A multi-step approach for the global sensitivity analysis of complex traffic simulation models. Application to the MITISIM model.
Submitted August 1st, 2012; Revised November 15th, 2012.

Biagio Ciuffo*
Institute for the Environment and Sustainability
European Commission – Joint Research Centre
Via E. Fermi, 2749 – 21027 Ispra (VA) – ITALY
Ph.: +39.0332.786782
Fax: +39. 0332.785236
E-mail address: biagio.ciuffo@ext.jrc.ec.europa.eu

Carlos Lima Azevedo
LNEC - National Laboratory of Civil Engineering
Transportation Department
Av. Brasil, 101 – 1700-066 Lisbon, PORTUGAL
Tel.: +351 21 844 3541
Fax: +351 21 844 3029
E-mail: cmazevedo@lnec.pt

Words: 7419
Plus 1 Tables: 250
Plus 4 Figures: 1000
TOTAL: 8669

*Corresponding author
ABSTRACT
In this work a multi-step approach for model sensitivity analysis has been discussed and applied to the MTISIM traffic simulation model. Throughout the paper it is argued that the application of sensitivity analysis is crucial for a true comprehension and the correct use of traffic simulation models, but it is also acknowledged that the main obstacle towards an extensive use of the most sophisticated techniques is jeopardized but the high number of model runs usually required, especially in the case of models with a high number of model parameters.

For this reason we have tested the possibility to perform a number of preliminary analyses by grouping model parameters on the basis of possible common features (in particular grouping all the parameters pertaining to the same sub-models) and, then, to use sensitivity analysis to discover which groups of parameters accounts for the highest share of the outputs’ variance. At the end of these preliminary steps a final sensitivity analysis on the parameters pertaining to the most influential groups can be performed to individuate the most important among them.

The proposed methodology has been applied to the MITSIM model (101 model parameters) and has allowed uncovering the role played by the different parameters and by the model stochasticity with 80% fewer model evaluations.
INTRODUCTION

In the last decades complex models have been developed in all scientific fields, also thanks to the continuously increasing computational capabilities. Design problems, processes control and political choices are more and more relying on the results they are able to provide. For this reason it is becoming of crucial importance analysing models, understanding how they work and, in particular, what influences their capability to reproduce physical phenomena. Global sensitivity analysis is the family of tools to be used with this aim.

Together with the uncertainty analysis, sensitivity analysis studies how the uncertainties in the model inputs affect the model response. In this picture, uncertainty analysis quantifies the output variability while sensitivity analysis describes the relative importance of each input in determining this variability (1,2).

The focus is, thus, on the model uncertainty: "What makes modelling and scientific inquiry in general so painful is uncertainty. Uncertainty is not an accident of the scientific method, but its substance" (1). Main sources of model uncertainty can be considered, i) the (in)adequacy of the models to the reality and ii) the (uncertain) model inputs.

Uncertainty due to the inadequacy of models arises from a number of sources like the modelling basic assumptions, the structural equations, the level of discretization, the numerical resolution method, etc. Reducing this part of uncertainty usually requires substantial modifications to the modelling structure and, therefore, can be hardly achieved by the model users, being mainly under the responsibility of model developers.

On the contrary, the portion of the uncertainty due to uncertain model inputs is almost totally on the shoulders of the model users. In order to reduce this part of the uncertainty, it is necessary to reduce as much as possible the uncertainty in the model inputs. Yet, different inputs can have a different clout on the model outputs and so is for the uncertainty they embody. When the number of these inputs increases to several hundreds, as it is the case of many complex models available today, understanding those on which it is necessary to focus the attention for their estimation becomes crucial.

In particular let distinguish between those inputs, which are observable, and those, which are not. Such distinction is crucial as it affects the possibility, or the cost, of reducing the uncertainty they are responsible for:

- As observable we intend those model inputs, which have a measurable equivalent in the reality. Thus they can be directly estimated and used to feed the models. In a traffic microscopic model, examples are: the network characteristics, the traffic lights timing, the traffic composition, the distribution of vehicles size, etc.

- Unobservable inputs are those, which either are hardly measurable, like the OD demand, or have not an actual equivalent in the reality. Concerning the latter case, most of traffic model parameters, for example, either do just not have a physical interpretation i.e. they are simply model constants, or they are deliberately considered uncertain by the modeller. In facts, as traffic models are necessarily only coarse representations of the real system, considering modelling parameters as uncertain inputs is commonly taken to cover both the epistemic uncertainty regarding the unmodelled details of the phenomena and the aleatory uncertainty not predicted by the average models (e.g. the variability in time of driver's behaviour). Such parameters can be therefore only indirectly estimated by means of inverse analysis, calibration, etc.

If it’s clear the effect of reducing the number of observable parameters on the cost and complexity of a transportation study, most of the times it is not clear which is the effect of including non-influential parameters in the calibration of a traffic simulation model.

First of all, when calibrating a transportation/traffic simulation model we mean solving a (constrained) optimization problem in which, i) the objective function is a measure of fit between some traffic measure observed in the reality (e.g. traffic counts, speed, travel times etc.) and the same measure derived from the model, ii) the problem variables are the model parameters (3,4).

Ciuffo and Punzo (4,5) have recently shown which is the effect of perturbing a sensitive and a non-sensitive parameter on the model behaviour, and, moreover, which is effect of including a non-influential parameter in the calibration problem. In particular, they have clearly shown that including a
non-influential parameter in the calibration problem may considerably affect the capability of whatever optimization algorithm to find the optimal parameters’ combination. Therefore, to the authors’ opinion, the identification of the model inputs which the analyst should focus the attention on, is a key step in the whole modelling activity.

In this framework, sensitivity analysis plays a fundamental role

As models are becoming more complex, also global sensitivity analysis techniques have made significant progresses in the last decade. Unfortunately it is common opinion that only a minority of sensitivity analysis practitioners make use of the most sophisticated techniques made available in the recent years (1). In fact, the most commonly adopted approach to sensitivity analysis still remains the One At Time (OAT). OAT measures, are based on the estimation of partial derivatives, and assess how uncertainty in one factor affects the model output keeping the other factors fixed to a nominal value. The main drawback of this approach is that interactions among factors cannot be detected, since they require the inputs to be changed simultaneously. In addition, this approach pertains to a family of sensitivity analysis techniques usually referred to as “local sensitivity analysis”, used to derive information on the behaviour of the model around a certain point (for example around the solution of the calibration problems to ascertain for the stability of such solution) rather than for exploring its input space. For this reason it should not be considered as a good practice (6). However, its simplicity and parsimony makes it the preferred choice for practitioners (7).

This picture holds in particular in the transportation/traffic simulation field. Few example of model sensitivity analysis for the individuation of the most influential model inputs are available.

The OAT approach has been applied with microscopic simulation model by Lownes and Mechemel (8) and Mathew and Radhakrishnan (9) in order, respectively, to prioritize model parameters in terms of their effects on model outputs, and to select the parameters to be calibrated. The effect of travel demand variability on travel times was investigated, using the same approach, by Bloomberg and Dale (10). In Kesting and Treiber (11) the same approach is also followed in order to get additional insight on the meaning of the values of parameters resulting from the calibration of two car-following models.

Another technique applied for the sensitivity analysis of traffic simulation model is the analysis of variance (ANOVA). It is a model independent probabilistic sensitivity analysis method used for determining whether there is a statistical association between an output and one or more inputs.

In ANOVA, model inputs are referred to as factors and their values are referred to as factor levels. An output, instead, is referred to as a response variable. Multifactor ANOVA studies the effect of two or more factors on the response variable and it is used to determine both the first-order and the interaction effect between factors and the response variable.

To apply this technique a number of evaluations of the responses against different values of the input parameters are required. From a statistical point of view, an appropriate way of performing these evaluations is defined by the experimental design techniques. In particular, a full factorial design can be properly applied in this case. In such a way, the full factorial experimental plan consists of \( n^k \) model evaluations, where \( k \) is the number of factors and \( n \) is the number of levels. The use of this experimental plan can also determine whether the factors interact with each other, that is, whether the effect of one factor depends on the levels of the others. The results that can be obtained from the ANOVA are twofold. First, they give an estimation of the model output variance explained by each parameter or by their combination. On the basis of this result, it is then possible to use a Fisher probability distribution to test the null hypothesis that the variance explained by a single parameter is negligible with respect to the whole model, that is, that the model is not sensitive (with a well-defined level of significance) to parameter changes. For further details on experimental design techniques and on the ANOVA implementation, the interested reader could refer to technical books such as (12)

ANOVA has been used in Bartin et al. (13) and Li et al. (14) to draw inference about the first order effect of the parameters of Paramics simulation model. Interaction effects were not captured since a two level full factorial design was adopted. A three level factorial design was used instead in Beegala et al. (15), in Ciuffo et al. (16) and in Punzo and Ciuffo (5,24) for the AIMSUN model. However, second order interactions effects of parameters could be evaluated only in the last two
studies where a full factorial design was adopted, on the contrary of Beegala et al. who adopted a fractional design.

Another study that used ANOVA to undertake a model sensitivity analysis was carried out by Park and Qi, (17): ANOVA was used to select parameters needing calibration. Five levels per parameter were taken into account. A Latin hypercube sampling algorithm was used to define the experimental design. However, also in this case, the interaction effect of the parameters was not evaluated.

Further than using the standard definition of ANOVA, a more efficient method based on variance decomposition as well can be used for model sensitivity analysis. This method consists of evaluating two types of sensitivity indices (that will be detailed in the remainder of the paper) and probably represents the most advanced and conceptually sound way of performing model sensitivity analysis. With respect to the experimental design used in ANOVA, the Monte Carlo approach ensures a more thorough exploration of the space of the model inputs. In traffic modeling, this approach has been used by Punzo and Ciuffo (18) for the sensitivity analysis of two car-following models. All the mentioned works refer to applications with either models with few parameters or considering just a sub-set of them. In particular, when dealing with complex traffic simulation models, it is, indeed, common practice making a selection of the parameters to involve in the analysis. The selection is based on a priori knowledge of the model, on developers advises or on common sense. This, however, is a fairly dangerous practice, as many interactions among groups of parameters may remain hidden even to the most expert model users. The problem is that complex traffic simulation models involve dozens of parameters and a sensitivity analysis would require too many model evaluations (already with 20 parameters an ANOVA based on a two levels full factorial design would require more than $10^6$ model evaluations for evaluating just the first order effect of each parameter). In fact, the access to both new advanced modelling techniques and detailed traffic and behavioural data, is increasing the level of detail of new and updated traffic simulation models, such as HUTSIM (19), SimMobility-ST (20) or MATSim (21). Furthermore, traffic simulators are increasingly being applied in many different traffic situations and consistency with the available data needs to be assured (22). These challenges have been linked to the need of a consistent understanding of the simulators performance, along with appropriate calibration and validation procedures. As already pointed out, the generic calibration of traffic models relies directly on the choice of the subset of parameters to calibrate. Although the importance of the accurate identification of this subset of parameters is well reported (23), only a few number of recent studies focus on the systematization of such procedures (18,24).

To deal with this issue, in the present paper we propose to use a multi step approach for model sensitivity analysis. In a first step, parameters are grouped with respect to the sub-models they are part of and a sensitivity analysis is carried out considering the different groups rather than the different parameters. Once that, from the sensitivity analysis, the most influential groups (namely the most influential sub-models) on the model outputs are individuated, a new sensitivity analysis on the parameters of the sub-models identified can be carried out. Again, if from the first screening still too many parameters result involved, in the second step one of the simplified approaches to model sensitivity analysis can be adopted and just after a further reduction of the parameters, a final sensitivity analysis can identify the subset of model inputs to be estimated with particular care.

In the present paper this methodology is applied to an advanced driver behavior model, MITSIM, which involves more than one hundred model parameters (25,26,27). Its integrated driver behavior model presented in Toledo et al. (25,26) is of particular interest due to the high interaction of all advanced models describing the driver behavior. This model has been successfully applied in different traffic scenarios and simulate complex driver behaviors (25,37,38). Among its parameters, some are usually considered as the most important ones, while some others are usually discarded from calibration and estimation studies, thus proving the usefulness of the approach adopted.

The paper is organized as follows. In the next section a brief description of the MITSIM model and of its sub-models is provided. Then, in the third section, the methodology adopted for the model sensitivity analysis with the description of the techniques used is presented. The case study is then described and the results of the analyses outlined and commented. A concluding section is then provided with also reference to the further steps of this analysis.
A HIGH DIMENSIONAL PROBLEM: MITSIM

The integrated driver behaviour model integrates four levels of decision-making: target lane, gap acceptance, target gap and acceleration, in a latent decision framework based on the concepts of short-term goal and short-term plan (see FIGURE 1).

FIGURE 1. Structure of the integrated driving behaviour model (26)

In previous research, this model was integrated in the microscopic traffic simulator MITSIM and extended with other sub-components of the driver behaviour such as the reaction time model and the merging, nosing and yielding acceleration models.

In the following sections, a brief review of the studied models is presented. The reader should however refer to (26,27,28) for a full description. All the parameters considered for potential calibration were classified in 15 different groups:

1. When a new vehicle enters the network, it is randomly assigned an update step size which specifies the frequency with which drivers update their driving behaviour. This value is drawn from a truncated normal distribution with mean, standard deviation, lower and upper bounds $\mu_{RT}, \sigma_{RT}, lb_{RT}, ub_{RT}$ (Group 1, 4 parameters: $\mu_{RT}, \sigma_{RT}, lb_{RT}, ub_{RT}$).

2. Different models describe the acceleration behaviour under the various situations. The stimulus-sensitivity framework, which the GM model is based on (29), was adapted for all the acceleration models considered in MITSIM. The car-following model, for both the acceleration and deceleration (see Group 10), is given by equation (23):

$$a_{cf}^g(t) = \alpha_{cf}^g \left[ V(t) \beta_{cf}^g \Delta x(t) - \gamma_{cf}^g \right] k(t) \delta_{cf}^g \Delta V(t) - \tau_{RT} \gamma_{cf}^g + \epsilon_{cf}^g(t)$$

(1)

where, $V$ is the speed of the subject vehicle; $\Delta x$ and $\Delta V$ are the gap and speed difference between the lead and subject vehicles; $k$ is the traffic density downstream of the subject vehicle; $\tau_{RT}$ its driver reaction time; and $\epsilon_{cf}^g$ the random error term. The car-following state depends on the headway between the subject and the front vehicle. In MITSIM each vehicle has its own headway threshold (see Group 10) However, general thresholds as the minimum response distance $\Delta x_{cf}^{min}$ and the general headway lower bound $h_{cf}^{lb}$ are also considered for this specific model (Group 2, 11 parameters: $\Delta x_{cf}^{min}$, $h_{cf}^{lb}$, $\alpha_{cf}^{acc}$, $\beta_{cf}^{acc}$, $\gamma_{cf}^{acc}$, $\delta_{cf}^{acc}$, $\alpha_{cf}^{dec}$, $\gamma_{cf}^{dec}$, $\delta_{cf}^{dec}$, $\epsilon_{cf}^{acc}$, $\epsilon_{cf}^{dec}$, $\rho_{cf}$).
3. When the headway between the subject and the lead vehicle is big enough and the subject vehicle speed is higher than a threshold \( V_{\text{min}} \), the free-flow state is set and the vehicle acceleration is given by:

\[
 a_{ff} = a_{ff}|V_{DS}(t - \tau_{RT}) - V(t - \tau_{RT})| + \varepsilon_{ff}(t) \\
 V_{DS}(t) = \beta_{ff} + V_{\text{lim}} + \gamma_{ff} \delta_{5} + \delta_{k} k(t) + \rho_{ff} V_{f}(t) 
\]

where, \( V \) is the speed of the subject vehicle; \( V_{DS} \) is its desired speed of the subject vehicle; \( V_{\text{lim}} \) is the local speed limit; \( \delta_{5} \) is 1 if the subject vehicle is heavy and 0 otherwise; \( k \) is 1 if the traffic density downstream is equal or less than a threshold \( \theta_{ff} \) and 0 otherwise; \( V_{f}(t) \) is the front vehicle speed; \( \tau_{RT} \) is the driver reaction time; and \( \varepsilon_{ff}(t) \) the random error term (Group 3, 7 parameters: \( V_{\text{min}}, a_{ff}, \beta_{ff}, \gamma_{ff}, \delta_{k}, \rho_{ff}, \theta_{ff} \)).

4. When a vehicle has reached a lane dropping area, it may be tagged as a merging vehicle. In this situation, the acceleration is calculated relaxing the car-following gap limitation and restricting overtaking when using the dropping lane. Upstream (\( \Delta x_{u} \)) and downstream (\( \Delta x_{d} \)) lengths from the beginning section of dropping lane set the total area where a vehicle can be tagged with merging state. The probability of being tagged merging is given by the fixed parameter \( p_{0} \), and only if the number of merging vehicles in the merging area is less than \( n_{\text{max}} \) (Group 4, 4 parameters: \( \Delta x_{u}, \Delta x_{d}, p_{0}, n_{\text{max}} \)).

5. The mandatory lane change state (MLC) is derived from previous models of MITSIM (27). When the general lane changing model proposed by Toledo (25,26) cannot be applied due to the lack of acceptable gaps (dense traffic conditions), an MLC may be initiated, limiting the lane alternatives in the lane choice and gap acceptance models. Additionally, a vehicle may switch to the MLC state only if its current lane is ending or does not connect to the next link in its path. The probability of initiation of such state is derived from the following equation when the distance to the downstream node is less than \( x_{\text{min}} \).

\[
 p_{\text{MLC}} = \exp \left( \frac{-\Delta x^{2}}{(a_{0}^{\text{MLC}} + a_{1}^{\text{MLC}n_{lc}} + a_{k}^{\text{MLC}}(t))^{2}} \right) 
\]

where, \( \Delta x \) is the distance to the downstream node limited by the lower bound \( x_{lb} \); \( n_{lc} \) is the number of lane changes required to reach the target lane; and \( k \) is the lane density. \( \Delta t_{\text{min}} \) is an additional parameter setting the minimum time in lane when tagged for MLC (Group 5, 5 parameters: \( \Delta x_{lb}, a_{0}^{\text{MLC}}, a_{1}^{\text{MLC}}, a_{k}^{\text{MLC}}, \Delta t_{\text{min}} \)).

6. When a vehicle is in nosing state, the lag vehicle is set to yielding with probability \( p_{\text{yes}} \) if it wasn’t previously yielding and \( p_{\text{yes}} \) otherwise (Group 6, 2 parameters: \( p_{\text{yes}}, \rho_{\text{yes}} \)).

7. When a vehicle has decided to change lanes and is in MLC state, a merging model that captures merging by gap creation, either through courtesy yielding of the lag vehicle or nosing of the subject vehicle, may be applied. The probability of a subject vehicle being set to the nosing state is given by:

\[
 p_{\text{nos}} = \frac{1}{1 + \exp \left( \alpha_{\text{nos}} + \beta_{\text{nos}} \Delta V_{-}(t) + \rho_{\text{nos}} \Delta x_{l}(t) + \rho_{\text{nos}} \Delta x_{u}(t) + \rho_{\text{gap}} l_{\text{gap}}(t) + \rho_{n_{lc}} n_{lc}(t) \right)} 
\]

where, \( \Delta V_{-} \) is the relative speed between the subject vehicle and the lead vehicle on the target lane; \( \Delta x_{l} \) is an impact factor depending on both the remaining distance to the point at which the lane change must be completed and on a parameter \( \lambda_{\text{nos}} \); \( l_{\text{gap}} \) is the total gap length; and \( n_{lc} \) is the number of lane changes required to reach the target lane (Group 7, 6 parameters: \( \alpha_{\text{nos}}, \beta_{\text{nos}}, \rho_{\text{nos}}, \lambda_{\text{nos}}, \rho_{\text{gap}}, \rho_{n_{lc}} \)).
8. The application of the nosing model is also restricted by a maximum waiting time before nosing \( t_{\text{nos}} \), a maximum and minimum distance for nosing, \( \Delta x_{\text{max}}^{\text{nos}} \) and \( \Delta x_{\text{min}}^{\text{nos}} \), and a maximum yielding time \( t_{\text{max}}^{\text{yield}} \) for the lag vehicle (Group 8, 4 parameters; \( t_{\text{max}}^{\text{nos}}, \Delta x_{\text{max}}^{\text{nos}}, \Delta x_{\text{min}}^{\text{nos}}, t_{\text{max}}^{\text{yield}} \)).

9. The courtesy yielding alternative is modelled as a fixed probability: \( p_0^{\text{yield}}, p_1^{\text{yield}}, p_2^{\text{yield}} \) and \( p_3^{\text{yield}} \) are the probabilities to yield to none, 1, 2 and three vehicles when tagged as MLC (Group 9, 4 parameters; \( p_0^{\text{yield}}, p_1^{\text{yield}}, p_2^{\text{yield}}, p_3^{\text{yield}} \)).

10. A high share of the simulation stochasticity comes from the driver population heterogeneity: the acceleration model error terms for the car following and free flow behaviour follow a normal distribution with mean zero and standard deviation \( \sigma_{\text{acc}}^{\text{cf}}, \sigma_{\text{acc}}^{\text{ff}} \) and \( \sigma_{\text{head}}^{\text{cf}}, \sigma_{\text{head}}^{\text{ff}} \) respectively; the headway threshold, which rules the choice between car-following and free flow acceleration models, is obtained from a truncated normally distributed with parameters \( \mu_{\text{dv}}^h, \sigma_{\text{dv}}^h \), and lower and upper bounds \( lb_{\text{dv}}^h \) and \( ub_{\text{dv}}^h \) (Group 10, 7 parameters; \( \mu_{\text{dv}}^h, \sigma_{\text{dv}}^h, lb_{\text{dv}}^h, ub_{\text{dv}}^h \)).

11. The target gap acceleration model captures the behaviour of drivers who target a lane change and already chose the corresponding target gap. This formulation is part of the integrated model proposed by Toledo et al. (25):

\[
a^{\text{TG}} = a_g^{\text{TG}} \left[D^{\text{TG}}(t - \tau)\beta^{\text{TG}} \cdot \exp\left(\beta_{\Delta v^+}^{\text{TG}} \Delta V_{\text{lag}}^+(t)\right) \cdot \exp\left(\beta_{\Delta v^-}^{\text{TG}} \Delta V_{\text{lead}}^-(t)\right)\right] + \epsilon^{\text{TG}}(t)
\]

where, \( D^{\text{TG}} \) is the distance to the to the desired position for the target gap \( TG \) (\( TG \in \{\text{backward, adjacent, forward}\} \)) and has different formulations for each of the possible \( TG \) but only depends on one parameter \( \beta_{DP}(\text{see } 26 \text{ for details}); \Delta V_{\text{lag}}^+ \text{ and } \Delta V_{\text{lead}}^-(t) \) are the positive and negative relative target lane leader speeds; \( \tau \) is the driver reaction time; and \( \epsilon^{\text{TG}} \sim \mathcal{N}(0, (\sigma^{\text{TG}})^2) \) is the random error term (Group 11, 13 parameters; \( \beta_{DP}, a_g^{\text{fwd}}, \beta_{\text{lag}}^{\text{fwd}}, \beta_{\text{lead}}^{\text{fwd}}, a_g^{\text{bkck}}, \beta_{\text{lag}}^{\text{bkck}}, \beta_{\text{lead}}^{\text{bkck}}, \sigma_a^{\text{adj}}, \sigma_a^{\text{adj}} \)).

12. The gap acceptance model evaluates the adjacent gaps in the target lane model and decides to switch lanes immediately or not. The adjacent gap is split into lead and lag gaps which both need to be acceptable for the lane change action. A gap is acceptable if it is greater than the corresponding critical gap, which mean is modelled as a random variable following a lognormal distribution (25,26):

\[
\ln(G_{\text{n}}^{\text{cr}}(t)) = a^l + \beta_{\Delta v^+}^{l} \Delta V_{\text{lag}}^l(t) + \beta_{\Delta v^-}^{l} \Delta V_{\text{lead}}^l(t) + \beta_{\text{EMU}}^{l} \text{EMU}^l(t) + a_{\text{v}}^l v_n(t) + \epsilon^l(t)
\]

where, \( G_{\text{n}}^{\text{cr}}(t) \) is the critical \( l \ (l \in \{\text{lead, lag}\}) \) gap; \( \Delta V_{\text{lag}}^l \) and \( \Delta V_{\text{lead}}^l \) are the positive or negative speed difference between the subject vehicle and the lag vehicle on the target lane limited by a threshold \( \Delta v_{\text{max}} \); \( \text{EMU}^l \) is the expected maximum utility of the target gap \( l \) when \( v_n(t) \) is the individual specific error term and \( \epsilon^l \sim \mathcal{N}(0, (\sigma^l)^2) \) the random error term (Group 12, 8 parameters; \( a^{\text{lead}}, a^{\text{lag}}, \beta_{\Delta v^+}^{\text{lead}}, \beta_{\Delta v^-}^{\text{lead}}, \beta_{\Delta v^+}^{\text{lag}}, \beta_{\Delta v^-}^{\text{lag}}, a_{\text{lead}}^{\text{lag}} \text{ and } a_{\text{lag}}^{\text{lag}} \)).

13. At the top of the drivers’ decision tree is the lane choice model. Modelled as a discrete choice problem, the probability of choosing a target lane is computed through a logit formulation using the following utility function (26):

\[TRB 2013 Annual Meeting Paper revised from original submittal.\]
where, $a^{TL}$ is a constant parameter for the target lane $TL$ ($TL \in \{left, current, right\}$); $\delta_{RML}$ is a dummy variable equal to one if $TL$ is the rightmost lane; $V_i$ is the speed of the lead vehicle on $TL$; $\Delta x$ is the gap between the lead and subject vehicles; $\delta_h$ is a dummy equal to one if the traffic density in $TL$ is higher than a threshold; $\delta_{add}$ is a dummy variable equal to one on the presence of bus and heavy good vehicles in $TL$; $\delta_{tail}$ is a dummy variable that captures drivers’ tendency to move out of their current lane if they are being tailgate and it’s equal to one if the backward gap is less than $\Delta x_{\text{floor}}$; $\Delta x_{\text{exit}}$ is the distance from the subject vehicle to the next exit; $\delta_{nic,i}$ are dummy variables equal to one for each $i$ number of lane changes required to reach $TL$, $\delta_{next}$ is a dummy for the need of exiting on the next off-ramp; $n_{add}$ is a dummy for the number of lane changes required from the TL to the off-ramp; $EMU^{TL}$ is the maximum utility of the available gaps in the $TL$ given by the target gap model; $v_n$ is the individual specific error term that captures correlations between observations over time and $e^{TL}$ the random error term (Group 13, 17 parameters: $a^{CL}$, $a^{RT}$, $\beta_{RML}^{TL}$, $\beta_{nic}^{TL}$, $\beta_b$, $\beta_k$, $k_{cell}$, $\beta_{tail}^{TL}$, $\Delta x_{\text{floor}}^{back}$, $\theta_{MLC}$, $\beta_{nic,1}$, $\beta_{nic,2}$, $\beta_{nic,3}$, $\beta_{next}$, $\beta_{add}$).

14. When a driver has decided to switch lanes, the target gap model captures the drivers’ intention on the lane changing decision process, when the adjacent gap is rejected $(25,26)$. The subject vehicle will then adjust its speed and position depending on the chosen target gap. Similarly to the lane choice model, the probability of choosing a target gap is modelled as a logit model using the following utility equation:

$$U^{TG} = a^{TG} + \beta_{\Delta x_{TG}} \Delta x_{TG}(t) + \beta_{l_{TG}} l_{TG}(t) + \beta_\delta \delta_f(t) + \beta_{\Delta v_{TG}} \Delta V_{TG}(t) + a_u^{TG} v_n + e^{TG}(t)$$

(9)

where $\Delta x_{TG}$ is the distance to the target gap $TG$ ($TG \in \{backward, adjacent, forward\}$); $l_{TG}$ is the effective gap length; $\delta_f^{TG}$ is a dummy for the presence of a front vehicle on the current lane; $\Delta V_{TG}$ is the relative gap speed; $v_n$ is the individual specific error term; and $e^{TG}$ the random error term (Group 14, 6 parameters: $a^{wd}$, $a^{back}$, $\beta_{\Delta x_{TG}}$, $\beta_{l_{TG}}$, $\beta_\delta$ and $\beta_{\Delta v_{TG}}$).

15. The origin and destination (OD) matrix is a key input on the variability of the simulation output. In this study the common stochasticity of the OD matrix was analysed by considering a common variance ($\sigma_{OD}^2$) for all OD paths and a distribution factor ($\beta_{OD}$), which determines the percentage of vehicles departing randomly (Poisson distribution instead of constant headway – Group 15, 2 parameters: $\sigma_{OD}$ and $\beta_{OD}$).

The complete list of the parameters considered for the sensitivity analysis is presented in TABLE 1.
TABLE 1 List of parameters considered in the sensitivity analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Definition</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction Time (G1)</td>
<td>$\mu_{RT}$</td>
<td>Reaction time mean (s)</td>
<td>0.5</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{RT}$</td>
<td>Reaction time standard deviation (s)</td>
<td>0.1</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>$w_{RT}$</td>
<td>Reaction time lower truncation (s)</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>$u_{RT}$</td>
<td>Reaction time upper bound (s)</td>
<td>2.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Car Following Model (G2)</td>
<td>$\Delta x_{min}$</td>
<td>Minimum response distance (m)</td>
<td>2.0</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>$h_{cf}$</td>
<td>Headway lower bound for car-following (CF) model (s)</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>$a_{dec}$</td>
<td>Constant parameter of the CF acceleration</td>
<td>0.015</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>$\beta_{dec}$</td>
<td>Speed parameter of the CF acceleration</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>$r_{dec}$</td>
<td>Headway parameter of the CF acceleration</td>
<td>-0.300</td>
<td>-0.125</td>
</tr>
<tr>
<td>Free Flow Model (G3)</td>
<td>$\Delta y_{d}$</td>
<td>Density parameter of the CF acceleration</td>
<td>0.45</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>$\rho_{dec}$</td>
<td>Speed difference parameter of the CF acceleration</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>$a_{cf}$</td>
<td>Constant parameter of the CF deceleration</td>
<td>-0.95</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>$y_{cf}$</td>
<td>Headway parameter of the CF deceleration</td>
<td>-0.75</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>$\rho_{cf}$</td>
<td>Density parameter of the CF deceleration</td>
<td>0.50</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>$\rho_{dec}$</td>
<td>Speed difference parameter of the CF deceleration</td>
<td>0.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Merging Model (G4)</td>
<td>$V_{min}$</td>
<td>Speed threshold for the free-flow (FF) model</td>
<td>5.0</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td>$\theta_{ff}$</td>
<td>Density threshold for the FF model (veh/km/lane)</td>
<td>15.0</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td>$a_{ff}$</td>
<td>Constant parameter of the FF acceleration</td>
<td>0.05</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>$\beta_{ff}$</td>
<td>Constant parameter of the desired speed model</td>
<td>-25.0</td>
<td>-5.0</td>
</tr>
<tr>
<td></td>
<td>$v_{ff}$</td>
<td>Speed parameter of the desired speed model</td>
<td>0.50</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>$\delta_{ff}$</td>
<td>Vehicle type parameter of the desired speed model</td>
<td>-2.00</td>
<td>-0.25</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{ff}$</td>
<td>Density parameter of the desired speed model</td>
<td>5.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Mandatory Lane Change Model</td>
<td>$\Delta x_{ib}$</td>
<td>Upstream distance of the merging point threshold (m)</td>
<td>25.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>$\Delta x_{d}$</td>
<td>Downstream distance of the merging point threshold (m)</td>
<td>50.0</td>
<td>200.0</td>
</tr>
<tr>
<td></td>
<td>$n_{max}$</td>
<td>Number of vehicles allowed in the merging area</td>
<td>4.0</td>
<td>25.0</td>
</tr>
<tr>
<td></td>
<td>$p_{0}$</td>
<td>Probability of aggressive merging</td>
<td>0.10</td>
<td>0.6</td>
</tr>
<tr>
<td>Nosing Model (G5)</td>
<td>$\Delta x_{ib}$</td>
<td>Lower bound of distance to decision point for MLC (m)</td>
<td>75.0</td>
<td>500.0</td>
</tr>
<tr>
<td></td>
<td>$a_{MLC}$</td>
<td>Constant parameter for MLC</td>
<td>500.0</td>
<td>1000.0</td>
</tr>
<tr>
<td></td>
<td>$n_{MLC}$</td>
<td>Number of lane changes parameter for MLC</td>
<td>0.25</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>$a_{MLC}$</td>
<td>Density parameter for MLC</td>
<td>0.75</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>$\Delta t_{min}$</td>
<td>Minimum time in lane in MLC (s)</td>
<td>0.75</td>
<td>1.50</td>
</tr>
<tr>
<td>Yielding Model (G6)</td>
<td>$p_{y, yield}$</td>
<td>Probability to yield if vehicle was previously not yielding</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>$p_{vy, yield}$</td>
<td>Probability to yield if vehicle was previously yielding</td>
<td>0.7500</td>
<td>1.0</td>
</tr>
<tr>
<td>Nosing Model (G7)</td>
<td>$a_{nos}$</td>
<td>Constant parameter of the nosing model</td>
<td>-5.0</td>
<td>-2.5</td>
</tr>
<tr>
<td></td>
<td>$\rho_{nos}$</td>
<td>Lead vehicle speed parameter of the nosing model</td>
<td>0.15</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>$\lambda_{nos}$</td>
<td>Scale parameter of the nosing model</td>
<td>-0.06</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>$\psi_{nos}$</td>
<td>Distance to critical decision point parameter</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>$\rho_{nos}$</td>
<td>Sum of lead and lag gaps parameter of the nosing model</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>$n_{nos}$</td>
<td>Number of lane changes parameter for nosing model</td>
<td>0.35</td>
<td>0.75</td>
</tr>
<tr>
<td>Nosing Control (G8)</td>
<td>$t_{yield}$</td>
<td>Maximum yielding time (s)</td>
<td>10.0</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>$t_{max}$</td>
<td>Maximum waiting time before nosing (s)</td>
<td>15.0</td>
<td>200.0</td>
</tr>
<tr>
<td></td>
<td>$\Delta x_{nos}$</td>
<td>Maximum distance for nosing (m)</td>
<td>100.0</td>
<td>300.0</td>
</tr>
<tr>
<td></td>
<td>$\Delta t_{nos}$</td>
<td>Minimum distance for nosing (m)</td>
<td>10.5</td>
<td>25.0</td>
</tr>
<tr>
<td>Courtesy Yielding (G9)</td>
<td>$p_{yield}$</td>
<td>Probability to not yield to any vehicle</td>
<td>0.05</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>$p_{1}$</td>
<td>Probability to yield to up to one vehicle</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>$p_{2}$</td>
<td>Probability to yield to up to two vehicles</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>$p_{3}$</td>
<td>Probability to yield to up to three vehicles</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>Driver Heterogeneity (G10)</td>
<td>$\sigma_{cf}^{acc}$</td>
<td>Car-following acceleration error term: standard deviation</td>
<td>0.5</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{cf}^{dec}$</td>
<td>Car-following deceleration error term: standard deviation</td>
<td>0.5</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>$a_{ff}$</td>
<td>Free-flow acceleration error term: standard deviation</td>
<td>1.00</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>$\rho_{n}$</td>
<td>Headway threshold mean</td>
<td>2.0</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{nv}$</td>
<td>Headway threshold standard deviation</td>
<td>0.25</td>
<td>1.00</td>
</tr>
<tr>
<td>Model</td>
<td>Parameter</td>
<td>Definition</td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Target Gap Model (G11)</td>
<td>$\theta_{h}^{b}$</td>
<td>Headway threshold lower bound</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>$\theta_{h}^{u}$</td>
<td>Headway threshold upper bound</td>
<td>4.0</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>$\beta_{dp}$</td>
<td>Desired position constant parameter for the forward gap</td>
<td>0.50</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{f}$</td>
<td>Forward gap constant parameter</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>$\beta_{p_{f}}^{f}$</td>
<td>Distance to desired position parameter for the forward gap</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{f}$</td>
<td>Positive speed difference for the forward gap</td>
<td>0.03</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>$\beta_{p_{f}}^{n}$</td>
<td>Negative speed difference for the forward gap</td>
<td>0.10</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{f}$</td>
<td>Standard deviation of the forward gap model</td>
<td>0.25</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>$\beta_{k}$</td>
<td>Backward gap constant parameter</td>
<td>-0.75</td>
<td>-0.40</td>
</tr>
<tr>
<td></td>
<td>$\beta_{d}$</td>
<td>Distance to desired position for the backward gap</td>
<td>-0.5</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>$\beta_{p_{b}}^{f}$</td>
<td>Positive speed difference parameter for the backward gap</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>$\beta_{p_{b}}^{n}$</td>
<td>Negative speed difference parameter for the backward gap</td>
<td>-0.25</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{b}$</td>
<td>Standard deviation for the backward gap acceleration model</td>
<td>0.8</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{b}$</td>
<td>Adjacent gap constant parameter</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{a}$</td>
<td>Standard deviation for the adjacent gap acceleration model</td>
<td>0.15</td>
<td>0.45</td>
</tr>
<tr>
<td>Gap Acceptance Model (G12)</td>
<td>$\alpha_{lead}$</td>
<td>Lead critical gap constant parameter</td>
<td>0.5</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>$\beta_{lead}$</td>
<td>Negative speed difference parameter - lead critical gap</td>
<td>-0.35</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>$\beta_{lead}$</td>
<td>Positive speed difference parameter - lead critical gap</td>
<td>-4.0</td>
<td>-1.5</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{lag}$</td>
<td>Standard deviation of the error term - lead critical gap</td>
<td>0.75</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>$\beta_{lag}$</td>
<td>Lag critical gap constant parameter</td>
<td>0.5</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>$\beta_{lag}$</td>
<td>Negative speed difference parameter - lag critical gap</td>
<td>-0.15</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{lag}$</td>
<td>Positive speed difference parameter - lag critical gap</td>
<td>0.2</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{lag}$</td>
<td>Standard deviation of the error term - lag critical gap</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>$\Delta_{max}$</td>
<td>Maximum speed difference (mph)</td>
<td>4.0</td>
<td>12.0</td>
</tr>
<tr>
<td>Lane Utility (G13)</td>
<td>$\alpha_{cl}$</td>
<td>Current lane constant parameter</td>
<td>2.0</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{rl}$</td>
<td>Right lane constant parameter</td>
<td>-0.5</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>$\beta_{rml}$</td>
<td>Right most lane dummy parameter</td>
<td>-1.50</td>
<td>-0.75</td>
</tr>
<tr>
<td></td>
<td>$\beta_{rcl}$</td>
<td>Front vehicle speed parameter</td>
<td>0.03</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>$\beta_{b}$</td>
<td>Bus following dummy parameter</td>
<td>-0.5</td>
<td>-1.5</td>
</tr>
<tr>
<td></td>
<td>$\beta_{b}$</td>
<td>Front vehicle spacing parameter</td>
<td>0.002</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>$\beta_{b}$</td>
<td>Heavy vehicle in neighbour in lane parameter</td>
<td>-0.350</td>
<td>-0.150</td>
</tr>
<tr>
<td></td>
<td>$\beta_{b}$</td>
<td>Density in lane parameter</td>
<td>-0.015</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>$\beta_{l}$</td>
<td>Tailgate dummy parameter</td>
<td>-5.50</td>
<td>-1.75</td>
</tr>
<tr>
<td></td>
<td>$\Delta_{back}$</td>
<td>Back gap threshold for tailgate dummy (m)</td>
<td>5.0</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td>$\kappa_{floor}$</td>
<td>Density threshold for tailgate dummy parameter</td>
<td>15.0</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td>$\beta_{1}$</td>
<td>One lane change required dummy parameter</td>
<td>-3.5</td>
<td>-1.0</td>
</tr>
<tr>
<td></td>
<td>$\beta_{2}$</td>
<td>Two lane change required dummy parameter</td>
<td>-6.5</td>
<td>-2.5</td>
</tr>
<tr>
<td></td>
<td>$\beta_{3}$</td>
<td>Each additional lane change required dummy parameter</td>
<td>-3.75</td>
<td>-1.25</td>
</tr>
<tr>
<td></td>
<td>$\beta_{exit}$</td>
<td>Next exit, one lane change required parameter</td>
<td>-2.00</td>
<td>-0.75</td>
</tr>
<tr>
<td></td>
<td>$\beta_{exit}$</td>
<td>Next exit, each additional lane changes required parameter</td>
<td>-1.0</td>
<td>-0.5</td>
</tr>
<tr>
<td></td>
<td>$\beta_{exit}$</td>
<td>Distance to exit parameter</td>
<td>-0.50</td>
<td>-0.15</td>
</tr>
<tr>
<td>Target Gap Model (G14)</td>
<td>$\alpha_{f}$</td>
<td>Forward gap constant parameter</td>
<td>-1.2</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>$\beta_{b}$</td>
<td>Backward gap constant parameter</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$\beta_{v}$</td>
<td>Effective gap length parameter</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>$\beta_{a}$</td>
<td>Relative gap speed parameter</td>
<td>-1.5</td>
<td>-1.0</td>
</tr>
<tr>
<td></td>
<td>$\beta_{s}$</td>
<td>Distance to gap parameter</td>
<td>-2.8</td>
<td>-1.8</td>
</tr>
<tr>
<td></td>
<td>$\beta_{p}$</td>
<td>Front vehicle dummy parameter</td>
<td>-2.2</td>
<td>-1.0</td>
</tr>
<tr>
<td>OD (G15)</td>
<td>$\sigma_{2}$</td>
<td>Variance of the OD matrix</td>
<td>0.00</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>$\beta_{od}$</td>
<td>Distribution Factor of the OD matrix</td>
<td>0.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Although almost all previous calibrations of MITSIM considered all demand parameters (OD entries), a small subset of supply parameters (driver behaviour) were
generally considered. The subset was typically defined based on the purpose of each calibration without any statistical analysis. Still, sensitivity analyses were found in a couple of previous studies with MITSIM. Sterzin (30) used an iterated OAT to a set of parameters from four specific models (car-following and free-flow acceleration, gap acceptance and lane utility models) in a previous version of MITSIM, for the analysis of weather factor in a freeway corridor in Virginia, USA. Although using this simplified models, similar constant parameters \( (\alpha^{cL}, \alpha^{RL}) \) of the car-following model, the desired speed constant parameter \( (\beta_{ff}) \) of the free-flow acceleration and the gap acceptance constant parameters \( (\alpha^{lead}, \alpha^{lag}) \) were found to be significant against sensor data (counts and speed). Kurian (31) used experimental design techniques to test the sensibility of eight parameters of the MITSIM car-following model against speed, counts and density sensor measures in a short congested corridor in California. He found a set of parameters of car-following deceleration \( (\mu_{cf}^{lb}, \alpha_{cf}^{des} \text{ and } \gamma_{cf}^{des}) \) to be the most sensitive to his dataset.

As already pointed out, computational constraints limited the scope of the mentioned works. In the following, an approach to simplify the sensitivity analysis, though without making any assumption on the model has been defined.

A MULTI-STEP GLOBSL SENSITIVITY ANALYSIS APPROACH

The following approach applies to any type of traffic simulation model and, in general, to any modelling framework composed by different independent sub-models interacting among each other. It is thought for models in which the total number of parameters makes the direct application of the selected sensitivity analysis technique unfeasible. As it will be detailed in the following, the sensitivity analysis approach used in the present paper is based on the computation of first order and total order sensitivity indices with a variance-based approach (1). The approach requires the evaluation of the model \( N^*(k+2) \) times, where \( k \) is the number of model parameters and \( N \) is the dimension of the Monte Carlo experiment (ranging from few hundreds to many thousands). The methodology proposed is composed by the following steps:

1) group model parameters basing on the basis of their similarities (e.g. parameters pertaining to the same sub-model, having the same physical interpretation and so on);
2) create a map between a number in the range \([0,1]\) (the value assigned to the group) and a combination of values for the parameters within the same group;
3) apply variance-based sensitivity analysis to the groups to individuate those accounting for the highest share of the model variance;
4) consider only the parameters in the influential groups.
   a. If the number is sufficiently small, apply variance-based techniques to the new set of parameters
   b. If the number is still too high, go to step 1.
5) define the set of parameters to include in the subsequent analyses.

Key step in the methodology is represented by point 2. The map between a number in the range \([0,1]\) and a combination of parameters’ values determine the quality of the sensitivity indices. In general it is necessary to have a sufficient exploration of the parameters space. In order to do so, in the present paper, per each group we have identified 1,024 parameters combinations using Sobol’s quasi-random sequences (36). At each step, the sequence define \( k \) numbers in the range \([0,1]\) in order to guarantee the best coverage of the parameters space. Then the value to be assigned to each parameter is extracted from the uniform distribution defined by the ranges reported in Table 1.
Once the 1,024 combinations have been defined, the map between a number in the [0,1] range is simply created assign to each of the combinations the same probability equal to 1/1024.

In the next section, the variance-based method for sensitivity analysis is described.

**Variance-based methods on the Sobol decomposition of variance**

The variance-based method based on the Sobol decomposition of variance is one of the most recent and effective global sensitivity analysis techniques. The original formulation of the method is due to Sobol’ (32,33) where he provided the analytical derivation and the Monte Carlo-based implementation of the concept. The latest setting for its practical implementation, instead, is due to Saltelli et al. (34).

Given a model in the form:

\[ Y = f(Z_1, Z_2, \ldots, Z_r) \]  \hspace{1cm} (10)

With \( Y \) a scalar, a variance based first order effect for a generic factor \( Z_i \) can be written as

\[ V_{Z_i} \left( E_{Z_{-i}} (Y|Z_i) \right) \]  \hspace{1cm} (11)

Where \( Z_i \) is the i-th factor and \( Z_{-i} \) is the matrix of all factors but \( Z_i \). Furthermore it is known that

\[ V(Y) = E_{Z_i} \left( V_{Z_{-i}} (Y|Z_i) \right) + V_{Z_i} \left( E_{Z_{-i}} (Y|Z_i) \right) \]  \hspace{1cm} (12)

Equation (3) shows that if for \( Z_i \) to be an important factor we need that \( E_{Z_i} \left( V_{Z_{-i}} (Y|Z_i) \right) \) is small, that it is to say that the closer \( V_{Z_i} \left( E_{Z_{-i}} (Y|Z_i) \right) \) to the unconditional variance \( V(Y) \) the higher the influence of \( Z_i \).

Thus we may define our first order sensitivity index of \( Z_i \) with respect to \( Y \) as:

\[ S_i = \frac{V_{Z_i} \left( E_{Z_{-i}} (Y|Z_i) \right)}{V(Y)} \]  \hspace{1cm} (13)

Sensitivity indices as in equation (13) can be calculated per each factor and per each factors combination. This, however, would need a huge amount of model evaluations. In order to reduce the efforts required, a synthetic indicator to be coupled with the first order sensitivity index is the total effects index, defined as follows (34):

\[ S_{T_i} = 1 - \frac{V_{Z_{-i}} \left( E_{Z_i} (Y|Z_{-i}) \right)}{V(Y)} = \frac{E_{Z_{-i}} \left( V_{Z_i} (Y|Z_{-i}) \right)}{V(Y)} \]  \hspace{1cm} (14)

Total effects index of the input factor \( i \) provides the sum of first and higher order effects (interactions) of factor \( Z_i \). When the total index is \( S_{T_i} = 0 \) the i-th factor can be fixed without affecting the outputs’ variance. If \( S_{T_i} \neq 0 \) the approximation made depends on the value of \( S_{T_i} \). It is worth noting that while \( \sum_{i=1}^{r} S_i \leq 1 \), \( \sum_{i=1}^{r} S_{T_i} \geq 1 \), both being equal to one only for additive models.

Since the analytical feasibility of traffic flow models limit the use of the formulas for the calculation of the variances reported in equation (3), the application of this method can be effectively performed in a Monte Carlo setting.

**Variance-based methods. Implementation**

The calculation, in a Monte Carlo framework, of the variance-based sensitivity indices presented in equations (13) and (14) has been object of research in the last decades. Different approaches and strategies may provide results with different accuracy and efficiency. The approach adopted in the present work has been specified in (1,34) and can be summarized in the following points:

- two \((N, r)\) matrices of quasi-random numbers (36) are generated. Using the random numbers two matrices of values for the input variables of the model as in equation (10) are generated (called \( A \) and \( B \) in the following).
a set of \( r \) matrices, \( C \), is obtained assembling \( r \) matrices equal to \( A \) except for the \( i \)-th column (with \( i \) varying from 1 to \( r \) among the \( r \) matrices) that is taken from \( B \).

\[
C_i = \begin{bmatrix}
Z_1^{(i)} & Z_2^{(i)} & \cdots & Z_r^{(i)} \\
Z_1^{(i+1)} & Z_2^{(i+1)} & \cdots & Z_r^{(i+1)} \\
\vdots & \vdots & \ddots & \vdots \\
Z_1^{(N)} & Z_2^{(N)} & \cdots & Z_r^{(N)} \\
\end{bmatrix}
\]

for \( i = 1 \ldots r \)

- the model is evaluated for all the \( [N\cdot(r+2)] \) combinations of input variables as given by matrices \( A, B, \) and \( C \) so as to produce the vectors of outputs \( y_A = f(A), y_B = f(B) \) and \( y_{C_i} = f(C_i) \) for \( i = 1 \ldots r \). These vectors are sufficient for the evaluation of all the first order and total effects indices. This is the reason why, the application of this technique for variance-based methods requires \( [N\cdot(r+2)] \). Since \( N \) is usually not lower than 1000, the number of evaluation required by this efficient approach is not, in any case, negligible, especially for complex and expensive models. For this reason, in the common practice, the approach presented in the present paper can be considered relevant.

Sensitivity indices can be then evaluated using the following formulations (34):

\[
S_i = \frac{1}{N} \sum_{j=1}^{N} y_{b}^{(j)} (y_{C_i}^{(j)} - y_{A}^{(j)})
\]

\[
S_{Ti} = \frac{1}{2N} \sum_{j=1}^{N} (y_{A+B}^{(j)})^2 - \left( \frac{1}{2N} \sum_{j=1}^{N} y_{A+B}^{(j)} \right)^2
\]

The choice of \( N \) is the last point to be discussed in this section. There are no universal recipes. \( N \) can vary from few hundred to several thousands. In order to assess if the indices calculated for a given \( N \) are sufficiently stable, it is worth calculating their confidence interval. This can be easily carried out via a parametric bootstrapping. In practice, in order to calculate sensitivity indices in equations (18) and (19), per each step of the process in the range \([1,N]\), the term in the summation at the numerator of both the equation need to be available. Performing a parametric bootstrapping of the indices means sampling \( N' \) combinations of these terms of the same size \( N \) with replacement. The confidence interval will be then created given the distribution of the \( N' \) indices. If the confidence interval will result sufficiently small, then the number of model evaluation can be considered sufficient. In the following, the results of the sensitivity analysis will be presented in this graphical form.
CASE STUDY

Freeway Layout

The network chosen for this study was the A44 road in the region of greater Porto, Portugal. It is a two-lane urban motorway with a total of 3940m and 5 main interchanges.

Located in the south bank of the Douro river, this road represents one of the main south entrances for the commuters living in this south-western region of greater Porto and to heavy vehicles heading to the port of Leixões. Each stretch average length is less than 1km. The A44 is a dual-carriageway motorway, with two 3.50m width lanes, 2.00m width shoulders in each direction, and an additional lane in just one of the five stretches. The two main interchanges at both edges of the A44 road are a cloverleaf interchange, at North, and a trumpet interchange at South. The other three main interchanges between A44 and other local roads are two elevated roundabouts and a partial diamond interchange. The main section has acceleration and deceleration lanes in all interchanges, although often as short as 150m. On and off-ramps are connected to local roads, generally with tight curves, intersections or pedestrian crossings, which tend to significantly reduce vehicle speeds (FIGURE1).

Available data

For the present case study, a data collection campaign was carried out in the first week of May 2011, with the specific purpose of collecting OD related data. The GLS simultaneous method presented by Cascetta et al. (35) was used in the OD flows estimation by combining link counts and a sampled of identified vehicle paths. Samples of OD paths were collected for several periods of the day, through the audio recording of license plate numbers in 10 different stations located at the entries and exits of the A44 road. Simultaneously, video recordings were collected at the same sites for sampling rate estimation. Loop counts aggregated by periods of 5 minutes from automatic detectors on the A44 main
sections were also provided by the road concessionaire for the data collection week. These three data-sets were combined in the simultaneous dynamic estimation of a generic weekday OD dynamic matrix.

**FIGURE 2.** Bar plots of first order and total order sensitivity indices with their 90% confidence intervals on the three GoFs calculated using counts (pictures on the left) and speed (pictures on the right) profiles from all the detectors.

**Computational resources**

For this study, MITSIM was installed under Scientific Linux in a cluster with 80 cores with 1 GB of RAM memory. This resource allowed for the fast processing of a high number of simulations.
Model outputs used for the sensitivity analysis
Since possible model outputs account for time series of counts and speed at 8 different detectors a strategy to aggregate them in a single measure need to be put in place. To this aim, per each simulation we computed a measure of Goodness of Fit (GoF) between real and simulated time series. In order to assess the dependence from the GoF measure selected, we used three of them, namely, the root mean squared error, RMSE, the root mean squared percentage error, RMSPE, and the Theil inequality coefficient, U, (please refer to 4 for their formulation and description). Willing also to assess the differences between the results achieved at different location, we considered 11 different outputs with the GoFs computed on each single detector, on all the detectors of each road direction and on all the detectors of the network. In total, therefore, we performed the sensitivity analysis of 66 different model outputs.

For what concerns the group analysis, groups were identified on the different sub-models of the MITSIM model as defined in Table 1. In the next section, groups will be identified with the number reported in Table 1.

RESULTS
In the present section main results of the model sensitivity analysis are reported. Although interesting, the differences arising from the sensitivity analysis of the model using outputs from different detectors have not been reported in the present version of this paper.

Group Analysis
In Figure 2, results of model sensitivity analysis considering the three different GoF measures calculated on counts and speed from all the detectors are reported. It is clear that counts and speed profiles are not influenced by exactly the same sub-models. In particular, from first order indices (the red bar) counts profiles are mainly driven by the parameters combinations of Group 2 (Car Following Model), 8 (Nosing Control Model) and 13 (Lane Utility Model), while speed profiles are mainly influenced by Groups 2 (Car Following Model), 10 (Driver Heterogeneity Model) and 13 (Lane Utility Model).

In addition groups interactions appear bigger with counts than with speed. In truth, from the analysis of the scatter plots of each group and their combination we recognise that this interaction is mainly due to the model stochasticity. In practice since it was not possible fixing the random seed of the simulation and no averages were computed almost all the parameters result interacting with the others. From the scatter plots this did not appear evident. On the contrary, from the analysis of the effect of the model stochasticity, it resulted accounting for almost 20-25% of variance with counts and 10-15% with speed, in agreement with what presented in Figure 1. This means that, actually, just few groups are influencing model outputs with their main effect or their interactions. The higher impact of model stochasticity using counts than using speed also suggested us to consider the speed as measure for future model calibration/estimation. This is also connected to the fact that traffic counts have been already used for the estimation of traffic demand and, therefore, it would be wise using speed for depicting traffic dynamics.

For what concerns the differences among the three GoFs, we noticed that, overall, the Theil statistic is able to more clearly identify the most influential parameters. For this reason in the following analyses we are only considering it.

Figure 4 also shows the lower capability of traffic counts in depicting traffic dynamics. Indeed results of the sensitivity analysis using counts from the detectors of the two different directions with different traffic conditions are approximately the same, while this is not the case for the speed. In this latter case, indeed, results are quantitatively different with just two groups (2 and 10) explaining more than the 80% of the output variance in the Northbound direction. This further confirmed the necessity of using speed profiles for reproducing traffic dynamics.
FIGURE 3. Bar plots of first order and total order sensitivity indices with their 90% confidence intervals on the Theil inquality coefficient calculated using counts (pictures on the left) and speed (pictures on the right) profiles from detectors of Northbound (pictures on top) and Southbound (pictures below) direction.

Second step: parameter analysis

The group analysis has allowed the identification of the most influential groups on the model outputs. The three selected groups are those influencing the speed profiles, therefore, groups 2, 10 and 13. These groups account for 34 parameters in total, with a consequent reduction of 2/3 in the number of parameters to analyze. Still this number is quite high for a comprehensive variance-base analysis. Another group analysis might be performed. However, we considered the possibility of performing the variance-based sensitivity analysis evaluating only total order sensitivity indices. In fact, as clearly pointed out in Saltelli et al. (34), total order indices reach stability much sooner than first order ones, thus requiring less model evaluations. We therefore tried using a size of the Monte Carlo experiment of 512, thus with 18,432 model evaluations.

Results are reported in Figure 4. In all the three pictures, the relatively narrow confidence intervals show the good quality of the indices’ estimation. As in the previous case, we consider the stochasticity accounting for around 10% of the outputs’ variance. In this way, almost all the parameters can be discarded from the subsequent calibration/estimation analyses. Considering the three pictures from three different detectors’ aggregations it is possible to ascertain that there are five parameters outperforming all the others in accounting for the output’s variance: $\alpha_{\text{dec}}^c$, $\gamma_{\text{dec}}^c$, $\sigma_{\text{dec}}^c$, $\sigma_{\text{dec}}^a$ and $\rho_{\text{dec}}^c$ (pertaining to groups 2 and 10). In addition, three other parameters account for a lower share of the total outputs’ variance: $\sigma_{\text{dec}}^b$, $\beta_{\text{dec}}^a$ and $\beta_{\text{add}}$. Together with their interactions, these parameters are able to account for a high share of the outputs’ variance (estimated to be around 90%), thus sufficient to
provide, once correctly estimated, a correct representation of traffic dynamics with just, around, 10% of uncertainty.

FIGURE 4. Bar plots of total order sensitivity indices with their 90% confidence intervals on the Theil inequality coefficient calculated using speed profiles from detectors of Northbound (picture on top left), Southbound (picture on top right) direction and from all the detectors (picture below).

In synthesis, the proposed approach has allowed identifying in a quantitative and objective way the most important parameters of the MITSIM model. On the specific case-study 8, out of 101 parameters resulted accounting for approximately the 90% of the total variance in the outputs of the model. In addition, the methodology required 88,064 model evaluations instead of the 421,888 (~80%) otherwise required for applying variance based techniques to the whole set of parameters.

CONCLUDING REMARKS
With the increasing complexity of the models involved in the decision making process it is becoming of crucial importance analysing them, understanding how they work and, in particular, what influences their capability to reproduce physical phenomena. Global sensitivity analysis is the family of tools to be used with this aim.

Sensitivity analysis studies how the uncertainties in the model inputs affect the model response. In addition, sensitivity analysis can be used to i) uncover technical errors in the model, ii) identify critical regions in the space of the inputs, iv) establish priorities for research, v) simplify models and vi) defend against falsifications of the analysis.

Unfortunately it is common opinion that only a minority of sensitivity analysis practitioners make use of the most sophisticated techniques made available in the recent years. The problem is that,
even with the most sophisticated sampling strategies, the exploration of the input space required by any
global sensitivity analysis approach requires many model runs to be performed. When the model is
computationally expensive, which is fairly common in the applications, sensitivity analysis becomes
unfeasible. This especially true when the number of model parameters becomes quite high.

To deal with this issue, practitioners usually perform sensitivity analysis on a subset of model
parameters chosen on the basis of their experience. This approach, however, is quite dangerous as
many interrelations may be hidden also to the most expert model users.

On the contrary, in this paper we have evaluated the possibility the possibility to perform a
number of preliminary analyses by grouping model parameters on the basis of their possible common
features (in particular grouping all the parameters pertaining to the same sub-models) and, then, to use
sensitivity analysis to discover which groups of parameters accounts for the highest share of the
outputs’ variance. At the end of these preliminary steps a final sensitivity analysis on the parameters
pertaining to the most influential groups can be performed to individuate the most important among
them.

The proposed methodology has been applied to the high-dimension MITSIM model (101
model parameters) and has allowed uncovering the role played by the different parameters and by the
model stochasticity with 80% fewer model evaluations. The group analysis has allowed individuating
the three most important sub-models, namely the car-following model, the lane utility model and the
drivers’ heterogeneity model. In addition it has allowed choosing among different possible measures of
goodness of fit and among different traffic measures those able to better depict traffic dynamics.

The final sensitivity analysis has then been performed with the last 35 model parameters and
has allowed individuating a group of 8 parameters accounting for almost the 90% of the output’s
variance, with a consequent significant simplification of the subsequent model calibration/estimation
phase.

Further work will be carried out to quantitatively assess the benefits introduced by the
parameters reduction on the calibration problem.

ACKNOWLEDGEMENTS
Research contained within this paper benefited from participation in EU COST Action TU0903
MULTITUDE – Methods and tools for supporting the Use caLibration and validaTIon of Traffic
simUlationmoDEls.

REFERENCES
1. Saltelli, A., Ratto, M., Anres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola,
Models with Point-to-Point Traffic Surveillance. Proceedings of the 88th TRB Annual Meeting,
Washington D.C. 2009
4. Ciuffo, B., Punzo, V., Quaglietta, E. Kriging meta-modelling to verify traffic micro-simulation
calibration methods. Proceedings of the 90th Annual Meeting of the Transportation Research
5. Punzo V., Ciuffo, B. How parameters of microscopic traffic flow models relate to traffic
conditions and implications on model calibration. Transportation Research Records, n. 2124, pp.
249-256. 2009
6. Campolongo, F. Saltelli, A., Cariboni, J. From screening to quantitative sensitivity analysis. A
7. Saltelli, A., Annoni, P. How to avoid a perfunctory sensitivity analysis, Environmental Modelling
and Software, vol. 25, pp. 1508–1517, 2010