



**HAL**  
open science

# Challenges in Applying Calibration Methods to Traffic Models

Peter Wagner, Christine Buisson, Ronald Nippold

► **To cite this version:**

Peter Wagner, Christine Buisson, Ronald Nippold. Challenges in Applying Calibration Methods to Traffic Models. TRB 2016, Transportation Research Board 95th annual meeting, Jan 2016, WASHINGTON D.C, United States. 13 p, 10.3141/2560-02 . hal-01411239v2

**HAL Id: hal-01411239**

**<https://hal.science/hal-01411239v2>**

Submitted on 21 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# CHALLENGES IN APPLYING CALIBRATION METHODS TO TRAFFIC MODELS

## **Peter Wagner\***

German Aerospace Center (DLR)  
Institute of Transportation Systems  
Rutherfordstrasse 2, 12489 Berlin, Germany  
Phone: +49 30 67055 e 237, Fax: +49 30 67055 - 291  
Email: peter.wagner@dlr.de, Web: <http://www.dlr.de/ts/en/>

## **Christine Buisson**

Université de Lyon, F-69622, Lyon, France  
IFSTTAR, LICIT, F-69675 Bron  
ENTPE, LICIT, F-69518 Vaulx en Velin  
Phone: +33 472 047 713, Fax: +33 472 047 712  
Email: christine.buisson@entpe.fr, Web: <http://mocopo.ifsttar.fr>

## **Ronald Nippold**

German Aerospace Center (DLR)  
Institute of Transportation Systems  
Rutherfordstrasse 2, 12489 Berlin, Germany  
Phone: +49 30 67055 - 263, Fax: +49 30 67055 - 291  
Email: ronald.nippold@dlr.de, Web: <http://www.dlr.de/ts/en/>

For Presentation and Publication  
95<sup>th</sup> Annual Meeting  
Transportation Research Board  
January 12 – 16, 2016  
Washington, D. C.

Submission date: July 30, 2015

words: 3 858  
plus 6 figures: 1 500  
plus 1 table: 250  
total count: 5 608  
word limit: 7 500

---

\*corresponding author

## 1 **ABSTRACT**

2        This text looks at calibration and validation as a means to understand traffic flow models  
3 better. It concentrates on the car-following part of it and demonstrates that the calibration of  
4 stochastic models under certain circumstances can become very difficult.

5        Three types of stochasticity are distinguished for microscopic traffic flow models: the one  
6 coming from noisy data, the one coming from the distribution of the parameters describing the  
7 driver's behavior and the one coming from the model itself, when a noise component is added  
8 to a deterministic differential equation governing the vehicle's movements.

9        By using four sub models comprising four different noise terms and an identical deterministic  
10 part this text shows that a calibration with synthetic – and therefore reproducible – data can  
11 lead to awry results. Parameters fitted by the calibration procedure are significantly different  
12 for deterministic and stochastic models. The text makes the conclusion that the stochasticity is  
13 the reason why the parameter estimation of stochastic models fails sometimes. Up to now, the  
14 authors were, unfortunately, not able to propose a solution to cope with this intrinsic pitfall of  
15 genuine stochastic models.

16 *Keywords:* Micro simulation, Car-following models, Trajectory data, Calibration

# 1 INTRODUCTION

2 A lot of work has been conducted recently to improve on the usage of microscopic traffic flow models and  
3 especially their calibration and validation to real world data, for an overview see [4]. This has worked with  
4 an astonishingly degree of quality, i. e. traffic flow models can be calibrated to real data with an r.m.s.-error  
5 of the order of around 10%.

6 While good enough for daily use, it may have a weak spot, and that is its unknown quality when it  
7 comes to extending such a calibrated model to situations and data it has not yet encountered. In order to be  
8 better off in such situations, calibration / validation should be used for another purpose: to help in a better  
9 understanding of these models by showing where they fall short.

10 This is especially important since we are entering an age with a mix of autonomous (of various degrees,  
11 even intelligent cruise control resembles an autonomous car) and human-driven vehicles, and this interaction  
12 is only poorly understood right now. Since the controller of an autonomous vehicle can be modeled with  
13 ease, the human drivers are a more challenging modeling endeavor. And it is important there to have the  
14 correct model, since so far we do not have good data that can be used to calibrate models that deal with this  
15 interaction.

16 Interestingly, as has already been mentioned in [3], performing a good calibration almost eliminates  
17 the differences between the models in terms of their ability to describe real-world phenomena. This result  
18 seems to be robust with respect to different data, different scenarios, different calibration methodologies,  
19 different objective functions, and different measures for the goodness of fit if the process is correctly suited.  
20 Note, however, that it was shown in [17] that some bad combinations (especially when the downhill simplex  
21 method is applied) can lead to bad results. Nevertheless, the homogeneity of the calibration results between  
22 the various models is surprising, and asks for an explanation.

23 The present text makes the hypothesis that this is due to a false treatment of stochastic models. In traffic  
24 flow there are two main types of stochastic models: car following models (CF) and lane changing models  
25 (LC). To demonstrate this in the case of CF we use a few synthetic examples. Using synthetic trajectories  
26 of cars following a synthetic driver according to a known model with given parameters was already done  
27 in [14, 15] and also in [16]. The idea of those two groups of authors was to test the complete calibration  
28 procedure to determine its ability to find the true optimal parameters of the original synthetic car-following  
29 behavior. This was done when measurement errors were added to synthetic trajectories [14, 15] and when  
30 different calibration procedure were used.

31 These papers tested the ability of the calibration procedure to cope with two of the various types of  
32 stochasticity. Indeed looking at the simulation procedure, one have to face three types of it: the noise in  
33 data (which generates an inaccuracy in the observation of the reality), the variability among parameters  
34 characterizing different drivers / vehicle couples and the stochasticity in the model itself. The various pa-  
35 pers referenced above cope with stochasticity in the data (adding noise into synthetic data for example) or  
36 examine the variability of parameters.

37 The aim in this present study is to exemplify that when the stochasticity is inside the model itself, then  
38 the calibration procedure can fail to reproduce the real parameters even with a correct calibration procedure  
39 and no added noise to the speed measurements of the synthetic trajectory. It demonstrates that under not too  
40 exotic conditions the parameter estimation of the calibration process can yield results that have nothing to  
41 do with the real parameters present, while still yielding a reasonable fit. This result is rather similar to what  
42 was already observed in the case of the stochastic LC model in [10]. So far, no remedy is known for work  
43 around such a result.

44 The present paper is organized as follows: Section 2 presents the characteristics of an example data  
45 set with special focus on the noise in empirical data. In contrast, section 3 considers the noise term in  
46 stochastic car-following models themselves. Subsequently, the parameter estimation of four variants of a

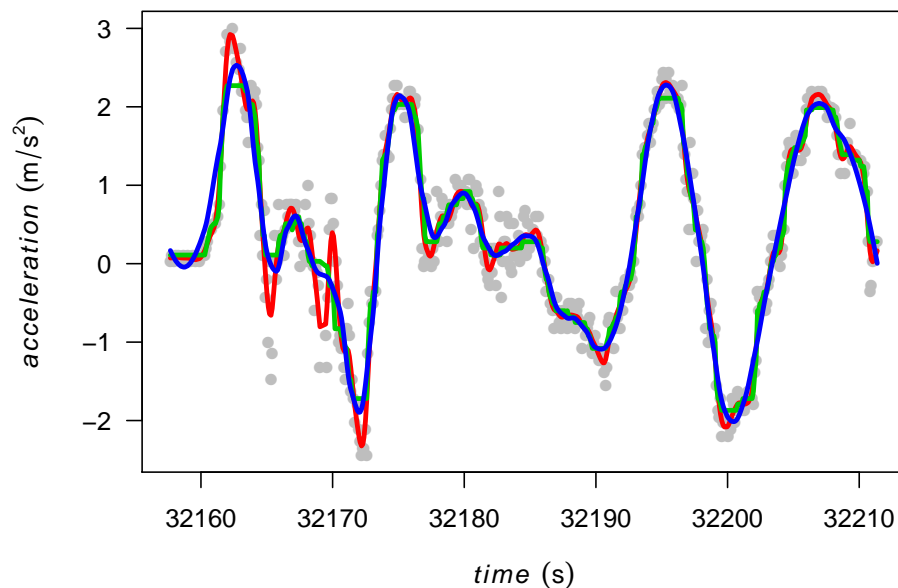
1 simple stochastic car-following model to synthetic trajectory data is analyzed in section 4. Finally, section  
 2 5 concludes the results of this paper.

## 3 2 STOCHASTICITY IN DATA

4 When models face reality, a number of issues have to be considered. All data contain experimental noise to a  
 5 certain extent. To exemplify this, the paper considers data from a large German project named simTD (Safe  
 6 and Intelligent Mobility Test Field Germany [1]), which aimed at a better understanding of communication  
 7 in traffic and sported a fleet of equipped vehicles that drove around in the Frankfurt region for several  
 8 months. Please take in mind that all data collection processes, especially when dealing with trajectory data  
 9 collection, result in noisy data.

10 The data in the simTD project have been sampled from August to December 2012 (on 97 days) from a  
 11 fleet of 125 vehicles that were driven by a few hundred different drivers. Each vehicle was equipped with  
 12 at least an acceleration sensor and a speed sensor, most of the vehicles also had equipment to monitor the  
 13 distance and speed-difference to a lead vehicle. In addition, they had communication devices on board and  
 14 monitored their position via GPS, and vast amounts of data from their CAN bus. All the data, including the  
 15 communication protocol and many more had been recorded by computers in the vehicles and subsequently  
 16 transferred to a common data-base. Note, however, that especially the distances and speed-differences have  
 17 been recorded by an equipment that is also in use for the driver assistance system in those cars. There were  
 18 no special measuring instruments designed for scientific experiments.

19 From this massive data-base (which contains in zipped format 1.3 TByte of data) a few examples have  
 20 been picked to be presented here. E. g., in Figure 1 the raw data from the acceleration sensor is displayed.  
 Since a vehicle is a heavy object, its acceleration versus time curve should be fairly smooth. However, as



**FIGURE 1** A short piece of acceleration as function of time for the vehicle 501 from the simTD data-base on August 27, 2012. The gray points are the raw data. The three curves represent different smoothing methods: the red curve is a spline smoothing, the green one from a median filter, and the blue curve is local polynomial regression of order two to the data. All these methods are implemented in “gnu R” [18].

1 can be seen from the raw data, there is a considerable amount of noise even in the acceleration raw data  
2 (gray points in Figure 1).

3 In Figure 1, different methods for smoothing the acceleration data from this data-base have been com-  
4 pared. Although they yield smoothed approximations to the raw data, they have the disadvantage of many  
5 smoothing method: they make assumptions about the underlying true process, and if one of these assump-  
6 tions is wrong, they fail. Nevertheless, from the different smoothed curves it can be concluded that the  
7 empirical noise in these data is between 0.1 and 0.3 m/s<sup>2</sup>. These numbers have been obtained by computing  
8 the root-mean-square (r.m.s.) distance between the raw data and the smoothed approximations. The value  
9 of 0.1 is for the spline interpolation, which follows the data quite closely, while the value of 0.3 is for the  
10 local polynomial regression, which in the eyes of these authors is the more reasonable choice.

11 To sum up, this section exemplifies that the measurement process produces intrinsically noisy data. The  
12 next section considers the stochasticity caused by the observation that drivers neither behave identically  
13 (this is why parameters are distributed) nor deterministically (this is why stochasticity is incorporated in the  
14 models themselves).

### 15 3 STOCHASTIC COMPONENTS IN MODELS

16 In addition to the noise from empirical data, the models with which traffic flow is described microscopically,  
17 can also have stochastic components. The noise can either be in the dynamic itself, or it is contained in the  
18 parameters: different drivers have different sets of parameters to describe their driving style. This is called  
19 driver heterogeneity. Even more confusingly, the parameters of one and the same driver may be subject to  
20 temporal changes, even during a short time-span.

21 Most of the models that have been described so far, however, are deterministic models. Indeed a model  
22 with no stochastic components in its core equations is deterministic even though its parameters are chosen  
23 randomly. To clarify what that means, the following model (which is modeled after [2]) will be considered:  
24 using  $\dot{v}$  as the acceleration  $a$  of the subject vehicle with speed  $v$  and  $g$  as the net headway to the vehicle in  
25 front, its basic version reads:

$$26 \quad \dot{v} = a = B(\omega^2(g - g^*(v, \Delta v))). \quad (1)$$

27 Here, the function  $B()$  limits the acceleration to fall in the interval  $[-\beta, \alpha(v)]$ , where  $\alpha(v)$  describes how  
28 the maximum acceleration decreases with speed. For this, any model might be acceptable. To be specific  
29  $\alpha(v) = \gamma(v_{\max} - v)$  is often used, thereby introducing the parameters  $v_{\max}$  and  $\gamma$ . See also Figure 2 for an  
30 example how the acceleration values in real data are distributed as function of speed.

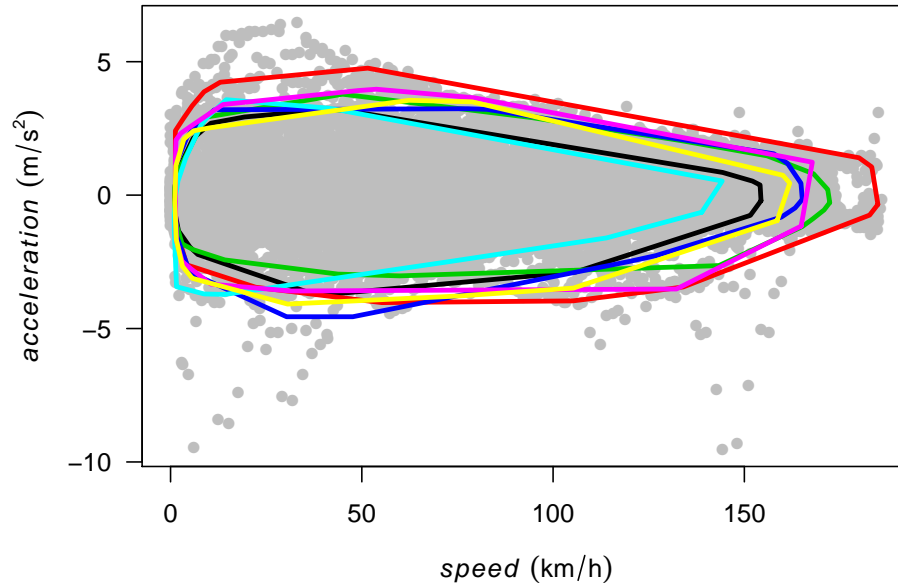
31 The preferred distance  $g^*$  depends on the preferred time headway  $T$  and, in addition, on the speed  
32 difference  $\Delta v = V - v$  between the lead vehicle's speed  $V$  and the following vehicles speed  $v$ :

$$33 \quad g^*(v, \Delta v) = v \left( T + \frac{\Delta v}{b} \right). \quad (2)$$

34 In this equation, the parameter  $b$  is some preferred deceleration the driver typically wants to apply in a  
35 normal car-following situation. This is well different from the parameter  $\beta$  that limits the maximum decel-  
36 eration either to the physically possible one, or to the maximum deceleration the driver applies even in a  
37 critical situation (which is typically smaller than the physical boundary of the vehicle itself).

38 This model has a number of more or less obvious relatives, like the Helly model [7], the Newell model  
39 [13], a cellular automaton model [12], or even a kind of brute force linearization of the Gipps and Krauß  
40 models [6, 11]. It also shares some similarity with the IDM [20].

41 The numbers  $\gamma, \omega$  are constants (for each driver!). To make the units in the equation correct, in addition,  
42 they are inverse relaxation times. So, each driver is described by the set of six parameters  $v_{\max}, b, T, \beta, \gamma, \omega$ .



**FIGURE 2** The convex hull of the acceleration and speed values of seven different drivers (colored lines) together with the raw values (gray points) of the “red” driver. The data were sampled on one day, October 22, 2012. To compute the convex hull, data-points which were hit less than 5 times have been omitted. This demonstrates that drivers occupy on average very similar regions in this  $(v, a)$ -space.

1 Note, that these parameters are in principle directly measurable, without getting them from a calibration  
 2 process. (Clearly, the result from such a direct estimation may differ from the results of a calibration.) The  
 3 acceleration bounds can be read off a time-series  $a(t)$  or from Figure 2, the parameters  $b$  and  $\omega$  from a plot  
 4 of acceleration versus speed difference or distance, respectively, and the preferred headway  $T$  from a fit of  $g$   
 5 versus speed  $v$ . In addition to that, they can also be estimated by fitting such a model to real data, e. g. from  
 6 car-following episodes.

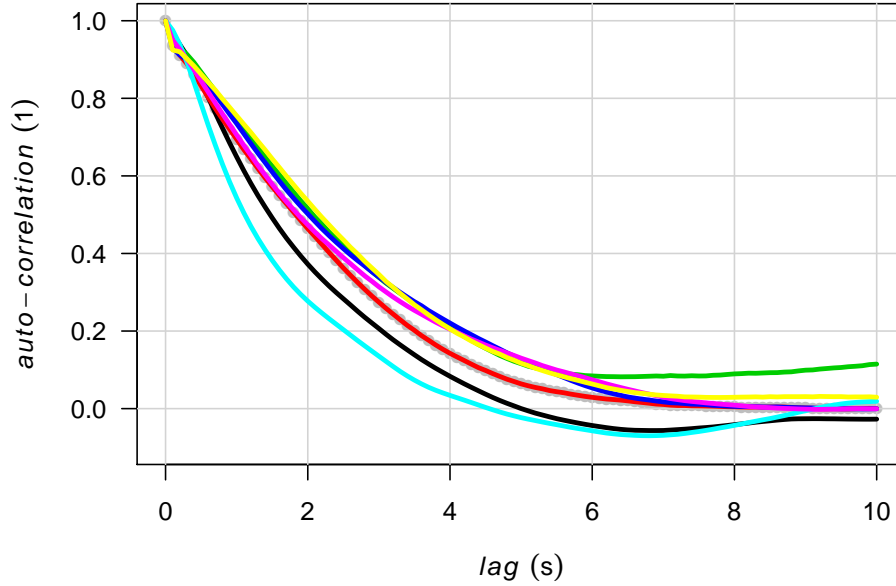
7 As mentioned already, it is assumed here that each driver has its own set of parameters, and these  
 8 parameters may change depending on external influences (weather, mood, level of stress etc.). However,  
 9 as long as the changes are slow compared with typical time-scales in the model, they can be considered as  
 10 constant and do not interfere with the dynamic. This is still a deterministic model, and it is formulated as a  
 11 differential equation. This means that the driver is applying her control in each instant of time.

12 Adding noise to such a model leads to a stochastic differential equation:

$$13 \quad \dot{v} = a = B(\omega^2(g - g^*(v, \Delta v))) + \sigma \xi. \quad (3)$$

14 The size of the noise  $\sigma$  is of course another parameter, while  $\xi$  stands for the noise term itself. It is important  
 15 to note that the noise term should be bounded (it cannot be normally distributed) and that it must have a  
 16 memory, which means that acceleration cannot change in 1 ms or so, but changes slowly, which is true for  
 17 vehicles with average masses above 1000kg. A noise term with such a memory is named colored noise.

18 This memory of the acceleration is an empirical feature. Real acceleration time-series have an auto-  
 19 correlation function that drops from 1 for time lag 0 to  $1/e$  for a time lag between two and four seconds.  
 20 This can be seen in Figure 3, where the same data as in Figure 2 have been used to compute the auto-



**FIGURE 3** Auto-correlation function of the acceleration of seven vehicles, again taken from the simTD data.

- 1 correlation function of the seven vehicles selected from the simTD data-base, also from October 22, 2012.  
 2 The auto-correlation function is computed by:

$$3 \quad c(\tau) = \frac{1}{(n-\tau)\sigma_a^2} \sum_0^{n-\tau} \hat{a}(t) \hat{a}(t+\tau),$$

4 where the variable  $\hat{a}(t)$  is the acceleration from which the mean value is subtracted, while  $\sigma_a$  is the standard  
 5 deviation. Again, this is computed using “gnu R” [18]. Using pure white noise gives an acceleration time  
 6 series with zero correlation time, which would not be valid under a physicist’s point of view.

7 A different model may be obtained by assuming that a driver does not react permanently, but only from  
 8 time to time. These so called action-point models [19] have discrete points in time where acceleration (more  
 9 precisely: driver’s control of it) changes quickly to a new value that might be based on equation (1). In this  
 10 case, the acceleration noise is added only at these action-points and by assuming that a driver is not very  
 11 good at setting the acceleration based on equation (1) but adds an error to it. In this case, the action-point  
 12 mechanism introduces the memory in the acceleration, since acceleration changes only little or not at all  
 13 between two subsequent action-points:

$$a_k = B(\omega^2(g(t) - g^*(v(t), \Delta v(t)))) + \sigma \xi(t) \quad t = t_k, \quad (4)$$

$$x(t) = x(t_k) + v(t_k)(t - t_k) + \frac{1}{2}a_k(t - t_k)^2 \quad t \in [t_k, t_{k+1}[. \quad (5)$$

14 Recently, a new class of models has been introduced that assume that the noise is not simply in the  
 15 acceleration (it may be there, in addition), but is in one or all of the parameters describing the driver [9].  
 16 The main culprit here is the preferred headway  $T$ , but other parameters might do as well. From empirical  
 17 data it is well-known that the headway distance in real traffic is a very volatile variable in a wide range of  
 18 numbers, typically covering a range between 0.5 and 2 times the mean value [21]. This is quite different  
 19 from the fluctuations in speed difference, which are on average around a few m/s compared to speeds of  
 20 20...40m/s.



1 So, these models assume that the parameter  $T$  changes due to a stochastic process. The consequences  
 2 of such a noise mechanism will be detailed in section 4.

### 3 **4 PARAMETERS ESTIMATION IN CASE OF A STOCHASTIC MODEL**

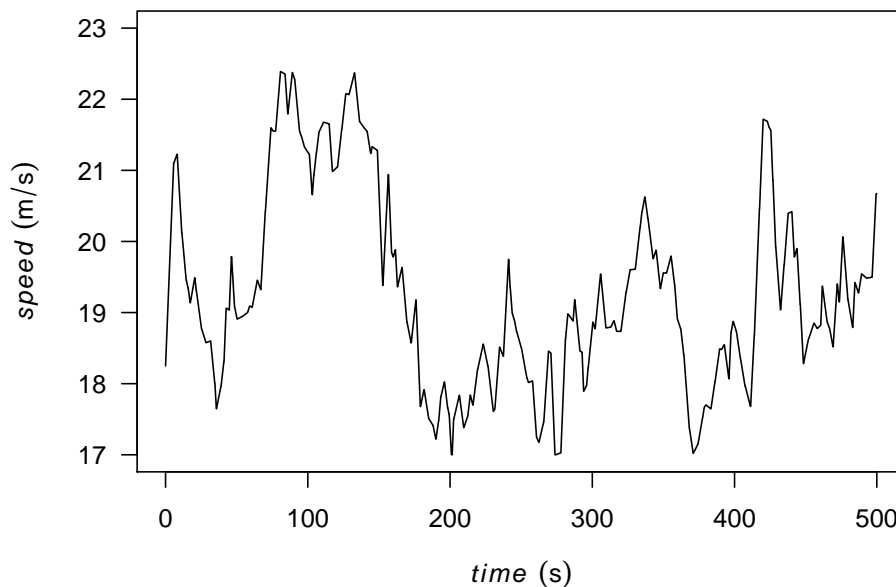
4 To fix ideas, the model in equation (1) is used without bounding the values of acceleration and speed,  
 5 i. e.  $B(\cdot)$  is the identity function (1) and is thus written:

$$6 \quad \dot{v} = \omega^2 \left( g - v \left( T + \frac{\Delta v}{b} \right) \right). \quad (6)$$

7 This reduces the number of parameters to three; the omitted three parameters are parameters that limit the  
 8 dynamics, while the remaining three parameters describe the interaction between the vehicles. Another  
 9 advantage of this reduction is that the model can be rewritten as a model that is linear in the three parameters  
 10  $p_i$  (and weakly non-linear in the dynamics itself):

$$11 \quad \dot{v} = p_1 g + p_2 v + p_3 v \Delta v. \quad (7)$$

12 A simple scenario is presented in the following. A vehicle that does not follow this model but drives  
 13 its own stochastic course will be used as a leading vehicle to this model. Note, that time is discretized in  
 14 chunks of 0.1 s, which is a common empirical resolution. The lead vehicle's trajectory is created by drawing  
 15 acceleration values  $A$  from a Laplace distribution  $p(A) \propto \exp(-|A|/a_0)$ , which will be changed at random  
 16 points in time whose distance is drawn randomly from the interval  $[0.5, 2]$  s. This yields an acceleration  
 17 trajectory  $A(t)$  of the lead vehicle, from which the speed  $V(t)$  of the lead vehicle is computed. In addition,  
 18 it is made sure that the resulting  $V(t)$  remains within two bounds  $[V_1, V_2]$ , an example is displayed in Figure  
 19 4. Such a trajectory looks similar to real-world speed functions, and it is important that it is not a simple  
 20 function. In comparison with the synthetic data previously used, this synthetic trajectory is much more  
 21 realistic (see Fig. 1 of [15]). If it were simple and mostly constant, the following vehicle's behavior in the



**FIGURE 4** Speed versus time of a small piece of the noisy synthetic trajectory of the lead vehicle. The trajectory is bounded between 17 and 23 m/s.

1 case of a deterministic model would be very simple, too, and markedly different from all empirical data we  
 2 have analyzed so far, which always appear very noisy. Note, however, that this noisy look of this curve is  
 3 only due to the random jumps at the action-points, in between the trajectory follows a simple linear function  
 4  $V(t) = V_i + a_i(t - t_i) \quad t \in [t_i, t_{i+1}[$ .

5 The subject vehicle follows this trajectory with a certain set of parameters, and this subject vehicle is  
 6 described by the model equation (7) and endowed with different noise mechanisms. In the following, the  
 7 triple  $p = (0.0169, -0.0239, 0.0172)$  will be used for equation (7), which has been obtained from a real-  
 8 world data-set by the calibration method specified below. Note, that this set corresponds to  $T = 1.41$  s  
 9 and  $b = 0.98$  m/s<sup>2</sup>, which seems a realistic choice of parameters. Running the model equation (7) with a  
 10 trajectory generated for the lead vehicle gives a certain simulated trajectory  $(g(t), v(t), a(t))$  of the following  
 11 vehicle.

12 By fitting this trajectory to the model equation (7) as a simple (robust) linear fit [18], the parameters can  
 13 be reproduced with a small error, which can be found in the column labeled “rms(*acc, fit*)” in Table 1. This  
 14 remains also true if the fitted parameters are used to run the model once more and then compare the accel-  
 15 eration gotten by this with the acceleration generated during the first run of the model. The corresponding  
 16 values can be found in the column of Table 1 labeled by “rms(*acc, sim*)”. All these results are in the second  
 17 row of Table 1, which is labeled “raw”.

18 So, when the model equation (7) is following a noisy lead vehicle with speed  $V(t)$ , it is possible to find  
 19 from the time-series the parameters that have gone into the model. This was the case for a deterministic  
 20 model, which was driven by a stochastic lead vehicle. It becomes way more difficult when the stochastic  
 21 variants of the model are being used. In total, the four different stochastic models have been used:

22 **Model-1:** Adding a white-noise term  $\sigma\xi$  to equation (7), which could be named the physicist’s approach  
 23 since it is lend from the modeling of the Brownian motion.

24 **Model-1a:** Just like model 1, but instead of white noise colored noise has been added to the acceleration  
 25 time-series [5, 8]. Colored noise can be understood as an exponentially smoothed white noise  
 26 process, the simplest approach that has been used here is  $n(t + \Delta t) = \alpha n(t) + (1 - \alpha)\sigma\xi(t)$ .

27 **Model-2:** An action-point type algorithm in the acceleration, which is given by equations (4) and (5).

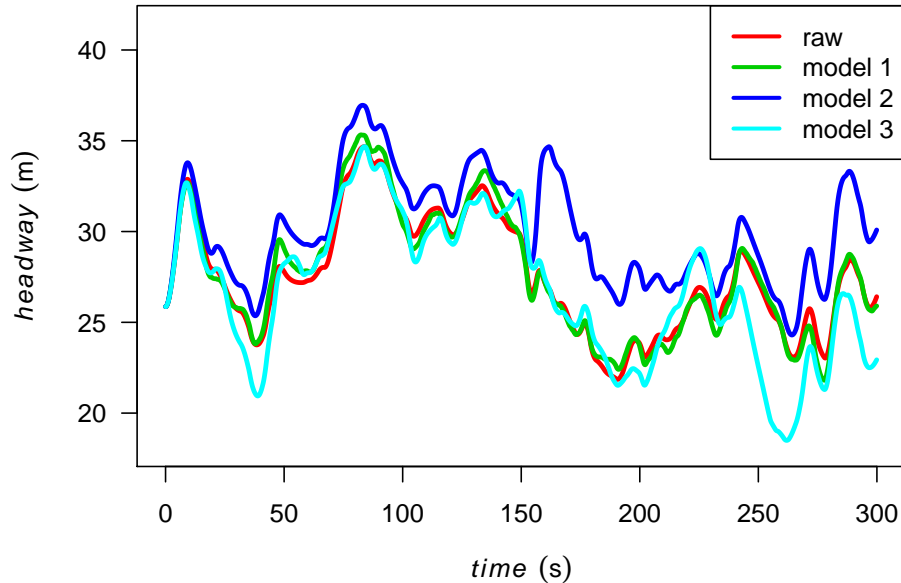
28 **Model-3:** In the so called 2D models [9, 21], a parameter could be changed randomly instead of fiddling  
 29 around with the dynamics. The method here uses a mechanism that changes the parameter  $p_2$   
 from time to time to a value that is drawn from a symmetric interval around the true value.

**TABLE 1 Results of the parameter estimation for the four models.**

	$p_1$	$p_2$	$p_3$	rms( <i>acc, fit</i> )	rms( <i>acc, sim</i> )
input	0.0169	-0.0239	0.0172		
raw	0.0183	-0.0259	0.0171	0.0064	0.0009
model 1	0.0170	-0.0241	0.0142	0.2981	0.4131
model 1a	0.0120	-0.0174	0.0005	0.5101	0.6079
model 2	0.0196	-0.0291	0.0158	0.1260	0.1244
model 3	0.0104	-0.0147	0.0109	0.1334	0.1416

30

31 One realization of the three models (model 1, model 2, and model 3) is displayed in Figure 5. Apart from the  
 32 raw model, which has no stochastic component and is almost coincident with the input data, the different



**FIGURE 5** Plot of the headway versus time for the three different models, for the same lead vehicle speed function. The red curve is the original (simulated) gap that the models 1 to 3 ought to reproduce.

1 models lead to distance-versus-time curves. Albeit these curves do not look too different, the parameter  
 2 estimation of all these stochastic models fails. The parameters are way off, and the r.m.s. error for the  
 3 acceleration becomes considerable (two orders of magnitude bigger than the one obtained for the raw data!).  
 4 Obviously, it depends on the parameters chosen for the size of the noise, so no general statement about its  
 5 size can be made. All the results can be found Table 1.

6 To see that this is not just the effect from a single simulation, 100 realizations of the process have been  
 7 created with the same fixed parameter set for model 3. The fitting of this simulation data to the model  
 8 equation then yields a different parameter set, whose distribution is displayed in Figure 6 along with the  
 9 input parameters (vertical arrows). The distribution is robust against changes in the lead vehicle's speed,  
 10 i. e. different realizations of  $V(t)$  give the same distribution of fitted parameters.

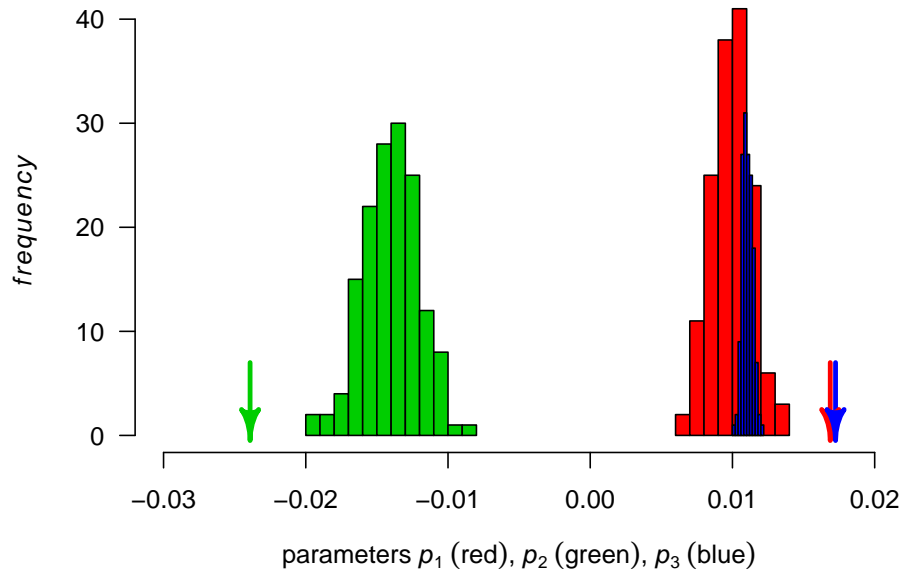
## 11 5 CONCLUSIONS

12 The literature is vast about testing calibration procedures and using simulation results to face noise in data  
 13 or behaviors variability among drivers. In both cases, the last decade produced a common agreement on the  
 14 solutions:

- 15 • To cope with noisy data, the simulation result is to be taken accompanied with an error estimation.
- 16 • To cope with the distribution of parameters, the simulations are repeated with random draws of the  
 17 parameter set, which result in a calibrated distribution.

18 On the contrary, literature is scarce about calibration procedure when the stochasticity is embedded in the  
 19 model itself.

20 This work demonstrates that most stochastic processes are potent enough to let a parameter estimation  
 21 process go awry. It is not only that the goodness of fit becomes worse, in addition, even the parameters do  
 22 not come out correctly. In order to minimize numerical problems, the model and the fitting procedure have



**FIGURE 6** Distribution of the parameters (for model 3) estimated from 100 different runs of creating a following trajectory with the same set of parameters and then fitting them to the model. The input parameters are indicated by the three arrows pointing to the x-axis.

1 been simplified strongly so that a linear fit was sufficient, with the full statistical power that such a method  
 2 provides (there are no false minima to be approached by this method, a linear fit also yields the optimum  
 3 solution). In addition, all the statistical quality measures like  $t$ -values of the parameters and their respective  
 4 significance levels displayed strong values indicating a really good fit nevertheless.

5 The important point here is that the bad fits do not manifest themselves. So, the researcher would be  
 6 convinced that the parameter estimation has led to a good result. However, all of the noise models used here  
 7 had the power to let the estimated parameter values come out wrongly. So far, we do not have a remedy for  
 8 this. The statistics, as well as the r.m.s. and even the visual inspection of the results look quite good, but  
 9 from the numerical experiments it could be seen that when fitting this type of models, the parameters cannot  
 10 be estimated correctly.

## 11 ACKNOWLEDGMENTS

12 The funding for this project has been provided by DLR's project I.MoVe. Our special thanks go to the team  
 13 of the TU Munich (Sebastian Gabloner, Martin Margreiter), who provided the simTD data. simTD itself  
 14 was funded by the Federal Ministry of Economics and Technology, the Ministry of Education and Research  
 15 as well as the Ministry of Transport, Building and Urban Development. To fit the model's parameters, yet  
 16 another data-set provided by T. Nakatsuji by have been used, which is gratefully acknowledged.

**REFERENCES**

- 2 [1] Safe and intelligent mobility – test field germany, 2012. URL:  
3 <http://www.simtd.de/index.dhtml/enEN/index.html>, last access 7.7.2015.
- 4 [2] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. Dynamical model of traffic con-  
5 gestion and numerical simulation. *Physical Review E*, 51:1035 – 1042, 2 1995.
- 6 [3] E. Brockfeld, R. D. Kühne, and P. Wagner. Calibration and validation of microscopic traffic flow  
7 models. *Transportation Research Records*, 1876:62 – 70, 2004.
- 8 [4] Winnie Daamen, Christine Buisson, and Serge P. Hoogendoorn, editors. *Traffic Simulation and Data*  
9 – *Validation Methods and Applications*. Taylor and Francis, CRC Press, 2014.
- 10 [5] Ronald F. Fox, Ian R. Gatland, Rajarshi Roy, and Gautam Vemuri. Fast, accurate algorithm for numer-  
11 ical simulation of exponentially correlated colored noise. *Phys. Rev. A*, 38:5938 – 5940, Dec 1988.
- 12 [6] P.G. Gipps. A behavioural car-following model for computer simulation. *Transportation Research*  
13 *Part B*, 15(2):105 – 111, 1981.
- 14 [7] W. Helly. Simulation of bottlenecks in single lane traffic flow. In *Proceedings of the symposium on*  
15 *theory of traffic flow.*, page 207 – 238, 1959.
- 16 [8] J. Honerkamp. *Stochastic Dynamical Systems*. VCH Weinheim, 1994.
- 17 [9] Rui Jiang, Mao-Bin Hu, HM Zhang, Zi-You Gao, Bin Jia, Qing-Song Wu, Bing Wang, and Ming Yang.  
18 Traffic experiment reveals the nature of car-following. *PloS one*, 9(4):e94351, 2014.
- 19 [10] Victor L. Knoop and Christine Buisson. Calibration and validation of probabilistic discretionary lane-  
20 change models. *IEEE Transactions on Intelligent Transportation Systems*, 16(2):834 – 843, 2015.
- 21 [11] Stefan Krauss, Peter Wagner, and Christian Gawron. Metastable states in a microscopic model of  
22 traffic flow. *Physical Review E*, 55(5):5597 – 5602, 1997.
- 23 [12] Kai Nagel and Michael Schreckenberg. A cellular automaton model for freeway traffic. *Journal de*  
24 *Physique I France*, 2:2221 – 2229, 1992. According to the web page of J Phys. I FRance, this paper  
25 has so far (12-9-2012) 829 cross references.
- 26 [13] G.F. Newell. A simplified car-following theory: a lower order model. *Transportation Research Part*  
27 *B*, 36(3):195 – 205, 2002.
- 28 [14] S. Ossen. *Longitudinal Driving Behavior: Theory and Empirics*. PhD thesis, TU Delft, 2008.
- 29 [15] S. Ossen and S. P. Hoogendoorn. Validity of trajectory-based calibration approach of car-following  
30 models in the presence of measurement errors. *Transportation Research Records*, 2088:117 – 125,  
31 2008.
- 32 [16] Biagio Ciuffo; Apollonia Miola; Vincenzo Punzo; and Serenella Sala. *Dealing with uncertainty in*  
33 *sustainability assessment*. Publications Office of the European Union, Luxembourg, 2012.
- 34 [17] Vincenzo Punzo, Biagio Ciuffo, and Marcello Montanino. Can results of car-following model calibra-  
35 tion based on trajectory data be trusted? *Transportation Research Record: Journal of the Transporta-*  
36 *tion Research Board*, (2315):11 – 24, 2012.

- 1 [18] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation  
2 for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.
- 3 [19] E. P. Todosiev and L. C. Barbosa. A proposed model for the driver-vehicle-system. *Traffic Engineering*,  
4 34:17 – 20, 1963/64.
- 5 [20] Martin Treiber, Ansgar Hennecke, and Dirk Helbing. Congested traffic states in empirical observations  
6 and microscopic simulations. *Physical Review E*, 62(2):1805, 2000.
- 7 [21] Peter Wagner. Analyzing fluctuations in car-following. *Transportation Research Part B*, 46:1384 –  
8 1392, October 2012.