Transportation Research Part B 91 (2016) 21–33

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Speed or spacing? Cumulative variables, and convolution of model errors and time in traffic flow models validation and calibration

Vincenzo Punzo*, Marcello Montanino

Department of Civil, Environmental and Architectural Engineering, Università degli Studi di Napoli Federico II, Via Claudio, 21, 80125 Napoli, Italy

ARTICLE INFO

Article history:
Received 15 September 2015
Revised 19 February 2016
Accepted 21 April 2016
Available online 13 May 2016

Keywords:
Cumulative variable
Convolution
Calibration
Validation
Error statistic
Car-following

ABSTRACT

This paper proves that in traffic flow model calibration and validation the cumulative sum of a variable has to be preferred to the variable itself as a measure of performance. As shown through analytical relationships, model residuals dynamics are preserved if discrepancy measures of a model against reality are calculated on a cumulative variable, rather than on the variable itself. Keeping memory of model residuals occurrence times is essential in traffic flow modelling where the ability of reproducing the dynamics of a phenomenon – as a bottleneck evolution or a vehicle deceleration profile – may count as much as the ability of reproducing its order of magnitude. According to the aforesaid finding, in a car-following models context, calibration on travelled space is more robust than calibration on speed or acceleration. Similarly in case of macroscopic traffic flow models validation and calibration, cumulative flows are to be preferred to flows. Actually, the findings above hold for any dynamic model.

1. Introduction

The measures of discrepancy between a simulation and the real world are at the core of any methodology aimed at reducing the uncertainty involved in scientific modelling, as model calibration and validation (in traffic simulation calibration and validation constitute the object of substantial research efforts see e.g. the EU COST Action TU0903-MULTITUDE and related outcomes, Daamen et al., 2013; Brackstone and Punzo, 2014, or the starting TRB's Task Force on Transportation System Simulation; AHB80T, 2015).

As for dynamic models, discrepancy is mainly measured on the interest variables time-series. With this kind of measure, the ability of a dynamic model to reproduce the temporal evolution of a system can be observed. Several error statistics – also known as 'global error statistics' – are generally used to quantify such discrepancy, or degree of match, between simulated and measured time-series. Examples include the sum of squared or absolute errors and all their linear transformations, and Theil's inequality coefficient (for a review about their use in traffic modelling, see Hollander and Liu, 2008; Brackstone and Punzo, 2014; Buisson et al., 2014).

In global error statistics the simulated value of an interest variable at each instant is compared with its measured value at the same instant. However, the temporal evolution of model residuals and its features, among them the autocorrelation

* Corresponding author. Tel.: +39 081 7683948; fax: +39 081 7683946.
E-mail addresses: vinpunzo@unina.it (V. Punzo), marcello.montanino@unina.it (M. Montanino).

http://dx.doi.org/10.1016/j.trb.2016.04.012
0191-2615/

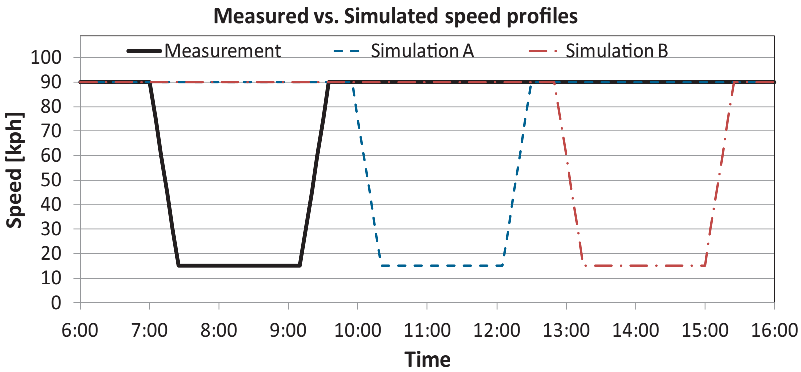

Fig. 1. Comparison between a measured speed time-series ("Measurement") and two simulated speed time-series ("Simulation A" and "Simulation B").

of residuals, do not play any role and, therefore, do not affect the calculation of the error statistic. It is only the magnitude of residuals that matters, not their occurrence time or sequence. For instance, if the sum of squared residuals has to be minimized in calibration, the minimization will be driven by the residuals with the highest magnitude (the residuals being squared) irrespective of their occurrence time. This is a major drawback in dynamic traffic models validation too, where the model ability to reproduce the duration of an event, as a bottleneck, may count as much as the ability to capture its magnitude.

This issue is illustrated in Fig. 1. It shows a time-series of measured speed data and two time-series of simulated speed data. It provides a qualitative representation of time-series of vehicle mean speeds at a spot detector, or of vehicle speed profiles (on a different time scale). The measured speed profile exhibits a drop in the period 7:00–9:30 a.m. while the two simulated speed profiles i.e. Simulation A and Simulation B exhibit a drop that is equal to the drop in the measured profile as for magnitude, but it is shifted in time (for simulations A and B the speed drop occurrence period is respectively 10:00 a.m. to 12:30 p.m. and 1:00–3:30 p.m.). It is easily understood that any of the above mentioned error statistics would yield the same value for both the simulated profiles, as those statistics have no memory of the residuals occurrence time. Actually profiles A and B are very different from one another, and A is closer to the measured profile than B.

With regards to model residuals autocorrelation, in particular, Hoogendoorn and Hoogendoorn (2010) acknowledged that it affects microscopic traffic flow model calibration. To avoid biased calibration results, they suggest an a posteriori transformation of measured and simulated time-series in the 'time-domain' that is proved to eliminate autocorrelation in case of linear models. Montanino et al. (2012) pointed instead at the aforementioned weakness of global error statistics in the time-domain suggesting to conceive error statistics in the 'frequency-domain'. Frequency-domain statistics make the most of the information about residuals autocorrelation in time-series data. Yet in a frequency-domain approach the magnitude of local errors is not directly taken into account.

Therefore, a feasible approach to overcome the issue raised (in the time-domain), i.e. keeping memory of the model residuals order, might be assigning weights to the residuals. In a systematic way, this can be obtained through a convolution of residuals and time. This is the basic idea inspiring this paper.

In general, given two functions f and g, their convolution is defined as the integral of their product after one is reversed and shifted (Damelin and Miller, 2011):

$$(f * g)(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{+\infty} f(t-\tau)g(\tau)d\tau$$

In our case, the discrete convolution of the model residuals time-series $g[k]$ and the time k, in the time interval [1, N], can be written as:

$$(f * g)[h] = \sum_{k=1}^{N} [h-k]g[k]$$

For a shift h equal to the total length N of the time-series the previous equation yields:

$$(f * g)[N] = \sum_{k=1}^{N} [N-k]g[k] \tag{1}$$

The discrete convolution in Eq. (1) represents the linear combination of the model residuals with the reversed discrete time. The first residual is weighted with the length N of the time-series, while the last residual is weighted with one. It is easily understood that the time convolution above allows preserving the memory of model residuals dynamics. The mathematical operation in Eq. (1) is therefore a possible method to overcome the aforementioned weakness of global error statistics.

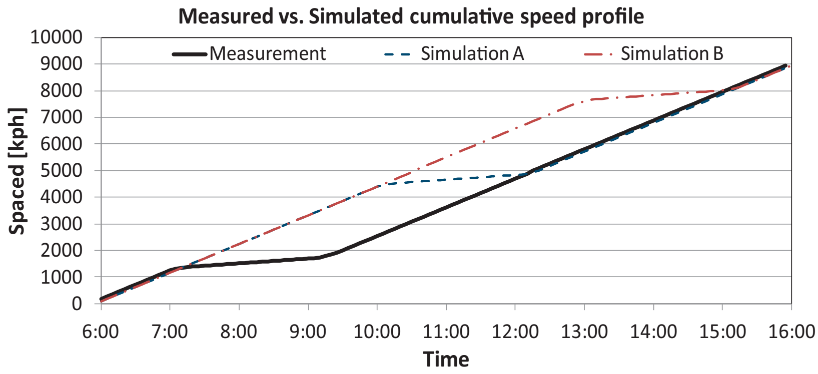

Fig. 2. Comparison between the measured cumulative speed profile ("Measurement") and the simulated cumulative speed profiles ("Simulation A" and "Simulation B").

Actually, in Section 2.2 we show that if in the calculation of a global error statistic a variable is replaced by its cumulative sum, the transformation of model residuals in Eq. (1) is implicitly obtained. In fact, given the two time-series of the model residuals of a variable Z and of its cumulative Y, the Sum of the Absolute or Squared Errors of Y (*SAE* or *SSE*; for their definition see Eqs. (5) and (6) include the time convolution of the residuals of Z. A first contribution of this paper, therefore, is a model for the propagation of model residuals from a time-discrete variable to its cumulative in *SAE* and *SSE* statistics and in all their linear transformations.

In Section 2.3 a specification of such error propagation model for traffic flow models is provided. Said property of a cumulative variable, i.e. keeping memory of the temporal evolution of its derivative residuals, suggests that cumulative variables are to be preferred as measures of performance in traffic flow models too. This finding solves longstanding methodological issues especially in car-following modelling as recalled by the question raised into the title of this paper (see Section 3 for a discussion on the field literature). This is the second major contribution of the paper.

In summary, the paper is organized as follows. In Section 2, a general model for the propagation of errors in *SAE* or *SSE* statistics is first provided. The model is then specified for car-following and macroscopic traffic flow models. The implications for the validation and calibration of traffic flow models of theoretical findings are discussed in Section 3, where a numerical example on car-following is included. As conclusions a summary and a brief discussion are provided.

2. Methodology

2.1. Introduction

In Fig. 2, the cumulative functions of the speed profiles shown in Fig. 1 are reported (in case the profiles in Fig. 1 are the speed profiles of a vehicle, the cumulative functions depicted in Fig. 2 are its measured and simulated trajectories, $x(t)$, i.e. the space travelled by the vehicle until instant t).

From the visual inspection it is easily understood that if we calculated discrepancy measures of these cumulative speed profiles A and B from the cumulative profile of measurements by means of error statistics as *SAE* or *SSE*, we would obtain a value for the statistic of A different by that of B (on the contrary of what happens when calculating these statistics on the speed profiles A and B in Fig. 1). In particular, we would obtain a value for the error statistic of A sensibly smaller than that of B (this result is intuitive when looking at the areas between the two simulated profiles and the measured one). This is the sought result as the speed drop in the simulated profile A is closer to the measured speed drop than that in the profile B. The rationale is clear: a discrepancy measure calculated on the speed profiles has no memory of the model residuals dynamics while it keeps this memory when calculated on the time integrals of the profiles themselves.

Properties of cumulative functions – monotonicity, in particular – and their graphical representation have been already largely exploited in studies on traffic flow theory – see for example Newell (1993a,b), Daganzo (1994), Cassidy (1998) and Cassidy and Bertini (1999) – and on queuing systems, e.g. Newell (1982) (some examples of common applications in transportation problems can be found in Daganzo, 1997).

In this paper, instead, we wish to investigate how the model residuals of a variable propagate to the model residuals of the cumulative sum of the variable in global error statistics. In fact, this will give a new insight in the calibration and validation of traffic flow models.

Therefore in the next subsection we first show how to obtain a general model for the propagation of the residuals from a time-discrete variable Z to its cumulative sum Y in *SAE* and in *SSE*. Then we show how this model can be specified for the case of car-following dynamics or, equivalently, for the case of flow measurements at point detectors.

2.2. General error propagation model

Given a continuous variable $z(t)$ and its time-discrete representation $z[k] \overset{def.}{=} z(k \cdot \Delta t)$, $\forall k \in \{0, ..., N\}$, with Δt the unit time step, we define the variable $y[k]$ as the cumulative sum of $z[k]$ as follows:

$$y[k] = \sum_{i=0}^{k} z[i] \tag{2}$$

In the following, for the sake of convenience, we refer to $z[k]$ as z_k, and to $y[k]$ as y_k.

We then indicate with z_k^{obs} the real world observations of the variable z_k, and with z_k^{sim} its simulated values.

The model error or residual on the variable z_k is defined as: $\varepsilon_k^Z = z_k^{sim} - z_k^{obs}$. Similarly, the model error on the variable y_k is defined as: $\varepsilon_k^Y = y_k^{sim} - y_k^{obs}$.

Assuming that a simulation starts at time $k = 0$, with $\varepsilon_0^Z = \varepsilon_0^Y = 0$, the model error evolution for the variables z_k and y_k can be derived recursively as follows:

$$\begin{aligned}
&\begin{cases} z_1^{sim} = z_1^{obs} + \varepsilon_1^Z \\ y_1^{sim} = y_0^{sim} + z_1^{sim} = y_0^{obs} + z_1^{obs} + \varepsilon_1^Z = y_1^{obs} + \varepsilon_1^Z \end{cases} \\
&\begin{cases} z_2^{sim} = z_2^{obs} + \varepsilon_2^Z \\ y_2^{sim} = y_1^{sim} + z_2^{sim} = y_0^{obs} + z_1^{obs} + z_2^{obs} + \varepsilon_1^Z + \varepsilon_2^Z = y_2^{obs} + \sum_{i=1}^{2} \varepsilon_i^Z \end{cases} \\
&\begin{cases} \ldots \\ \ldots \end{cases} \\
&\begin{cases} z_k^{sim} = z_k^{obs} + \varepsilon_k^Z \\ y_k^{sim} = y_k^{obs} + \sum_{i=1}^{k} \varepsilon_i^Z \end{cases}
\end{aligned} \tag{3}$$

Thus, at the generic step k, the error ε_k^Y on the cumulative variable can be expressed as follows:

$$\varepsilon_k^Y = \sum_{i=1}^{k} \varepsilon_i^Z \tag{4}$$

To evaluate the discrepancy between model outputs and measurements, we may adopt error measures like the Sum of the Absolute Errors, SAE, or the Sum of the Squared Errors, SSE:

$$SAE = \sum_{k=1}^{N} |\varepsilon_k| \tag{5}$$

$$SSE = \sum_{k=1}^{N} (\varepsilon_k)^2 \tag{6}$$

where ε_k stands for ε_k^z or ε_k^y when the error measures are evaluated on $z[k]$ or on $y[k]$, respectively. Accordingly, corresponding error measures will be indicated in the following as SAE^Z, SSE^Z and SAE^Y, SSE^Y.

In case the SAE is calculated on the variables $z[k]$ and $y[k]$, from Eq. (5) it is obtained:

$$\begin{aligned} SAE^Z &= \sum_{k=1}^{N} \left|\epsilon_k^Z\right| \\ SAE^Y &= \sum_{k=1}^{N} \left|\epsilon_k^Y\right| \end{aligned} \tag{7}$$

However, the SAE^Y can be expressed also as a function of $z[k]$ through a recursive application of Eq. (4):

$$\begin{aligned}
&\left|\varepsilon_1^Y\right| = \left|\varepsilon_1^Z\right| \\
&\left|\varepsilon_2^Y\right| = \left|\varepsilon_1^Z + \varepsilon_2^Z\right| \\
&\left|\varepsilon_3^Y\right| = \left|\varepsilon_1^Z + \varepsilon_2^Z + \varepsilon_3^Z\right| \\
&\ldots \\
&\left|\varepsilon_k^Y\right| = \left|\sum_{i=1}^{k} \varepsilon_i^Z\right|
\end{aligned} \tag{8}$$

As $|\sum_{i=1}^{k} \varepsilon_i^Z| \le \sum_{i=1}^{k} |\varepsilon_i^Z|$, it results:

$$\begin{aligned}&\left|\varepsilon_1^Y\right| \le \left|\varepsilon_1^Z\right|\\&\left|\varepsilon_2^Y\right| \le \left|\varepsilon_1^Z\right| + \left|\varepsilon_2^Z\right|\\&\left|\varepsilon_3^Y\right| \le \left|\varepsilon_1^Z\right| + \left|\varepsilon_2^Z\right| + \left|\varepsilon_3^Z\right|\\&\ldots\\&\left|\varepsilon_k^Y\right| \le \textstyle\sum_{i=1}^{k} \left|\varepsilon_i^Z\right|\end{aligned} \tag{9}$$

Summing up the left- and the right-hand-side terms in Eq. (9) and rearranging terms, it yields:

$$SAE^Y \le \sum_{k=1}^{N} (N-k+1)\cdot\left|\varepsilon_k^Z\right| = \sum_{k=1}^{N}\left|\varepsilon_k^Z\right| + \sum_{k=1}^{N}(N-k)\cdot\left|\varepsilon_k^Z\right| \tag{10}$$

Then, recalling the definition of SAE^Z in Eq. (7) and the one of convolution in Eq. (1), it is obtained:

$$SAE^Y \le SAE^Z + \left(k * \left|\varepsilon_k^Z\right|\right)[N] \tag{11}$$

that is the sought relationship between the error statistic calculated on the variable *Z*, SAE^Z, and the same statistic calculated on its cumulative variable *Y*, SAE^Y. It is worth noting that the upper bound for the difference between the values of the two error measures is given by the convolution with time of the absolute residuals of *Z*. Therefore, in the case of the *SAE* and assuming equality in Eq. (11), the *SAE* on a cumulative variable *Y* is given by the sum of two terms: the *SAE* calculated on the derivative of the variable *Y* i.e. *Z*, and the convolution of time and model residuals on the variable *Z*. Eq. (11) explains why a statistic calculated on an integral variable is able to keep memory of the model residual dynamics, as shown in the example of Figs. 1 and 2.

As for the *SSE*, in case it is calculated on the variables $z[k]$ and $y[k]$, from Eq. (6) it is obtained:

$$\begin{aligned}SSE^Z &= \sum_{k=1}^{N}\left(\epsilon_k^Z\right)^2\\SSE^Y &= \sum_{k=1}^{N}\left(\epsilon_k^Y\right)^2\end{aligned} \tag{12}$$

Similarly to the *SAE*, the SSE^Y can be expressed as a function of $z[k]$ through a recursive application of Eq. (4):

$$\begin{aligned}
&\left(\varepsilon_1^Y\right)^2 = \left(\varepsilon_1^Z\right)^2 &&= \left(\varepsilon_1^Z\right)^2\\
&\left(\varepsilon_2^Y\right)^2 = \left(\varepsilon_1^Z+\varepsilon_2^Z\right)^2 &&= \left(\varepsilon_1^Z\right)^2+\left(\varepsilon_2^Z\right)^2 \quad +2\left(\varepsilon_1^Z\varepsilon_2^Z\right)\\
&\left(\varepsilon_3^Y\right)^2 = \left(\varepsilon_1^Z+\varepsilon_2^Z+\varepsilon_3^Z\right)^2 &&= \left(\varepsilon_1^Z\right)^2+\left(\varepsilon_2^Z\right)^2+\left(\varepsilon_3^Z\right)^2 \quad +2\left(\varepsilon_1^Z\varepsilon_2^Z\right)+2\left(\varepsilon_1^Z\varepsilon_3^Z\right) \quad +2\left(\varepsilon_2^Z\varepsilon_3^Z\right)\\
&\left(\varepsilon_4^Y\right)^2 = \left(\varepsilon_1^Z+\varepsilon_2^Z+\varepsilon_3^Z+\varepsilon_4^Z\right)^2 &&= \left(\varepsilon_1^Z\right)^2+\left(\varepsilon_2^Z\right)^2+\left(\varepsilon_3^Z\right)^2+\left(\varepsilon_4^Z\right)^2 \quad +2\left(\varepsilon_1^Z\varepsilon_2^Z\right)+2\left(\varepsilon_1^Z\varepsilon_3^Z\right)+2\left(\varepsilon_1^Z\varepsilon_4^Z\right) \quad +2\left(\varepsilon_2^Z\varepsilon_3^Z\right)+2\left(\varepsilon_2^Z\varepsilon_4^Z\right)+2\left(\varepsilon_3^Z\varepsilon_4^Z\right)\\
&\ldots\\
&\left(\varepsilon_k^Y\right)^2 = \cdots\\
&\ldots
\end{aligned} \tag{13}$$

Summing up the terms on the left hand-side and those on the right-hand-side in Eq. (13), and rearranging, it is obtained:

$$\begin{aligned}SSE^Y &= \sum_{k=1}^{N}(N-k+1)\cdot\left(\varepsilon_k^Z\right)^2 + 2\sum_{k=1}^{N-1}\sum_{i=k+1}^{N}(N-i+1)\cdot\left(\varepsilon_k^Z\cdot\varepsilon_i^Z\right)\\&= \sum_{k=1}^{N}\left(\varepsilon_k^Z\right)^2 + \sum_{k=1}^{N}(N-k)\cdot\left(\varepsilon_k^Z\right)^2 + 2\sum_{k=1}^{N-1}\varepsilon_k^Z\cdot\sum_{i=k+1}^{N}(N-i+1)\cdot\varepsilon_i^Z\end{aligned} \tag{14}$$

Then, recalling the definition of the SSE^Z in Eq. (12) and the one of convolution in Eq. (1), we can express the SSE^Y as a function of ε_k^Z and k:

$$SSE^Y = SSE^Z + \left(k * \left(\varepsilon_k^Z\right)^2\right)[N] + 2\sum_{k=1}^{N-1}\varepsilon_k^Z\cdot\sum_{i=k+1}^{N}(N-i+1)\cdot\varepsilon_i^Z \tag{15}$$

Similarly to Eq. (11), Eq. (15) defines the relationship between the error statistic *SSE* on the variable *Z* and the same statistic on the cumulative variable of *Z* i.e. *Y*. Here the structure of the error model is clear when looking at Eq. (13). The first term in Eq. (15) i.e. SSE^z is given by the squared errors in the diagonal of the first block of terms on the right-hand-side.

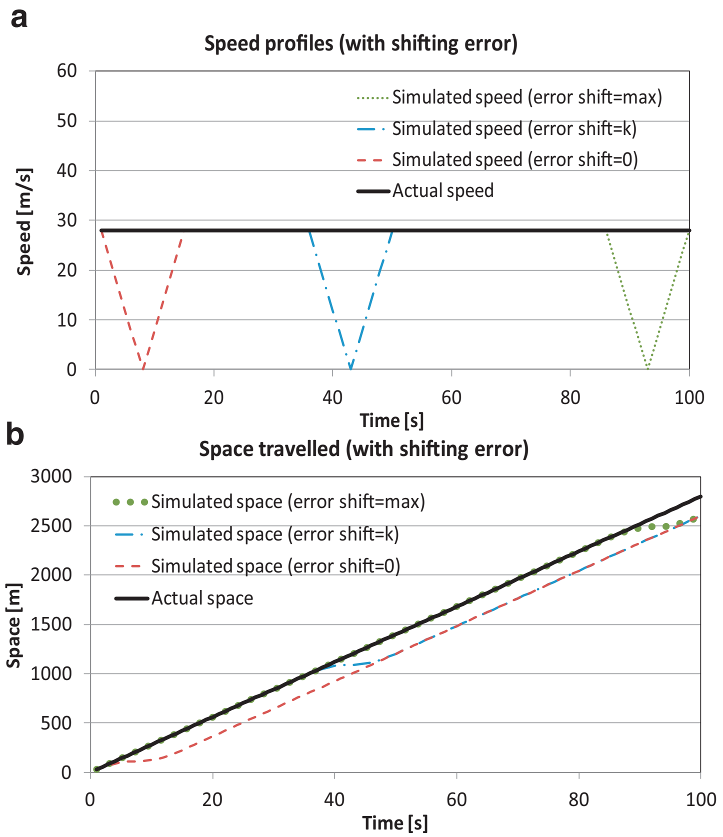

Fig. 3. Speed profiles and space travelled in case of a shifting error. (a) The actual speed profile (black solid line) and three simulated profiles that present an error that is the same as for magnitude but occurring at different times (dashed/dotted lines); and (b) the four corresponding integral functions that is the travelled spaces.

The second term, that is the sum of the remaining squared residuals in the same block, is the convolution of the squared residuals and time. The third term instead, also containing a convolution of errors and time, accounts for the sum of all the double products (and can be negative, of course).

It is worth noting that Eqs. (11) and (15) are model-independent. In fact, they apply to the time-series of an output variable of any model, once it has been simulated. Of course, if the variable Y (that is the cumulative variable in Eq. (2)) is also an input for the model and thus affects its numerical solution, the definition of Y in Eq. (2) must be consistent with the one adopted in the model solution scheme. This is cleared in Section 2.3 where the specifications of Eqs. (11) and (15) for traffic flow models are presented.

It is worth observing also that previous results can be easily extended to any linear transformation of the two error statistics discussed, such as the mean absolute error (*MAE*), the mean absolute normalized or percentage error (*MANE*/*MAPE*), the root mean squared error (*RMSE*), the root mean square normalized or percentage error (*RMSNE*/*RMSPE*). Differently, Theil's inequality coefficient would require a slightly more complex error propagation model.

An intuitive explanation of the above results is provided in Figs. 3–5. In Fig. 3, in particular, an actual vehicle speed profile (the flat line) and three simulated profiles are reported. The three simulated profiles present the same error in speed as for magnitude but shifted in time. It is easily verified that in case of the speed profile with a zero shift, the error on the cumulative variable i.e. on the space travelled is kept constant for almost all the simulation (the error at each instant is given by the vertical distance between the measured space and the simulated space). Instead, for the other two profiles the error accumulation is shorter. It is easily understood that the process of error accumulation in a cumulative variable is given by the convolution terms in Eqs. (11) and (15).

To get further insight into the significance of the three terms in Eq. (15), a notable case is presented in Figs. 4 and 5. The case of simulated speed profiles that present two errors with the same magnitude but with opposite sign is presented. In the following these errors are referred as compensatory errors. The case is interesting because, at a first glance, one might argue that error statistics calculated on cumulative variables are not able to account for such kind of error (a good example is that of two cars following one another at a constant speed: in this case, one might argue that the inter-vehicle spacing before and after the occurrence of two compensatory errors on the speed profile of the follower is the same). In Fig. 4(a) the dashed (red) line represents a speed profile with a zero lag between the positive residual and the negative one; the

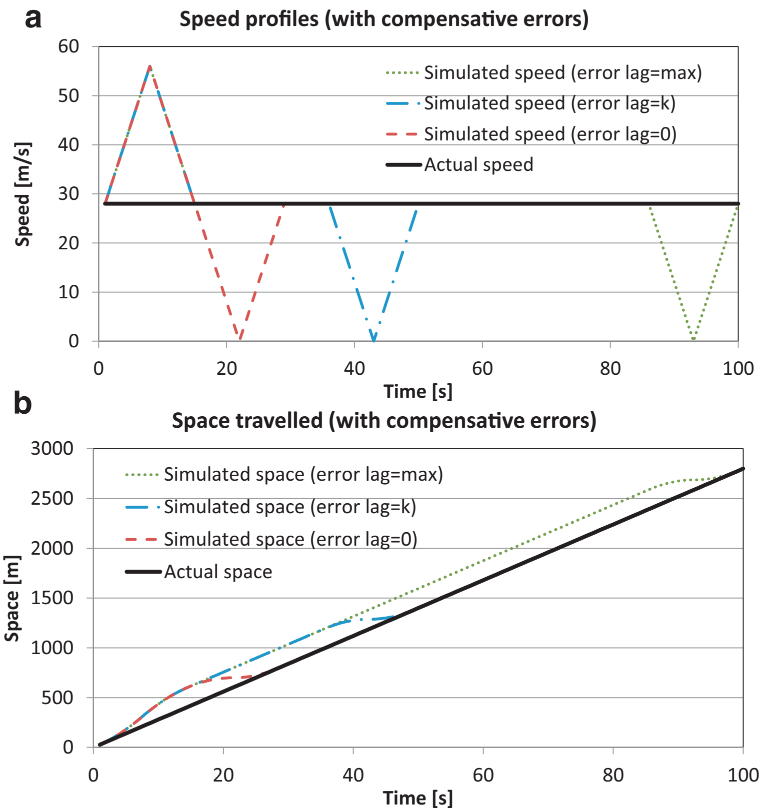

Fig. 4. Speed profiles and space travelled with positive and negative residuals of equal magnitude. (a) Three simulated speed profiles with different lag widths between the two compensatory errors (dashed/dotted lines) and the actual speed profile (black solid line); and (b) the corresponding integral functions that is the travelled spaces. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

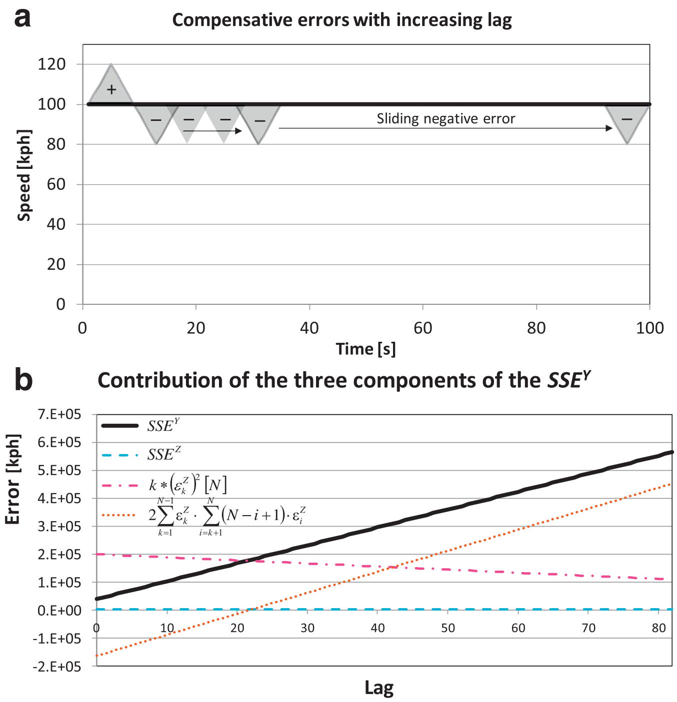

Fig. 5. (a) Simulated speed profiles with increasing lag between two 'compensatory' errors; and (b) individual contribution of each term of Eq. (15), and total value of the SSE^Y, vs. the lag.

dashed dotted (blue) line the case of a k time steps lag, while the dotted (green) line the case of a lag covering the entire simulation period. In fact, in Fig. 4(b) it is easily recognized that the integral functions of the three simulated speed profiles overlap only after the occurrence of the compensatory error (and at the beginning of the simulation). Therefore, any global statistic calculated on the integral measure will return error values increasing with the lag. The same statistic calculated on speed, instead, does not make any distinction between errors taking place in different instants, as already discussed.

This result is generalized in Fig. 5(a) where the case of a negative residual drifting apart from the positive one is depicted. Fig. 5(b), instead, shows the relative contribution of each term in Eq. (15) as a function of the lag. Notably, the first term on the right hand side of Eq. (15), i.e. the SSE^Z, is insensitive to the lag width as explained before. The second term weights early errors more than late ones. Therefore, it decreases with the lag increase that is with the shift of the speed error towards the end of the simulation period (note that the error terms are all positive, being squared). Eventually, the third term, which can include both positive and negative error terms, sharply increases. In fact, with the shift of the negative error towards the end of the simulation period, positive terms increase while negative ones decrease.

2.3. Specification for traffic flow models

In case of microscopic traffic flow models, usually referred as car-following models, the variable $z[k]$ can be thought as the speed of the follower vehicle while the cumulative variable $y[k]$ is the trajectory of the vehicle i.e. the space travelled. In the following, these variables are referred as v_k and x_k. This implies a definition of the cumulative variable different from the one given in Eq. (2), as cleared in the following.

A general formulation for a car-following model can be given as (Wilson, 2008):

$$a_k = f(s_k, v_k, \Delta v_k, \beta) \tag{16}$$

where $\Delta v_k = v_k - v_k^L$ is the relative speed, that is the difference between the follower's (v_k) and the leader's (v_k^L) speeds; $s_k = x_k^L - L_k - x_k$ is the net inter-vehicle spacing between the leader (x_k^L) and the follower (x_k) positions, with L_k the length of the leader vehicle; β is the vector of model parameters, and a_k is the instantaneous acceleration of the follower vehicle at time k returned by the model.

Assuming that a_k is constant between instant k and $k+1$, the follower's speed at time $k+1$ can be derived straightforwardly (see Eq. (17)(a)). The position of the follower vehicle at time $k+1$ has to be updated through the motion equation. In the field literature, mostly used integration schemes to solve such equation are the forward Euler method and the ballistic update (see Treiber and Kanagaraj, 2015) that yield:

$$\begin{cases} v_{k+1} = v_k + a_k \cdot \Delta t & & \text{(a)} \\ x_{k+1} = x_k + v_{k+1} \cdot \Delta t & \text{Eulerian} & \text{(b)} \\ or & & \\ x_{k+1} = x_k + \dfrac{v_{k+1} + v_k}{2} \cdot \Delta t & \text{Ballistic} & \text{(c)} \end{cases} \tag{17}$$

where Δt is the integration step. It is worth observing that the cumulative variable to be used in the derivation of an error propagation model (see Eq. (2) for the general model), has to be the same as the one defined by the model integration scheme i.e. by Eq. (17)(a) or (17)(b).

At the generic instant k, the model errors on speed and position, i.e. ε_k^V and ε_k^X, can be derived following the approach in Eq. (3). With such error definition, the relationship between the SSE^X on the travelled space and the SSE^V on speed, for both the Eulerian and the ballistic integration schemes can be derived as follows (the derivation for the SAE is here omitted for brevity, given the minor relevance of the SAE for the calibration of car-following models; see Punzo et al., 2012).

In case of the Euler method, the follower's position x_k can be intended as the cumulative sum of the space travelled at each interval k, i.e. $\sum_{i=1}^{k} (v_k \cdot \Delta t)$ yielding a result equivalent to that in Eq. (4):

$$\varepsilon_k^X = \Delta t \cdot \sum_{i=1}^{k} \varepsilon_i^V \tag{18}$$

Conversely, a ballistic integration scheme yields:

$$\begin{cases} v_1^{sim} = v_1^{obs} + \varepsilon_1^V \\ x_1^{sim} = x_0^{sim} + \frac{v_1^{sim}+v_0^{sim}}{2} \cdot \Delta t = x_0^{obs} + \frac{v_1^{obs}+v_0^{obs}}{2} \cdot \Delta t + \frac{\varepsilon_1^V}{2} \cdot \Delta t = x_1^{obs} + \frac{\varepsilon_1^V}{2} \cdot \Delta t \\ v_2^{sim} = v_2^{obs} + \varepsilon_2^V \\ x_2^{sim} = x_1^{sim} + \frac{v_2^{sim}+v_1^{sim}}{2} \cdot \Delta t = x_1^{obs} + \frac{\varepsilon_1^V}{2} \cdot \Delta t + \frac{v_2^{obs}+\varepsilon_2^V+v_1^{obs}+\varepsilon_1^V}{2} \cdot \Delta t = x_2^{obs} + \varepsilon_1^V \cdot \Delta t + \frac{\varepsilon_2^V}{2} \cdot \Delta t \\ \ldots \\ \ldots \\ v_k^{sim} = v_k^{obs} + \varepsilon_k^V \\ x_k^{sim} = x_k^{obs} + \Delta t \cdot \left(\sum_{i=1}^{k-1} \varepsilon_i^V + \frac{\varepsilon_k^V}{2} \right) \end{cases} \tag{19}$$

Thus, at the generic step k, the error ε_k^X can be expressed as follows:

$$\varepsilon_k^X = \Delta t \cdot \left(\sum_{i=1}^{k-1} \varepsilon_i^V + \frac{\varepsilon_k^V}{2} \right) \tag{20}$$

On these bases, following the approach described in Eq. (13), the relationships between SSE^X and SSE^V for the Euler method and the ballistic update scheme are, respectively:

$$SSE^X = \Delta t^2 \cdot \left[SSE^V + \left(k * \left(\varepsilon_k^Z\right)^2 \right)[N] + 2 \sum_{k=1}^{N-1} \varepsilon_k^V \cdot \sum_{i=k+1}^{N} (N - i + 1) \cdot \varepsilon_i^V \right] \tag{21}$$

$$SSE^X = \Delta t^2 \cdot \left[\frac{SSE^V}{4} + \left(k * \left(\varepsilon_k^Z\right)^2 \right)[N] + 2 \sum_{k=1}^{N-1} \varepsilon_k^V \cdot \sum_{i=k+1}^{N} \left(N - i + \frac{1}{2} \right) \cdot \varepsilon_i^V \right] \tag{22}$$

Eqs. (21) and (22) are the sought relationships between the SSE calculated on travelled space and the same statistic calculated on speed, for the two integration schemes considered. They are similar to Eq. (15) so that considerations made on structure and terms of the Eq. (15) remain valid.

It is worth noting that in the field of car-following an approach to weight speed errors with traffic regimes (that is deemed to overcome the overestimation of errors at low/high speeds when using normalized/non-normalized discrepancy measures) has been proposed by Kesting and Treiber (2008) (see also Hamdar et al., 2015). In fact, also such measures are linear transformation of Eq. (6) and thus the structures of Eqs. (21) and (22) do not change.

In case of macroscopic traffic flow models, the variable Z can be thought as the flow rate at a point detector ($Z = flow$) while the variable Y is the cumulative number of vehicles arrived at the same detector until time k:

$$Y = cum = \sum_{i=1}^{k} (flow_i \cdot \Delta t) \tag{23}$$

It is straightforward to verify that the derivation of the relationship between SSE^{cum} and SSE^{flow} is the same as the one between SSE^X and SSE^V obtained for the case of the forward Euler method applied to the motion equation in the car-following example, thus yielding again Eq. (21). This result – i.e. the error propagation between a variable and its cumulative counterpart in an error statistic – holds for any measured traffic characteristic.

3. Implications on model validation and calibration

As for validation, Eqs. (21) and (22) and the earlier discussion suggest that, if a dynamic model is to be validated against a specific variable, it is preferable to calculate the statistic on the cumulative of the variable itself. In fact, through this device model residuals dynamics are taken into account and meaningful information on the capability of a dynamic model to reproduce the temporal evolution of a phenomenon is obtained. In case of car-following dynamics, for instance, validating models on the errors made on the space travelled implicitly takes into account the speed errors too, while the opposite is not true. In turn, for macroscopic traffic flow models the use of cumulative flows is preferable to that of flows. In a single link, for instance, at each instant, the sum of cumulative inflows and outflows gives the number of vehicles accumulated in the link at that instant, i.e. the link density (when divided by the link length). Therefore, the time-series of residuals between measured and simulated cumulative flows returns the evolution over time of the error in the link density that represents a meaningful measure of performance for validating dynamic traffic flow models. It is worth observing that when using flows as a measure of performance, instead, two simulated flow profiles returning the same value for an error statistic might depict completely different evolutions of the traffic density on a link.

Also in case of model calibration the adoption of a cumulative variable in the objective function is more effective than the adoption of the variable itself, as shown below. This holds for any dynamic model of a system.

In case of car-following models, for instance, the problem of selecting a variable in the objective function intertwined with many issues hindering the calibration problem solution. These includes the scarceness, incompleteness, or inconsistency of data as to the model complexity (see e.g. Hoogendoorn and Hoogendoorn, 2010; Treiber and Kesting, 2013; Punzo et al., 2015), the measurement errors in the data (Ossen and Hoogendoorn, 2008; Montanino and Punzo, 2015), the computational complexity of the analysis (Ciuffo et al., 2008), the probabilistic modelling of parameters, including their covariance structure (Kim and Mahmassani, 2011) and the asymmetry in the importance of model parameters (Punzo et al., 2015). Overall, these make the calibration problem all but trivial, with over-fitting of models and non-transferability of parameter values as possible drawbacks (Brockfeld et al., 2004; Punzo Simonelli, 2005; van Hinsbergen et al., 2015; Zheng et al., 2013).

Leaving aside these problems and focusing only on the choice of the variable in the measure of discrepancy, feasible variables have been the space travelled by a vehicle (or, equivalently, the spacing between vehicles), the vehicle speed (or, equivalently, the relative speed) and the acceleration. In a ‘global approach’ for calibration (see Treiber and Kesting, 2013), where each single evaluation of the objective function entails a simulation of an entire trajectory, and a Least Squares (LS) estimation is the favourite solving method, speed and spacing are usually preferred to acceleration. In fact, usually being indirectly obtained by position or speed measurements through differentiation, the acceleration is affected by strong noise (high frequency error components amplify in the differentiation process; see Punzo et al., 2011).

However, the choice between speed and spacing has been controversial in the field literature with many studies making use of one variable or the other, and some even of their combination (see e.g. Ossen and Hoogendoorn, 2008; Kim and Mahmassani, 2011). Although some attempts to provide arguments in favour of the spacing have been produced, a sound theoretical interpretation of results has never been provided. For instance, Punzo and Simonelli (2005) first claimed that spacing is preferable to speed as the variable in the objective. The authors provided an intuitive explanation supported by results of a cross-comparison of model calibration against speed and spacing. In Ossen and Hoogendoorn (2007), the numerical analysis showed that the calibration against speed does not allow estimating some of Gipps’ model parameters, confirming a preference for the spacing. Furthermore, the use of a combination of speed and spacing did not improve the calibration results. Kesting and Treiber (2008) also suggest the use of spacing and compare the impact on results of using different error measures in the objective function (see also Hamdar et al., 2015). Eventually, in the exploratory study of Punzo et al. (2012), the preference for spacing was confirmed through substantial empirical evidence.

From a conceptual point of view, if we could rely on a ‘perfect’ car-following model, it would not matter whether to calibrate on speed or space. We would have a unique global minimum, in which the objective function would be null, and a unique set of optimal parameters being the model calibrated either against speed or space. In reality, as no model is error-free we have two different minima and two different set of optimal parameters when calibrating against speed or space. Clearly, if the model is calibrated against speed, we will never find another set of parameters returning a speed profile closer to the real speed profile than the one calibrated. However, we can find a set of parameters that returns a space profile closer to the real space travelled than the one obtained with parameters calibrated against speed. The same holds when calibrating on space, so that either optimality is reached on speed or on space.

The only way to define the dominance of a variable in the objective function over the other is by measuring the robustness of calibration results with respect to the other variable that is making a cross-validation, as proposed in Punzo and Simonelli (2005). In fact, if the model is calibrated against speed, it will be possible to calculate the error it makes in reproducing the spacing, and vice versa. We can therefore calculate the spread, as measured by a relative error statistic, between the spacing resulting from the speed calibration and the optimal spacing i.e. the one obtained calibrating directly against spacing (Eq. (24)(a)), and likewise on speed (Eq. (24)(b)):

$$\begin{aligned} &\frac{RMSE(s,\beta_V) - RMSE(s,\beta_S)}{RMSE(s,\beta_S)} \qquad \text{(a)} \\ &\frac{RMSE(v,\beta_S) - RMSE(v,\beta_V)}{RMSE(v,\beta_V)} \qquad \text{(b)} \end{aligned} \tag{24}$$

where $RMSE(i,\ \beta_j)$, with $i, j \in \{s, v\}$, is the Root Mean Square Error of the variable i obtained by running the model with the optimal parameters calibrated against the variable j.

We performed such cross-validation by calibrating the Intelligent Driver Model (IDM) by Treiber et al. (2000) against both speed and inter-vehicle spacing for all the 2037 trajectories in one of the I80 NGSIM datasets, therefore covering a wide range of driving dynamics. The methodology for the IDM calibration applied here is the one described in Punzo et al. (2015) where all the calibration phases, including criteria for the definition of parameters ranges of car-following models and the set-up of the calibration problem are provided.

It is worth noting that the dataset used here is the ‘reconstructed’ NGSIM I80-1 dataset described in Montanino and Punzo (2015) and not the original one (NGSIM, 2005) which is known to be affected by massive errors that corrupt the accuracy of vehicle kinematics and microscopic traffic dynamics (see Thiemann et al., 2008; Punzo et al., 2011). In fact, the analyses performed in Montanino and Punzo (2015), which also provide the methodology to process the original NGSIM data, unequivocally suggest the use of such new processed dataset. The paper compares, for the two datasets, the impacts of the measurements errors in trajectories on the calibration and validation of traffic flow models (against microscopic and

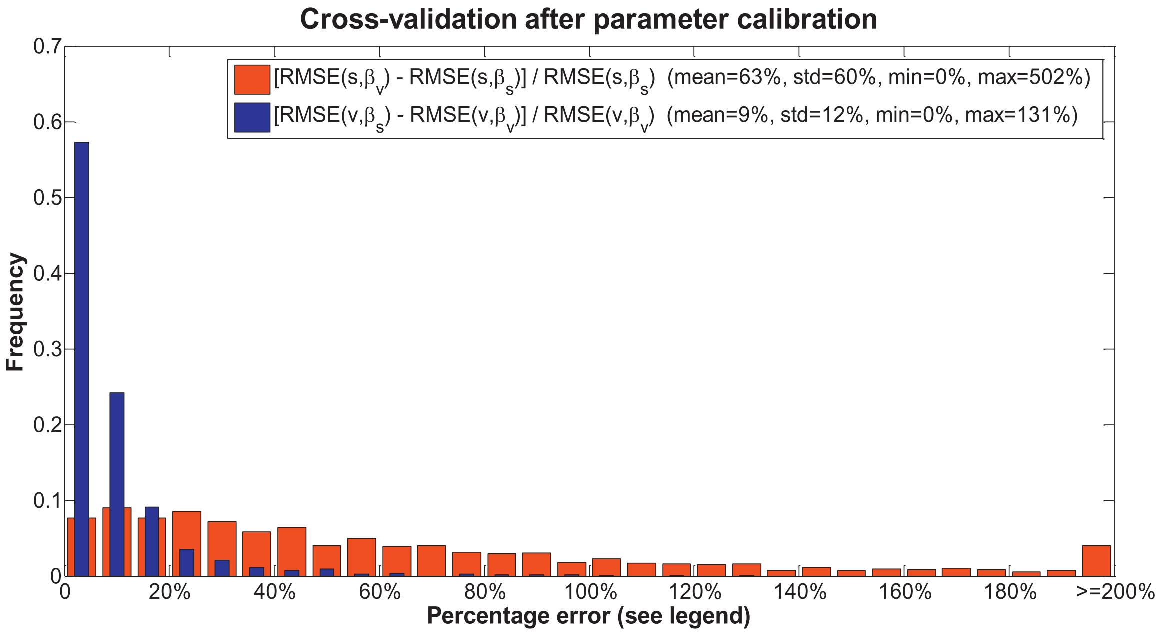

Fig. 6. Cross-validation of calibration results. In red, the histogram of relative errors on spacing obtained when calibrating on speed. In blue, the histogram of relative errors on speed when calibrating on spacing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

macroscopic measures, at both the scale of individual traffic models and traffic simulation) quantifying the benefit for traffic modelling of using the 'reconstructed' dataset (this dataset is available at MULTITUDE, 2014).

The result of the cross-validation experiment is shown in Fig. 6. It can be clearly observed that results of calibration on spacing (see the blue bars) are much more robust than those obtained when calibrating on speed (see the red bars). In fact, the mean relative error on speed that is obtained by calibrating on spacing is equal to 9%, while the mean relative error made on spacing when calibrating on speed is equal to 63%. Similar results hold for the standard deviation.

The explanation comes clear from the findings in the previous section. In fact, calibrating a car-following model against speed or spacing, using the Sum of Squared Errors as measure of discrepancy, means minimizing the following objective functions, respectively:

$$\min\left\{SSE^{V}\right\} \tag{25}$$

$$\min\left\{SSE^{X}\right\} = \min\left\{SSE^{V} + \left(k * \left(\varepsilon_{k}^{Z}\right)^{2}\right)[N] + 2\sum_{k=1}^{N-1}\varepsilon_{k}^{V}\cdot\sum_{i=k+1}^{N}(N-i+1)\cdot\varepsilon_{i}^{V}\right\} \tag{26}$$

It is clear that calibrating on spacing (see Eq. (26)) an algorithm will try to minimize residuals on speed through the SSE^V, as well as, their optimal allocation in time through the other two terms in Eq. (26), that is the time allocation of speed errors that minimizes errors on spacing (see again examples in Figs. 3–5). The result of calibration on spacing will be therefore optimal for spacing and suboptimal for speed. On the contrary, when calibrating on speed (see Eq. (25)), the algorithm will be insensitive to spacing. In fact, the objective function in Eq. (25) makes no difference between speed profiles that have the same SSE^V but different values of the error on spacing i.e. different values of the SSE^X. In this case results of calibration on speed will be optimal for speed and indeterminate for spacing, for the same reasons arising in Figs. 1 and 2.

A further and straightforward mathematical argument in favour of using spacing in place of speed in calibration (see Treiber and Kesting, 2013) is that, assuming that the spacing was perfectly fit by the model (i.e. the model is 'perfect') and that trajectory data were internally and platoon consistent, calibrating on spacing would fit perfectly also the speeds. However, the contrary is not true because of the integration constant.

From the above discussion is also clear that combining speed and spacing in the objective function, through normalized measure of discrepancy, introduces a bias on the results of validation and calibration even in case of adopting equal weights. In case of calibration, in particular, a numerical evidence of the impact of such bias was given in Punzo et al. (2012) for the case of Theil's inequality coefficient.

4. Summary and conclusions

The comparison of model outputs and real world measurements is at the core of the building and corroboration of any scientific model. To this aim time-series of output variables are adopted. Discrepancy between simulated and real world variable time-series is usually measured through global error statistics whose major drawback is to ignore model residuals dynamics. That global statistics keep no memory of the occurrence times of model errors, or of their order, is an unfortunate condition for the evaluation of dynamic models accuracy as models of this kind are chiefly used for their ability to describe the temporal evolution of phenomena.

In this paper we show that this weakness may be solved by considering the convolution of model errors and time. Further we demonstrate that a convolution of this kind can be achieved by replacing a time-discrete variable by its cumulative in global error statistics as the Sum of Absolute Errors or the Sum of Squared Errors (*SAE* and *SSE*). In fact, given the two time-series of the model residuals of a variable Z and of its cumulative Y, the above statistics calculated on Y include the convolution of model residuals of Z and time. Model residuals of a cumulative variable are therefore able to keep memory of the temporal evolution of its derivative model residuals.

To prove this condition a general model for the propagation of model residuals in the above error statistics from a variable to its cumulative has been developed. The model holds for all the linear transformations of *SAE* and *SSE*. It has been specified for the car-following modelling in case of two integration schemes (i.e. Euler method and ballistic update), and in case of macroscopic modelling of traffic flows. The model yields mathematical relationships between the above error statistics applied to a variable and the same statistics applied to the cumulative of the variable itself. It is worth observing that the obtained relationships are model-independent, as they apply to the time-series of output variables of any model once it has been simulated.

These results have significant implications on the validation and calibration of traffic flow models. As for validation, if a dynamic model is to be validated against a specific variable it is recommended to calculate the error statistic on the cumulative of the variable itself e.g. for car-following models, that means on space rather than speed; for macroscopic traffic flow models, on cumulative flows rather than flows. In fact, as a cumulative variable keeps memory of model residuals dynamics it returns more meaningful information on the capability of a dynamic model to reproduce the temporal evolution of a phenomenon than the variable itself.

As for calibration, if a model is to be calibrated on a variable, it is recommended to adopt the cumulative of the variable as a measure of performance in the objective function. In fact, when adopting a variable in the objective function, the calibration is insensitive to the cumulative of the variable itself while it is affected by its derivative. This result provides a solution to a long-standing issue in the calibration of car-following models: inter-vehicle spacing is to be preferred to vehicle speed or acceleration as a variable in the objective function. The impact of this finding has been quantified through an extensive calibration exercise of the Intelligent Driver Model on all trajectories of a NGSIM dataset.

Aknowledgements

The paper highly benefited from the valuable comments of the anonymous reviewers. 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 simUlation moDEls.

References

AHB80T, 2015. TRB's Task Force on Transportation System Simulation (AHB80T) www.tft.ceng.calpoly.edu/operations.htm (last accessed January, 2016).
Brackstone, M., Punzo, V., (Eds.), 2014. Traffic Simulation: Case for Guidelines, JRC Scientific and Technical Report, EUR 26534 EN, Publications Office of the European Union, Luxembourg.
Brockfeld, E., Kühne, R.D., Wagner, P., 2004. Calibration and validation of microscopic traffic flow models. Transportation Research Record 1876, 62–70.
Buisson, C., Daamen, W., Punzo, V, Montanino, M., 2014. Calibration and validation principles. In: Daamen, W., Buisson, C., Hoogendoorn, S.P. (Eds.), Traffic Simulation and Data. CRC Press, Taylor & Francis Group, pp. 89–117.
Cassidy, M.J., 1998. Bivariate relations in nearly stationary highway traffic. Transportation Research Part B 32, 49–59.
Cassidy, M.J., Bertini, R.L., 1999. Some traffic features at freeway bottlenecks. Transportation Research Part B 33, 25–42.
Ciuffo, B., Punzo, V., Torrieri, V., 2008. Comparison between simulation-based and model-based calibrations of traffic flow micro-simulation models. Transportation Research Record 2088, 36–44.
Daamen, W., Buisson, C., Hoogendoorn, S.P. (Eds.), 2013, Traffic Simulation and Data. CRC Press, Taylor & Francis Group.
Daganzo, C.F., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B 28 (4), 269–287.
Daganzo, C.F., 1997. Fundamentals of Transportation and Traffic Operations, Pergamon.
Damelin, S., Miller, W., 2011. The Mathematics of Signal Processing. Cambridge University Press.
Hamdar, S.H., Mahmassani, H.S., Treiber, M., 2015. From behavioral psychology to acceleration modeling: calibration, validation, and exploration of drivers' cognitive and safety parameters in a risk-taking environment. Transportation Research Part B 78, 32–53.
Hollander, Y., Liu, R., 2008. The principles of calibrating traffic microsimulation models. Transportation 35 (3), 347–362.
Hoogendoorn, S., Hoogendoorn, R., 2010. Calibration of microscopic traffic-flow models using multiple data sources. Philosophical Transactions of the Royal Society A 368 (1928), 4497–4517.
Kesting, A., Treiber, M., 2008. Calibrating car-following models by using trajectory data. Transportation Research Record 2088, 148–156.
Kim, J., Mahmassani, H.S., 2011. Correlated parameters in driving behavior models: car-following example and implications for traffic microsimulation. Transportation Research Record 2249, 62–77.
Montanino, M., Ciuffo, B., Punzo, V., 2012. Calibration of microscopic traffic flow models against time-series data. In: Proceedings of IEEE Intelligent Transportation Systems Conference, 16–19 September, 2012. Anchorage, Alaska, USA.

Montanino, M., Punzo, V., 2015. Trajectory data reconstruction and simulation-based validation against macroscopic traffic patterns. Transportation Research Part B 80, 82–106.
MULTITUDE, 2014 Synthesis Technical Report. www.multitude-project.eu/enhanced-ngsim.html (accessed September, 2015).
NGSIM, 2005. Next Generation SIMulation. US Department of Transportation, Federal Highway Administration http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm (accessed 30. 01.16.).
Ossen, S., Hoogendoorn, S.P., 2007. Calibrating Car-following Models Using Microscopic Trajectory Data: A Critical Analysis of Both Microscopic Trajectory Data Collection Methods and Calibration Studies Based on These Data. Report T&P 2006.010. Delft University of Technology, Delft.
Ossen, S., Hoogendoorn, S.P., 2008. Validity of trajectory-based calibration approach of car-following models in presence of measurement errors. Transportation Research Record 2088, 117–125.
Newell, G.F., 1982. Applications of Queueing Theory. Chapman and Hall, New York.
Newell, G.F., 1993a. A simplified theory of kinematic waves in highway traffic, Part I: general theory. Transportation Research Part B 27 (4), 281–287.
Newell, G.F., 1993b. A simplified theory of kinematic waves in highway traffic, part II: Queueing at freeway bottlenecks. Transportation Research Part B 27 (4), 289–303.
Punzo, V., Simonelli, F., 2005. Analysis and comparison of microscopic traffic flow models with real traffic microscopic data. Transportation Research Record 1934, 53–63.
Punzo, V., Borzacchiello, M., Ciuffo, B., 2011. On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data. Transportation Research Part C 19, 1243–1262.
Punzo, V., Ciuffo, B., Montanino, M., 2012. Can results of car-following models calibration based on trajectory data be trusted? Transportation Research Record 2315, 11–24.
Punzo, V., Montanino, M., Ciuffo, B., 2015. Do we really need to calibrate all the parameters? Variance-based sensitivity analysis to simplify microscopic traffic flow models. IEEE Transactions on Intelligent Transportation Systems 16 (1), 184–193.
Thiemann, C., Treiber, M., Kesting, A., 2008. Estimating acceleration and lane-changing dynamics based on NGSIM trajectory data. Transportation Research Record 2088, 90–101.
Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. Physical Review E 62 (2), 1805–1824.
Treiber, M., Kanagaraj, V., 2015. Comparing numerical integration schemes for time-continuous car-following models. Physica A 419, 183–195.
Treiber, M., Kesting, A., 2013. Traffic Flow Dynamics: Data, Models and Simulation. Springer.
van Hinsbergen, C.P.I.J., Schakel, W.J., Knoop, V.L., van Lint, J.W.C., Hoogendoorn, S.P., 2015. A general framework for calibrating and comparing car-following models. Transportmetrica A: Transport Science 21 (in press).
Wilson, R.E., 2008. Mechanisms for spatio-temporal pattern formation in highway traffic models. Philosophical Transactions of the Royal Society A 366 (1872), 2017–2032.
Zheng, Z., Ahn, S., Chen, D., Laval, J., 2013. The effects of lane-changing on the immediate follower: anticipation, relaxation, and change in driver characteristics. Transportation Research Part C 26, 367–379.