or decreasing the value of one parameter by a certain value Δ based on P^0 . Hence, the coordinates of P^1 in this case, for example, are [$X_1^0 + \Delta, X_2^0$]. The third point P^2 is generated in the same way but by changing another parameter based on P^1 , and its coordinates are [$X_1^0 + \Delta, X_2^0 + \Delta$]. According to the above definition of EE, two EEs can be obtained based on the three sampling points:

$$EE(X_1) = \frac{Y(P^1) - Y(P^0)}{\Delta} = \frac{Y(X_1^0 + \Delta, X_2^0) - Y(X_1^0, X_2^0)}{\Delta} \quad (5)$$

$$EE(X_2) = \frac{Y(P^2) - Y(P^1)}{\Delta} = \frac{Y(X_1^0 + \Delta, X_2^0 + \Delta) - Y(X_1^0 + \Delta, X_2^0)}{\Delta} \quad (6)$$

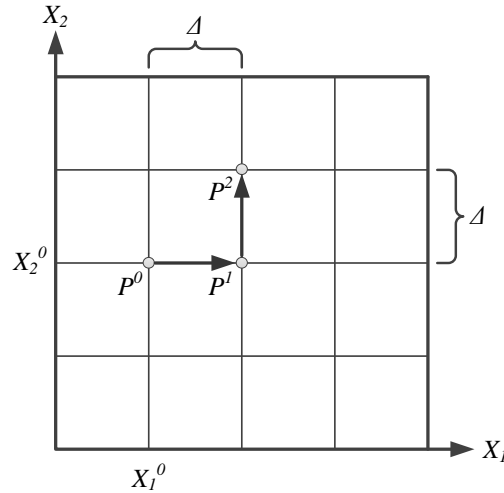


Figure 1. Illustration of One Morris trajectory Created for a 2-parameter Model

The example above indicates that for a k -parameter model, a trajectory with $k+1$ points can provide k EEs (one per parameter). Therefore, by randomly sampling m trajectories in the input space, each parameter can get the same amount of EEs (*i.e.*, m) as before, but only $m(k+1)$ model runs are needed.

2.3. Improvement by Sampling with Optimized Trajectories

Since all Morris trajectories are randomly generated, it is possible that many of them overlap in a certain area of the input space. Using the overlapping trajectories for data sampling may result in a poor representation of the input space while requiring many unnecessary model executions. To overcome such drawback, an improved sampling design was proposed in [13]: based on a high number of Morris trajectories, a subset of the most spread trajectories is taken out as the Optimized Trajectories (OT) set. The concept “spread” is defined based on the Euclidean distance d_{T_i, T_j} between any two trajectories T_i and T_j :

$$d_{T_i, T_j} = \begin{cases} \sqrt{\sum_{r=1}^k [X_r^p(i) - X_r^q(j)]^2} & i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where k is the number of parameters, $X_r^p(i)$ is the r -th coordinate of the p -th point in trajectory T_i . For a set (*i.e.*, S) of n trajectories (*i.e.*, $S = \{T_1, T_2, \dots, T_n\}$), the total distance D_S of this set is defined as:

$$D_S = \sqrt{0.5 * (\sum_{i=1}^n \sum_{j=1}^n d_{T_i, T_j}^2)} \quad (8)$$

By enumerating all sets that contain n trajectories from the original set of m random Morris trajectories ($n \ll m$), the set with the maximum D_S is selected as the OT set, and the number

of model runs is accordingly reduced to $n(k+1)$. The number of Morris trajectories (*i.e.*, m) is usually chosen between 500 and 1000 (*e.g.*, [2], [13] and [14]), but there is currently no standard rule for the choice of n . Although some empirical studies such as [13] and [15] have shown that 10 to 20 OT are sufficient to provide satisfactory results, the authors of [14] and [16] argued that a high value of n (*e.g.*, 100) should be considered, especially when there are many nonlinear interactions among the parameters. As a rule-of-thumb, one possible way to find the optimal n is to increase the number of optimized trajectories gradually, and calculate the corresponding rankings of the parameters based on the SI. The optimal n is reached once the rankings (especially the rankings of the most influential parameters) converge. It should be noted that although a high value of n is expected to raise the accuracy of the SA, it may make the initial screening unfeasible when running the simulation is computationally expensive and/or there are many parameters. In practice, a reasonable n should be adopted by considering both the accuracy and efficiency of the SA.

The advantage of using OT for sampling is that with the same number of trajectories, they can cover more sampling points than any other non-OT. Hence, it facilitates a better scanning of the input space without increasing the number of model runs [2]. As a result, the same SI can be achieved by sampling with OT, but the number of model executions is greatly reduced. Moreover, because the OT sampling can ensure a good scan of the input space by itself, it is not necessary to use other sampling approaches such as Latin Hypercube Sampling (LHS, for more details see [2]) to produce the Morris trajectories at the beginning [13].

However, when the model is complex and has many inputs, the combinatorial optimization problem of OT selection can make the OT sampling unfeasible: to find the optimized set of n trajectories out of the m Morris trajectories, the total number of possible combinations is calculated as $m!/[n!(m-n)!]$. In some cases the number of combinations could be huge. For example, when $n = 10$, $m = 200$, there are over 2.2×10^{16} combinations. In a case like this, although running the model might be computationally feasible, it is unfeasible for a normal PC to check all possible combinations to find the OT. In this paper, we improve the OT sampling approach to overcome such drawback. The details are given in the next section.

3. Quasi-Optimized Trajectories based Elementary Effects Method

3.1. Introduction to the Algorithm

In this paper, we develop the quasi-Optimized Trajectories based Elementary Effects (quasi-OTEE) method. This method inherits the same measures of the original EE approach for screening parameters, but overcomes the aforementioned combinatorial optimization problem of the OT sampling, *i.e.*, the quasi-OT sampling significantly improves the efficiency in data sampling. The algorithm of the quasi-OT sampling is described below.

In order to find the “optimal” trajectories, a set (named S_m) of m Morris trajectories, *i.e.*, T_1, T_2, \dots, T_m , is randomly generated in the input space. According to Eq. (7), the distance matrix for all trajectories is defined as:

$$d_{S_m} = \begin{bmatrix} 0 & d_{T_1, T_2} & \dots & d_{T_1, T_m} \\ d_{T_2, T_1} & 0 & \dots & d_{T_2, T_m} \\ \dots & \dots & \dots & \dots \\ d_{T_m, T_1} & d_{T_m, T_2} & \dots & 0 \end{bmatrix} \quad (9)$$

In the first step, the total distance (*i.e.*, D_{S_m}) of the m trajectories belonging to S_m is calculated as:

$$D_{S_m} = \sqrt{0.5 * (\sum_{i=1}^m \sum_{j=1}^m d_{T_i, T_j}^2)} \quad (10)$$

In the second step, all trajectory sets that contain $m-1$ trajectories are generated based on S_m . In total there are m possible sets, named $S_{m-1}(1), S_{m-1}(2), \dots, S_{m-1}(m)$ respectively. The trajectory set $S_{m-1}(p)$ ($1 \leq p \leq m$) contains all trajectories in S_m except the trajectory T_p (for example, $S_{m-1}(1) = \{T_2, T_3, \dots, T_m\}$). The total distance of the $m-1$ trajectories belonging to set $S_{m-1}(p)$ is calculated as:

$$D_{S_{m-1}(p)} = \sqrt{D_{S_m}^2 - \sum_{i=1}^m d_{T_i, T_p}^2} \quad (11)$$

Then the trajectory set $S_{m-1}(H_1) = \{T_1^{H_1}, T_2^{H_1}, \dots, T_{m-1}^{H_1}\}$ ($1 \leq H_1 \leq m$) that has the highest total distance is selected as the optimal set in this step.

In the third step, based on $S_{m-1}(H_1)$, we can enumerate all trajectory sets containing $m-2$ trajectories, *i.e.*, $S_{m-2}(1), S_{m-2}(2), \dots, S_{m-2}(m-1)$, and find the optimal set $S_{m-2}(H_2)$ ($1 \leq H_2 \leq m-1$) with the $m-2$ most spread trajectories in $S_{m-1}(H_1)$.

In the next step, we do the same but based on $S_{m-2}(H_2)$, etc. In each step except the first one, the total distance is calculated based on the total distance derived in the previous step. For example, if the optimal set $S_{m-r}(H_r)$ is found and it contains the $m-r$ most spread trajectories, *i.e.*, $T_1^{H_r}, T_2^{H_r}, \dots, T_{m-r}^{H_r}$, then in the next step, the total distance of the $m-r-1$ trajectories in set $S_{m-r-1}(p^{H_r})$ is calculated as:

$$D_{S_{m-r-1}(p^{H_r})} = \sqrt{D_{S_{m-r}(H_r)}^2 - \sum_{i=1}^{m-r} d_{T_i^{H_r}, T_p^{H_r}}^2} \quad (12)$$

where $1 \leq r \leq m-n-1$, $1 \leq p \leq m-r$, and $S_{m-r-1}(p^{H_r})$ is the set containing all trajectories in $S_{m-r}(H_r)$ except trajectory $T_p^{H_r}$ (notice that the superscript H_r is used here just for the ease of describing different subsets based on the optimal set $S_{m-r}(H_r)$, while the trajectories remain the same). Compared to Eq. (10), Eq. (12) saves significant computational effort in enumerating the combinations of any two trajectories belonging to set $S_{m-r-1}(p^{H_r})$ when calculating the total distance. Hence, the computational cost of each additional step is greatly reduced.

The number of trajectories within the optimal trajectory set is decreased by one after each step. Finally, there is a set $S_n(H_{m-n})$ containing the n most spread trajectories from the previous optimal set $S_{n+1}(H_{m-n-1})$. As these n trajectories are not necessarily the same ones found by the aforementioned OT approach, they are called herein quasi-Optimized Trajectories (quasi-OT).

In the quasi-OT generation process, only $m-1$ possible combinations need to be considered when picking $S_{m-1}(H_1)$ from S_m , and $m-2$ possible combinations are considered when picking $S_{m-2}(H_2)$ from $S_{m-1}(H_1)$, etc. The total number of combinations considered to get the final optimal set $S_n(H_{m-n})$ is $(m-n+1)*(m+n)/2$. This number is much smaller than the number of combinations required by the original OT approach (*i.e.*, $m!/[(n!)*(m-n)!]$). Using the same example as in the last section, the quasi-OT approach requires only 20,055 combinations, much less than the 2.2×10^{16} combinations required by the original OT approach.

3.2. Validation of the quasi-OTEE Approach

Here we conduct two experiments to demonstrate the efficiency and accuracy of the quasi-OTEE approach. The first experiment focuses on the performance of the proposed quasi-OT sampling algorithm, by cross-comparing it with the original OT sampling in [13] and a recent developed sampling approach in [14]. The second experiment aims to check the validity of the quasi-OTEE approach, by comparing its results with those obtained from the well-known variance-based approach.

1) Experiment 1

In the first validation experiment 500 trajectory sets (*i.e.*, cases) were randomly generated in the input space. Each set contained 20 Morris trajectories (*i.e.*, $m = 20$). In each trajectory set, 10 trajectories (*i.e.*, $n = 10$) were selected respectively via the original OT approach (*i.e.*, Approach_1) proposed by Campolongo *et al.*, [13], via the quasi-OT approach (*i.e.*, Approach_2) proposed in this study, and via another optimized trajectory approach (*i.e.*, Approach_3) proposed by Ruano *et al.* [14].

Since Approach_1 uses the exhaustive search algorithm to find the true OT, it often requires an extremely high computational cost. For the sake of cross-comparing the three approaches, we chose $n = 10$ and $m = 20$ in this validation test intentionally, so that it is computationally feasible with a normal PC (with a 2.80 GHz Intel Core i5 processor) to check all possible combinations with the original OT approach. Different settings of n and m might be considered in other tests when a High Performance Computer (HPC) is available, although similar conclusions are expected. The comparison results are shown in Table 1.

Table 1. Comparison of the Three Trajectory Sampling Approaches

	Approach_1	Approach_2	Approach_3
Cases with no differences in trajectories compared to Approach-1	--	441 (88%)	15 (3%)
Cases with a difference of 1 trajectory compared to Approach-1	--	42 (8%)	23 (5%)
Maximum difference in trajectories compared to Approach-1	--	4 trajectories in 1 case	7 trajectories in 10 cases
Average total distance obtained by the 10 optimal trajectories	3101.8	3100.7	3074.9
Total computation time of the 500 cases	15278s	12s	14s

Note: Approach_1 = original OT approach proposed by Campolongo *et al.* in [13]

Approach_2 = quasi-OT approach proposed in this paper

Approach_3 = approach proposed by Ruano *et al.* in [14]

Table 1 shows that the quasi-OT are very close to the trajectories obtained by the original OT approach. In most cases (88%) the quasi-OT and the true OT are exactly the same. Moreover, the average total distances are almost the same. Compared with the trajectories obtained by the optimization approach presented in [14], although both approaches spent similar time in generating the trajectories, the quasi-OT always have higher dispersion (*i.e.*, the total distance achieved is longer), which means they have less overlaps and hence can provide a better coverage of the input space. This can be explained as below.

In Approach_3, a set of t trajectories ($t \leq n$) with the highest Euclidean distance (see Eq. (7)) are first chosen from each row of the distance matrix (see Eq. (9)), and the final optimized trajectories are obtained by adding trajectories to this set through iterations (more details are given in [14]). However, as only the Euclidean distance but not the total distance (see Eq. (8)) is considered in the first step of Approach_3, it is possible that the t trajectories

are actually not the most spread trajectories. As an example, assume there are four trajectories (i.e., T_1 , T_2 , T_3 , and T_4), and the Euclidean distances d_{T_1,T_2} , d_{T_1,T_3} , d_{T_1,T_4} are ordered as $d_{T_1,T_2} > d_{T_1,T_3} > d_{T_1,T_4}$. Then in the first step of Approach_3, the trajectory set $\{T_1, T_2, T_3\}$ will be chosen because d_{T_1,T_2} and d_{T_1,T_3} are the two largest elements in the first row of the distance matrix. However, if d_{T_1,T_3} is just slightly larger than d_{T_1,T_4} , but $d_{T_2,T_3} \ll d_{T_2,T_4}$, the total distance D_{T_1,T_2,T_3} (i.e., $[0.5*(d_{T_1,T_2}^2 + d_{T_1,T_3}^2 + d_{T_2,T_3}^2)]^{0.5}$) can be much shorter than the total distance D_{T_1,T_2,T_4} (i.e., $[0.5*(d_{T_1,T_2}^2 + d_{T_1,T_4}^2 + d_{T_2,T_4}^2)]^{0.5}$). In such case, as the initial trajectory set found by Approach_3 is non-optimal, it is possible that the final trajectory set obtained would be less spread than the quasi-OT set obtained by Approach_2 (recall that the quasi-OT set is obtained by always discarding the worst trajectory in each iteration step).

In addition, it is obvious in Table 1 that the quasi-OT sampling took the shortest computational time (12s) in this experiment, especially compared to the original OT sampling which took over 1,000 times longer.

This experiment has shown that the sampling results from the quasi-OT are highly consistent with those from the original OT method in [13], and better than those from the approach introduced in [14]. In addition, the quasi-OT sampling has the highest efficiency (i.e., least computation time) among the three approaches, and this advantage is especially significant when compared to the original OT method.

2) Experiment 2

In the second validation experiment, we used a small synthetic network (see Figure 2) modeled in Aimsun.

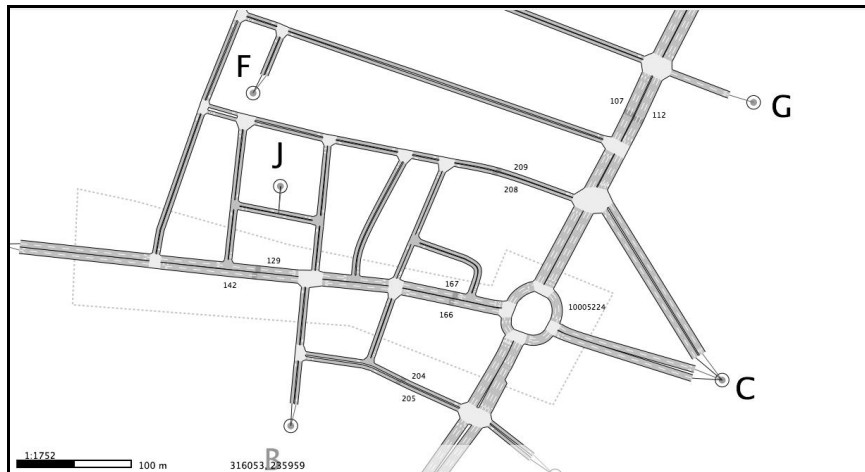


Figure 2. Layout of the Aimsun Network in the Second Validation Experiment

We included 8 parameters for the SA in this experiment: vehicle reaction time (rT), vehicle reaction time at stop (rTs), vehicle length (vL), jam density (jD), give-way time (gWt), maximum acceleration (mA), random seed of the simulation ($seed$), and the demand factor (OD). Four model outputs were collected from the whole network: delay, density, traffic flow, and travel time. The SA using the variance-based approach (for details see [2]) was carried out first. It took 40,960 runs of the Aimsun model, and the results were used as reference. Then, 10 tests using the quasi-OTEE were performed. In each test, 50 quasi-OT were chosen from 500 randomly generated Morris trajectories. Every test with the quasi-OTEE took 450 model runs.

The total sensitivity indexes (for details see [2]) calculated in the reference simulations with the variance-based approach are used to give the true rank of a parameter according to its importance (*i.e.*, the most important parameter has the highest total sensitivity index). In the quasi-OTEE analysis, the sensitivity index μ^* is used to sort the parameters from most important to least important (*e.g.*, the parameter with the highest μ^* is regarded as the most important parameter). The detection rates of the quasi-OTEE approach across different outputs and tests are obtained based on the results of the variance-based approach. A brief explanation of the method used to compute the detection rate is given below.

Let V and Q be two parameters sets obtained with the variance-based approach and the quasi-OTEE approach respectively. Each set contains I parameters with the highest total sensitivity indexes or μ^* . Then the quasi-OTEE detection rate of the I important parameters is calculated as the cardinality of the intersection of V and Q (*i.e.*, $V \cap Q$) divided by I .

Figure 3 shows the aggregated results based on all outputs and tests. It is evident from the figure, the quasi-OTEE approach has a very high precision in identifying the important parameters of the model. In most cases, it can give the same results of the most important parameters as the variance-based approach, but with much less model runs (almost 1% of the total runs used by the variance-based approach). In other words, the quasi-OTEE approach is able to effectively find the whole set of important parameters and decide which parameters to discard.

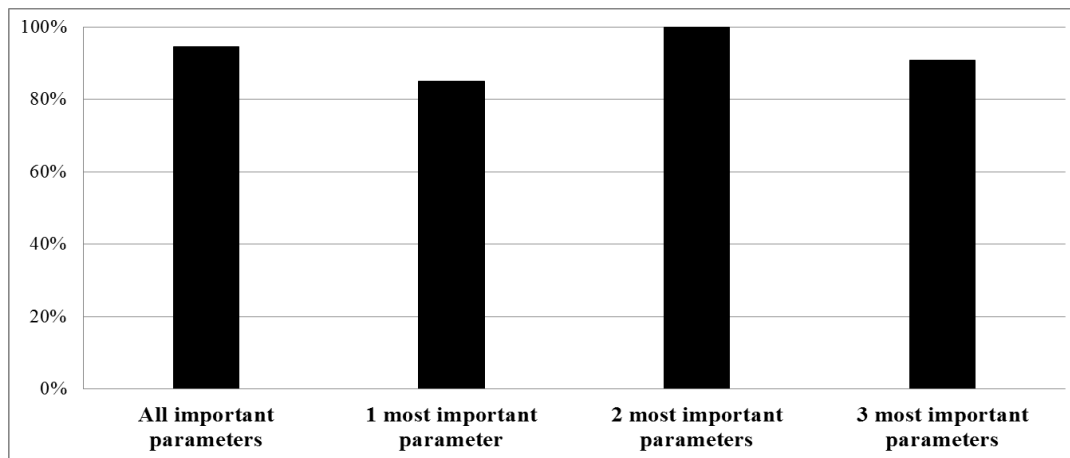


Figure 3. Detection Rates of Important Parameters in the Second Validation Experiment. Results are Aggregated Across all Outputs and Tests

To sum up, the two validation experiments have shown that: 1) the quasi-OT sampling significantly improves the efficiency of the original OT sampling in [13]; 2) the sampling by quasi-OT is better than the sampling approach proposed in [14] in terms of more space coverage; and 3) the quasi-OTEE has very high precision in identifying important parameters, comparable with the variance-based approach but with much less computational cost. Thus, the quasi-OTEE can be considered as an ideal approach for the preliminary SA of computationally expensive models, especially when it is unfeasible to apply the variance-based techniques from the beginning.

4. Case study: Sensitivity Analysis of Zurich Network in VISSIM

4.1. Case Study Background

In this section we illustrate the application of the proposed quasi-OTEE approach with a case study. For that we employ a VISSIM network of Zurich, Switzerland. This network encompasses the city downtown, with a total area of around 2.6 km². It is a complex urban layout with narrow streets, hills, mixed transportation modes, a large amount of pedestrians, *etc.* Due to these complexities, the simulation has a very high computational demand (around 30 minutes for a 1-hour traffic simulation), especially when compared to the studies that simulate individual links, single intersections or a few intersections along a single road (*e.g.*, [17], [18], [19], [20], [21], [22]) and [23]). To make the calibration of this complex model feasible, the SA is employed prior to the calibration for identifying the highly influential parameters.

Since the number of parameters contained in the original model is too large (192 parameters, see [24]), it is hard to make any SA feasible. Thus, an initial screening was performed first to narrow the set of critical parameters. The first cut from 192 parameters in the original model down to 148 was based on the study of Zurich inner city's traffic patterns and characteristics, model intended uses, and available data. For example, parameters related to bicycle lanes were discarded because bicycles were not included in the model. Then the remaining 148 parameters were classified based on previous studies, common sense, and our own experience: 55 parameters were obtained directly from the macroscopic traffic demand model, and another 79 parameters were regarded as not very influential or not highly variable, so they were fixed to their default values in all simulations. At last, 14 parameters (see Table 2) were chosen as the most critical SA parameters. These parameters were then used for the subsequent analysis with the quasi-OTEE approach.

Table 2. 14 Parameters used in the Sensitivity Analysis

#	Parameter	Range	Unit
1	Average Standstill Distance	[1, 3]	m
2	Additive Part of Desired Safety Distance	[0, 4]	--
3	Multiplicative Part of Desired Safety Distance	[1, 5]	--
4	Maximum Deceleration (Own)	[-6, -2]	m/s ²
5	Accepted Deceleration (Own)	[-1.5, -0.5]	m/s ²
6	-1 m/s ² per Distance (Own)	[50, 150]	m
7	Max. Deceleration (Trailing)	[-5, -1]	m/s ²
8	Accepted Deceleration (Trailing)	[-1.5, -0.5]	m/s ²
9	-1 m/s ² per Distance (Trailing)	[50, 150]	m
10	Minimum Headway	[0.3, 1]	m
11	Safety Distance Reduction Factor	[0, 1]	--
12	Maximum Deceleration for Cooperative Braking	[-5, -1]	m/s ²
13	Lane Change Distance	[150, 250]	m
14	Emergency Stop Distance	[3, 7]	m

4.2. Design of Sensitivity Analysis

In this case study, the SA aimed at finding the parameters whose variation had the greatest impact on travel time. For this, 20 road sections in the inner city of Zurich (Figure 4) were chosen as the travel time measurement sections.

The simulation resolution (*i.e.*, the number of times the vehicle's position is calculated within one simulated second) was set to 10 as required by the adaptive signal control program used by the city of Zurich. For each simulation, based on the quasi-OT sampling, the 14 parameters took values from uniform distributions within the ranges shown in Table 2. All

other VISSIM parameters were set to their default values [24], or to the values derived from the demand model.

The vehicles' travel time for passing through each road section were measured for a period of one hour after a 15-minutes warm-up period. As the lengths of the twenty travel time measurement sections are not the same, we used the *Travel Time per Meter Travelled* (TTMT, calculated as travel time divided by the length of the corresponding section, and equivalent to the inverse of the speed) to aggregate the results measured from different sections. The final sensitivity indexes are accordingly computed based on the TTMT aggregated over all sections.

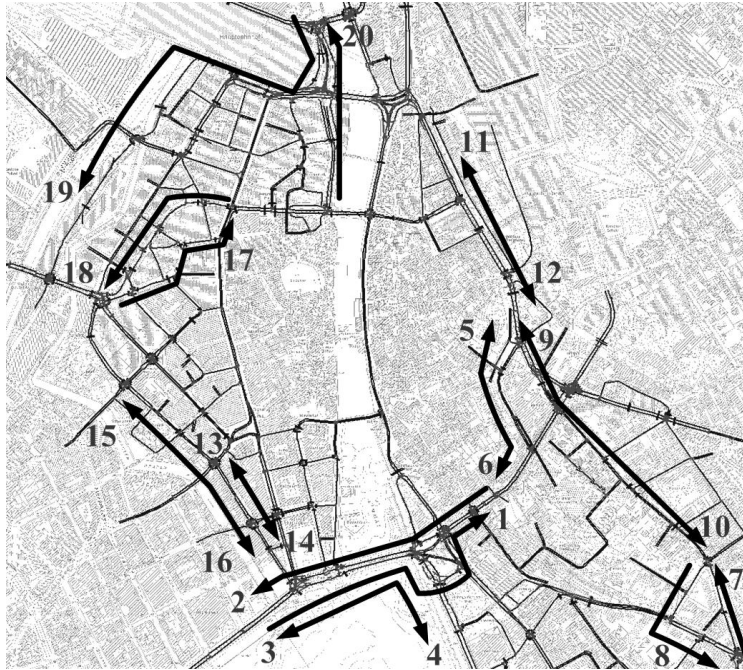


Figure 4. Travel Time Measurement Sections

4.3. Computational Cost Estimation

The quasi-OTEE analysis in this case study used 40 quasi-OT that were selected from 500 random Morris trajectories. As there were 14 parameters involved in the SA, in total 600 simulations were required in this case study. The total computation time was around 12 days (including 20 minutes for the quasi-OT selection). Below are some rough estimates of the computational costs of different EE approaches described in Section 2 (using the same PC as in Section 3.2):

- EE method using 200 random samples for each parameter: 116 days
- EE method with 200 Morris trajectories: 62 days
- EE method with 40 true OT: 12 days + $5 \cdot 10^{44}$ days for the selection of true OT

Therefore, if considering that the computation time has been greatly reduced by the quasi-OTEE, although it does not guarantee exactly the same data sampling points as the true OT, the quasi-OTEE is a good compromise between accuracy and efficiency.

Notice that it is possible to use a HPC (*e.g.*, 50G cores with parallel computing) to overcome the computation problem described above. However, the cost and availability of the HPC, and the limits from the simulation itself (*e.g.*, software compatibility with the HPC, security considerations in transferring the data) should not be neglected. Furthermore, even if

a HPC is available, the quasi-OTEE approach still has its own advantages due to its high efficiency. Compared to the same number of randomly generated Morris trajectories, the quasi-OT will surely have less overlaps and provide a better coverage of the sampling space. Compared to the true OT, due to the high efficiency of the quasi-OTEE, when spending the same computational cost, a much bigger set of the original Morris trajectories (see Section 2.2 and 2.3) could be considered with the quasi-OT sampling (*i.e.*, $m_{\text{quasi-OT}} \gg m_{\text{true-OT}}$). Using a bigger set of the original Morris trajectories could lead to a more thorough exploration of the sampling space, and further increase the reliability of the results.

4.4. SA Results

The SI (*i.e.*, μ^* , μ , and σ of EE) were calculated after finishing the 600 simulations. The results for all 14 parameters are plotted in Figure 5. It should be noted that the SI values presented in Figure 5 are context dependent [11], and their scale could be totally different in other studies (*e.g.*, [12] and [13]). Hence, when analyzing the results, the practitioners shall not focus on any particular value/scale of the SI, but the relative differences among them. To the authors' knowledge, there is currently no quantitative technique for interpreting the results rather than qualitatively comparing the SI of different parameters.

For the ease of demonstrating the results, the parameters with similar SI in the μ^* - σ plot (Figure 5a) were grouped by using the K-Means Clustering approach [25]. Five clusters were determined accordingly:

- Cluster 1: Parameter 2 ($\mu^* \approx 0.113$, $\sigma \approx 0.089$);
- Cluster 2: Parameter 1 ($\mu^* \approx 0.081$, $\sigma \approx 0.107$);
- Cluster 3: Parameter 3 ($\mu^* \approx 0.072$, $\sigma \approx 0.073$);
- Cluster 4: Parameters 4, 5, 6, 8, 9, 10, 12, 13 and 14 ($\mu^* \approx 0.065$, $\sigma \approx 0.102$);
- Cluster 5: Parameters 7 and 11 ($\mu^* \approx 0.051$, $\sigma \approx 0.068$).

Parameters 7 and 11 in Cluster 5 have the lowest μ^* and σ , therefore they are categorized as the least influential parameters, *i.e.*, they have the least amount of interactions with other parameters and the model output is rarely influenced by their variations. It is also clear that Parameter 2 has the highest μ^* , hence it is considered as the most influential parameter. Moreover, Parameter 1 has the second highest μ^* but the highest σ , which indicates that the variation of this parameter can cause strong non-linearity effects and/or interactions with other parameters. Therefore, Parameter 1 is also considered as one of the most influential parameters.

To examine the sensitivity of the rest parameters, we plotted a wedge formed by two lines corresponding to $\mu = \pm 2\text{SEM}$ (Standard Error of the Mean, calculated as σ/\sqrt{n} where n is the sample size, in this case 40) in the μ - σ plot (Figure 5b). According to [11], the variation of any parameter lying outside the wedge will surely influence the model output, independently of its adopted value, the range of the variation, and the adopted values of the other parameters. Therefore, Parameter 3 is also one influential parameter, although comparing to Parameters 1 and 2 its variation has fewer impacts on the model output, and it is less correlated with other parameters. As for the parameters belonging to Cluster 4, a comparison between Figure 5a and Figure 5b shows that they all have relatively low μ but high μ^* and σ . This indicates that these parameters will have both positive and negative effects depending on the adopted values of other parameters. Therefore, in order to avoid the Type II error (*i.e.*, considering an important parameter as non-important, see [2]), it is reasonable to expand the set of most influential parameters by including Parameters 4, 13 and 14 (*i.e.*, the parameters with the highest μ^* and σ in Cluster 4) in this study. However, this may require extra efforts in the next steps of the analysis (*e.g.*, model calibration).

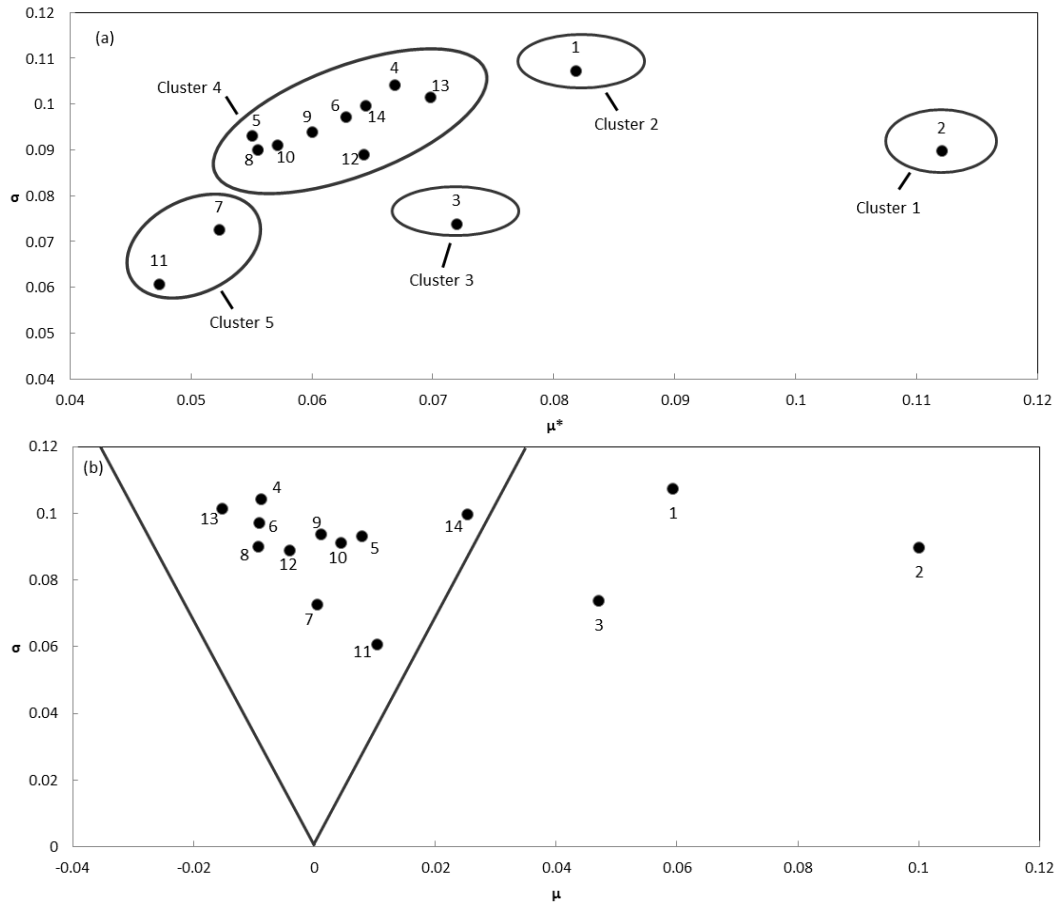


Figure 5. (a) Plots of μ^* versus σ of the EE with Respect to TTMT for the 14 Parameters. The Plots are separated into 5 Clusters based on the K-Means Clustering. (b) Plots of μ versus σ of the EE with Respect to TTMT for the 14 Parameters. Lines in the Figure Correspond to $\mu = \pm 2SEM$

It is worth mentioning that the boundary lines in Figure 5b were arbitrarily selected in this case as well as in [11] and [14]. Hence, other studies may use different thresholds according to their specific needs. To find the proper thresholds, one feasible way is to include some parameters with known sensitivity information in the EE calculation, so the boundaries can be determined based on the corresponding SI of these parameters.

To sum up, based on the above SA results, the variations of Parameters 1, 2, 3 (from car-following model), 4 (from lane-changing model), and 13 and 14 (from lane model) are prone to have the greatest impact on travel time as found in this case study. It is rather difficult to explain the choice of the six parameters from the physical point of view, since like many other commercial simulators, the non-open source nature of VISSIM also hinders the model users to learn its internal calculation process. Yet, this further highlights the value of the SA approach proposed here, through which the complex relationships among parameters in this “black box” can be efficiently discovered.

Finally, we want to caution that the above case study is just used to illustrate the application of the proposed quasi-OTEE approach. The influential parameters discussed here may not necessarily be the same ones in other VISSIM SA studies, as different settings of the

same model (e.g., changes in network layout) could drive minor changes in the results. For example, the SA performed in the studies [26] and [27] were also based on the travel time of the same model, but used a slightly modified network layout, different signal control logic, and simulation resolutions. Hence, the found sets of influential parameters were slightly different (e.g., although Parameters 1, 2, 3 and 13 also belong to the set of most influential parameters in [26] and [27], their ranks were not exactly the same). In addition, as the original EE approach also takes the statistical distribution assumed for each input parameter as an uncertain factor into account (for details see [2]), the parameters' ranges and distributions (see Table 2) can also have strong impacts on the final SA results.

Nevertheless, this case study has shown that the quasi-OTEE approach is able to deal with the SA of computationally expensive models such as VISSIM, and it can give reasonable sensitivity results with relatively cheap computational cost, which is ideal when other SA techniques are too expensive to be applied from the beginning.

5. Conclusions

In this paper we developed the quasi-OTEE method based on the Elementary Effects (EE) method. Through computing and comparing the Sensitivity Indexes, this approach can efficiently screen the most influential parameters in a complex model.

The two validation experiments showed that: (i) Sampling with quasi-OT is much more efficient than with the original OT [13]. (ii) The quasi-OT can achieve better coverage of the input space than other OT modifications proposed in the literature (e.g., [14]). (iii) The quasi-OTEE approach has a very high precision in identifying the important parameters.

The SA performed for the Zurich network in VISSIM, a complex microscopic traffic model with many parameters, further highlighted the advantages of this method. By applying the quasi-OTEE approach, the most influential parameters in the case study were properly identified, while the computation time was greatly reduced.

To summarize, the quasi-OTEE approach can be used in finding the important parameters, and deciding which parameters to discard in model calibration. Due to its high efficiency and accuracy, it is ideal for the preliminary SA of computationally expensive models, especially when those quantitative, yet more complex SA techniques (e.g., variance-based approaches) are unfeasible. It can be used as a stand-alone tool when the purpose of the study is to find the important parameters rather than their sensitivity ranks. It could also be used to set the foundation for other quantitative techniques if a more refined analysis is desired afterwards (e.g., investigate the non-negligible interactions, identify the correct sensitivity ranks of influential parameters). Furthermore, since the quasi-OTEE is a general SA approach, independent of any specific traffic model, it can be applied as a practical and efficient screening tool for any computationally expensive model in the wider scientific community.

The future research could be devoted to converting this screening approach into a quantitative SA approach based on the same design and sampling process, so that the results from SA could actually be further used for more in-depth investigations of the input parameters.

Acknowledgments

The authors are very grateful to Dr. Biagio Ciuffo from the Joint Research Center of European Commission, for providing valuable information on sensitivity analysis, and computing the results of the variance-based approach used as reference in Section 3 for the validation of the quasi-OTEE method.

We also would like to thank Dr. Christian Heimgartner and the City of Zurich, for providing us the VISSIM model in Section 4 for the SA experiment.

Research contained within this paper also benefited from participation in EU COST Action TU0903 – Methods and tools for supporting the Use caLibration and validaTion of Traffic simUlation moDEls.

References

- [1] MULTITUDE. State-of-the-art: Review of traffic data collection and estimation techniques and review of methodologies for traffic estimation, calibration and validation. MULTITUDE report (in press), (2011).
- [2] B. Ciuffo, V. Punzo and E. Qualietta, “Kriging meta-modelling to verify traffic micro-simulation calibration methods”, Proceedings of the 90th Transportation Research Board Annual Meeting, Washington DC., USA, (2011) January 22-26.
- [3] A. Saltelli, M. Ratto, T. Anres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana and S. Tarantola, “Global Sensitivity Analysis - The Primer”, John Wiley & Sons Ltd, Cheichester England, (2008).
- [4] N. E. Lownes and R. B. Machemehl, “VISSIM: a multi-parameter sensitivity analysis”, Proceedings of the 2006 Winter Simulation Conference, Monterey, California, USA, (2006) December 3-6.
- [5] T. V. Mathew and P. Radhakrishnan, “Calibration of microsimulation models for nonlane-based heterogeneous traffic at signalized intersections”, Journal of Urban Planning and Development, vol. 136, no. 1, (2010), pp. 59-66.
- [6] F. Cunto and F. F. Saccomanno, “Calibration and validation of simulated vehicle safety performance at signalized intersections”, Accident Analysis and Prevention, vol. 40, no. 3, (2008), pp. 1171–1179.
- [7] A. Beegala, J. Hourdakakis and P. G. Michalopoulos, “Methodology for performance optimization of ramp control strategies through microsimulation”, Transportation Research Record: Journal of the Transportation Research Board, No. 1925, Transportation Research Board of the National Academies, Washington DC, (2005), pp. 87-98.
- [8] B. Ciuffo, V. Punzo and V. Torrieri, “A framework for calibrating a microscopic simulation model”, Proceedings of the 86th Annual Meeting of the Transportation Research Board, Washington DC., USA, (2007) January 21-25.
- [9] V. Punzo and B. Ciuffo, “How parameters of microscopic traffic flow models relate to traffic conditions and implications on model calibration”, Transportation Research Record: Journal of the Transportation Research Board, No. 2124, Transportation Research Board of the National Academies, Washington DC, (2009), pp. 249-256.
- [10] B. Bartin, K. Ozbay, O. Yanmaz-Tuzel and G. List, “Modeling and simulation of unconventional traffic circles”, Transportation Research Record: Journal of the Transportation Research Board, No. 1965, Transportation Research Board of the National Academies, Washington DC, (2006), pp. 201-209.
- [11] Z. Li, H. Liu and K. Zhang, “Sensitivity analysis of Paramics based on 2K-P fractional factorial design”, 2nd International Conference on Transportaton Engineering, ASCE, (2009), pp. 3633-3638.
- [12] M. D. Morris, “Factorial sampling plans for preliminary computational experiments”, Technometrics, vol. 33, no. 2, (1991), pp. 161-174.
- [13] F. Campolongo and A. Saltelli, “Sensitivity analysis of an environmental model: a worked application of different analysis methods”, Reliability Engineering and System Safety, vol. 57, no. 1, (1997), pp. 49-69.
- [14] F. Campolongo, J. Cariboni and A. Saltelli, “An effective screening design for sensitivity analysis of large models”, Environmental Modelling and Software, vol. 22, no. 10, (2007), pp. 1509-1518.
- [15] M. V. Ruano, J. Ribes, A. Seco and J. Ferrer, “An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors”, Environmental Modelling & Software, vol. 37, pp. 103-109.
- [16] M. V. Ruano, J. Ribes, J. Ferrer and G. Sin, “Application of the Morris method for screening the influential parameters of fuzzy controllers applied to wastewater treatment plants”, Water Science & Technology, vol. 63, no. 10, pp. 2199-2206, (2011).
- [17] Nossent. Assessing the convergence of a Morris-like screening method for a complex environmental model. Presentation at 7th International Conference on Sensitivity Analysis of Model Output (SAMO), Nice, France, (2013) July 1-4.
- [18] B. Park and J. D. Schneeberger, “Microscopic simulation model calibration and validation case study of VISSIM simulation model for a coordinated actuated signal system”, Transportation Research Record: Journal of the Transportation Research Board, No. 1856, Transportation Research Board of the National Academies, Washington DC, (2003), pp. 185-192.
- [19] G. Gomes, A. May and R. Horowitz, “Calibration of VISSIM for a Congested Freeway”, Research Reports Proceedings 103238, Institute of Transportation Studies, University of California at Berkeley, (2004).

- [20] S. A. Ahmed, "Calibration of VISSIM to the traffic conditions of Khobar and Dammam", Master Thesis at King Fahd University of Petroleum & Minerals, Saudi Arabia, (2005).
- [21] Z. Z. Wu, J. Sun and X. G. Yang, "Calibration of VISSIM for Shanghai Expressway Using Genetic Algorithm", Proceedings of the 2005 Winter Simulation Conference, Orlando, Florida, USA, (2005) December 4-7.
- [22] B. Park and H. Qi, "Development and Evaluation of a Calibration and Validation Procedure for Microscopic Simulation Models", In Transportation Research Record: Journal of the Transportation Research Board, No. 1934, Transportation Research Board of the National Academies, Washington DC, (2005), pp. 208-217.
- [23] L. Yu, X. M. Chen, T. Wan and J. F. Guo, "Calibration of VISSIM for Bus Rapid Transit Systems in Beijing Using GPS Data", Journal of Public Transportation, vol. 9, no. 3, pp. 239-257, (2006).
- [24] D. M. Miller, "Developing a procedure to identify parameters for calibration of a VISSIM model", Master Thesis at Georgia Institute of Technology, USA, (2009).
- [25] PTV.VISSIM 5.40-03 User Manual. PTV Planung Transport Verkehr AG, Karlsruhe, (2012).
- [26] J. MacQueen, "Some methods for classification and analysis of multivariate observations", Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, vol. 1, (1967), pp. 281-297.
- [27] Q. Ge and M. Menendez, "Sensitivity Analysis for Calibrating VISSIM in Modeling the Zurich Network", 12th Swiss Transport Research Conference, Ascona, Switzerland, (2012) May 2-4.
- [28] Q. Ge and M. Menendez, "An improved approach for the sensitivity analysis of computationally expensive microscopic traffic models: a case study of the Zurich network in VISSIM", Proceedings of the 92th Annual Meeting of the Transportation Research Board, (2013) January 13-17; Washington DC., USA.

Authors



Qiao Ge, received the B.S. degree in civil engineering from Tsinghua University, Beijing, China, in 2006. He received the M.S. degree in geodesy and survey engineering from Tsinghua University, Beijing, China, in 2008, and the M.S. degree in transportation systems from Technical University Munich, Munich, Germany in 2010.

He is currently pursuing the Ph.D. degree in traffic engineering at ETH Zurich, Zurich, Switzerland. His research interests include the sensitivity analysis and the calibration of traffic models, as well as the microscopic traffic simulation of traffic systems based on V2V and V2I communication.



Monica Menendez, joined the Swiss Federal Institute of Technology (ETH) Zurich as the Director of the Research Group "Traffic Engineering" in October 2010. Prior to that, she was a Management Consultant at Bain & Company's San Francisco office. She joined Bain after receiving a Ph.D. and an M.S. in Civil and Environmental Engineering (focus on Transportation – Traffic Flow Theory and Operations) from the University of California, Berkeley in 2006. During her studies at UC Berkeley she received, among other awards, an "NSF Fellowship" in 2002, the "Gordon F. Newell Award" in 2006, and the "Outstanding Graduate Student Instructor Award" in 2006. She also holds a dual degree in Civil Engineering and Architectural Engineering from the University of Miami, from where she graduated Summa Cum Laude in 2002.

Her research interests include traffic flow theory and operations, sustainable transportation, and logistics. She is the author/co-author of multiple publications in the areas of transportation, operations research, and construction management.