THE STRUCTURE OF THE PARAMETER SPACE OF CAR-FOLLOWING 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/

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
95th Annual Meeting
Transportation Research Board
January 12 – 16, 2016
Washington, D. C.

Submission date: July 30, 2015

words:	4 252
plus 9 figures:	2 250
total count:	6 502
word limit:	7 500

*corresponding author

ABSTRACT

Typically, the calibration of microscopic traffic flow models seeks to minimize the speed or the distance error, but not both of them simultaneously. By analyzing the car-following part of a couple of microscopic traffic flow models with data from car-following situations, it is demonstrated that the two objectives depend on each other in a complicated way: Minimizing the speed error does not lead to a minimum distance error and vice versa. This means the existence of a Pareto front in the speed-error-distance-error space. The Pareto front in particular, and the parameter space in general, can be probed in a systematic manner by exploring the parameter space of a given model with quasi-random sequences. Furthermore, this systematic scanning of the parameter space allows for a ranking of the importance of the parameters of the model, albeit in the context of the data given. It is demonstrated that only a sub-set of the parameters of a given model are actually needed. So far, this is only theoretical work, since it relies strongly on a fast implementation of the microscopic models. However, with the advance of faster and more parallel computers such an approach can even be implemented with ease in realistic settings and with any traffic simulation program that allows for an external setting of its parameters.

Keywords: Car-following Models, Micro Simulation, Calibration, Parameter Space

1 INTRODUCTION

In the recent years a large amount of work has been performed on the calibration and validation of traffic flow models, see [3, 4, 5, 8, 21, 30] to name but a few. However, finding the right set of parameters for precisely simulating traffic described by empirical data can still be a challenging endeavor. Depending on the specific situation not all model parameters might contribute to the results with the same degree. In such cases, identifying less "important" parameters and setting them to constant values may simplify the process of computing an optimal solution for the remaining more "important" parameters. For further investigation, this paper analysis the effect, i. e. the importance for the model's output, of the model parameters for a number of car-following models in a systematic way. Therefore, the most basic scenario of one vehicle following behind another one is chosen. Note, that other traffic situations can be treated in a similar manner.

In case of car-following, calibration tries to minimize the speed-error e_v (with v_t the measured speed at time-step t, T_n the total number of time-steps with data available and $\hat{v}_t$ the resulting speed given by the model; in general, "hatted" variables denote simulated variables in the following):

$$e_v = \frac{1}{T_n} \sum_{t=1}^{T_n} |v_t - \hat{v}_t|, \tag{1}$$

or the distance-error e_g (with g_t the measured net distance or gap to the vehicle in front):

$$e_g = \frac{1}{T_n} \sum_{t=1}^{T_n} |g_t - \hat{g}_t|. \tag{2}$$

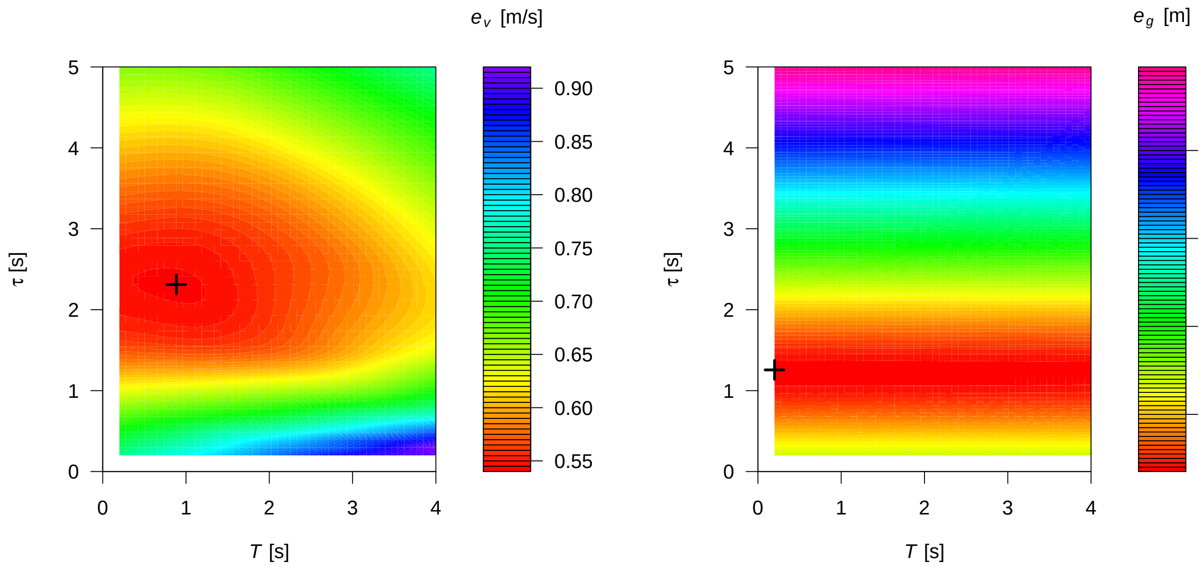
Of course, other error measures could have been used instead of the mean absolute deviation (MAD) defined here. Surprisingly, in most cases minimizing e_v and e_g together is not possible: The set of parameters that minimizes e_v is not the same set that minimizes e_g. This is demonstrated explicitly by the following example. As a simple model, a linear optimal velocity model is used, which is defined as follows:

$$\frac{d}{dt}\hat{v}_t = \frac{1}{T}\left(\frac{\hat{g}_t}{\tau} - \hat{v}_t\right). \tag{3}$$

This model is similar to the Newell model [24] and has a relation to the first CA-model [23]. Here, the model is used because it has just two parameters: the relaxation time T and preferred headway τ. (Note, that this model has another parameter, the vehicle length; however, since the data used in this paper contain only the net distance to the vehicle in front, this parameter is neglected for all the models under consideration.) This allows a complete visualization of the results, even in parameter space. In principle, the parameters can be sampled from a positive range; here both of them will be distributed in the range $[0,4]$ s. In this case, both error functions under consideration (eqn. (1), (2)) can be displayed as a function of the parameters as shown in Figure 1. It can be seen that the minima in speed and gap differ from each other: the minimum of the speed error is at $(T = 0.86,\ \tau = 2.29)$ s, while the gap-minimum is at $(T = 0.05,\ \tau = 1.23)$ s. Let us mention in passing that for the root mean square error both minima have again different values.

Note that the gap-error in Figure 1 (b) does depend on τ, but only very weakly on T. (However, as the gap error equals the integrated speed error, a small dependence on T is still given.) Therefore, T is an example for an "unimportant" parameter with respect to the analyzed car-following situation. Later, a more precise definition for the importance of a parameter will be derived.

In general, free-flow parameters (e. g. desired speed v_0, desired acceleration a etc.) are "unimportant" in pure car-following situations. Conversely, parameters assigned to car-following (e. g. desired time gap T_D or minimum headway s_0) are irrelevant for solely free-flow situations. Parameters describing approaches to



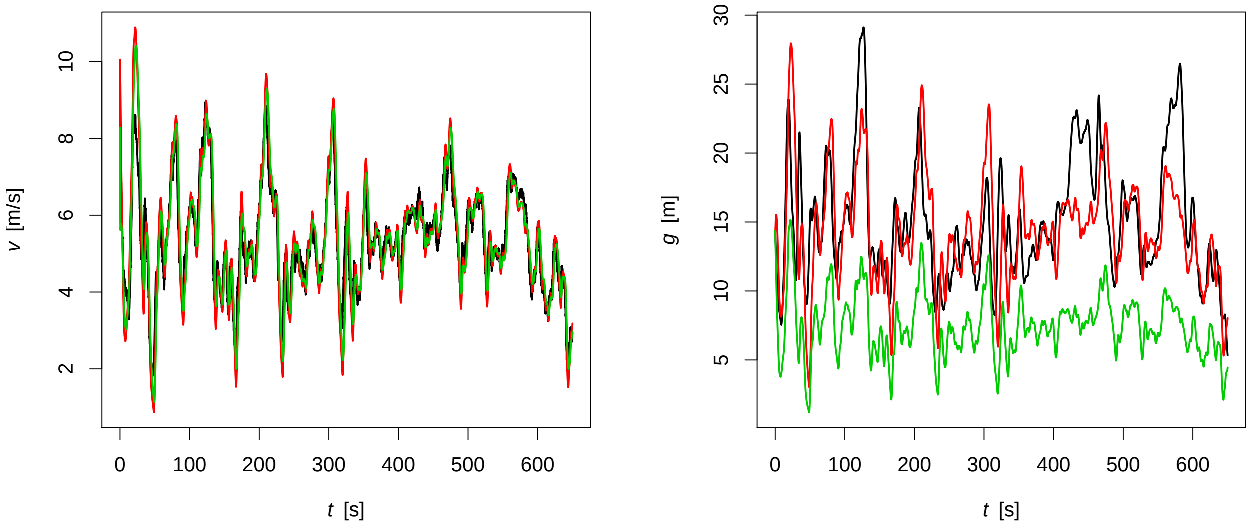
(a) Speed error e_v **(b) Gap error e_g**

FIGURE 1 Visualization of speed and gap error. The error is the MAD as defined in eqn. (1), (2) for the model eq. (3) as a function of the two parameters T and τ. The position of the minimum is denoted by a black cross.

slower cars, standing obstacles (e. g. desired deceleration b in Gipps' model or IDM), or other non-stationary situations (relaxation time T) are irrelevant in both free-flow and car-following situations if no large speed differences are contained in the data set.

Figure 2 shows a comparison between the measured time series (black) and the simulation (red / green) with different optimal parameter sets regarding speed and gap error. In Figure 2 (a) the red curve represents simulated speeds $\hat{v}_t$ with the parameter set at the minimum of the speed error e_v. Consequently, for Figure 2 (b) the red curve corresponds to the simulated gaps $\hat{g}_t$ with the parameter set at the minimum of the distance error e_g. The green curve refers to the simulated speed $\hat{v}_t$ with the parameter set at the minimum gap-error e_g in Figure 2 (a). In conclusion, the green curve in Figure 2 (b) shows the simulated gaps $\hat{g}_t$ with the parameter set at the minimum speed-error e_v. It could be seen that this model is quite good at simulating the speed (speed-error $e_v = 0.55\,\mathrm{m/s}$), but not very well at simulating the distance between the two vehicles (gap-error $e_g = 7.07\,\mathrm{m}$).

Therefore, a more in-depth analysis is needed. This is done by scanning the whole reasonable parameter space by generating quasi-random sequences described below. The approach pursued here shares some similarities with sensitivity analysis [28, 7, 27, 13, 32], but is more general. In fact, this approach reveals the existence of a so-called Pareto front in the (e_v, e_g)-space, which is typical for multi-criteria optimization problems [29]. A Pareto front describes a line in the parameter space where one parameter cannot be improved further without degrading another parameter. Note also in passing that minimizing e_g is much more difficult than the minimization of e_v, i. e. the distance error is larger than the speed error. This can be seen if the respective error-terms are compared to the mean values of speed and distance: They are of comparable size since they are usually connected by $g = v\tau$, where τ is the time headway, which is for car-following situations in the range of $1\ldots2$ seconds. The reason for this fact is not so easy to uncover: The distance is essentially the time integral of the speed difference between the two vehicles, therefore an existing speed-error might be amplified by this integration. Another reason might be the idea formulated in



(a) Measured and simulated speed **(b) Measured and simulated headway**

FIGURE 2 Comparison between data and simulation. Left: simulated speed $\hat{v}_t$ (red) vs. measured speed v_t (black). Right: simulated distance $\hat{g}_t$ (red) vs. measured distance g_t (black). The green curves denote the speed for the minimum gap-error e_g (left) and the minimum speed-error e_v (right).

[34] and analyzed in [14]: That the preferred time headway of a driver fluctuates permanently, making it therefore difficult for any model to describe the distance behavior in a consistent manner, which does not include these fluctuations.

The present paper is organized as follows: Section 2 presents the underlying data for this analysis as well as the car-following models used. Moreover, in the last sub-section 2.4 the idea of the systematic scanning of the parameter space is introduced. Section 3 assays the results of this work. Finally, section 4 concludes this paper.

2 METHODS

2.1 Data Set

For this analysis we used empirical data collected by a high-precision GPS-measurement in a real-world experiment in Hefei, China on January 19, 2013. This experiment featured 25 vehicles in several car-following situations with various speed and acceleration patterns to simulate different levels of congestion. At some points the resulting data set presents shock waves and reveals significant arbitrary fluctuations in the headways g while the speed difference Δv between following vehicles are constantly small.

Because this work focuses on the car-following process, the absence of any external influences or disturbances is important to ensure the collected data contain pure car-following episodes. For this paper, 24 different drivers following a leading vehicle in 11 different episodes with a length of about 1 hour were analyzed. Therefore, only speed v_t, distance to the vehicle in front g_t, speed difference to the vehicle in front Δv_t, and acceleration a_t have been used. For more details about the data set please refer to [14].

2.2 Generating Simulation Output

The simulated data are computed as follows: For each of the car-following episodes and each of the models described by an acceleration function $A(v_t, g_t, \Delta v_t)$ the ordinary differential equation (ODE)

$$\frac{d}{dt}\hat{v}_t = A(\hat{v}_t, \hat{g}_t, V_t - \hat{v}_t), \tag{4}$$

$$\frac{d}{dt}\hat{g}_t = V_t - \hat{v}_t \tag{5}$$

is solved with the input of the speed of the lead vehicle V_t (a driven ODE). There are other methods that generate such a sequence (see [33] for a discussion), but only by solving the ODE above a faithful representation of the car-following dynamics is obtained that can be compared with the measured variables (g_t, v_t). The ODE above is being solved by a second order update scheme with the time-step size set to $\Delta t = 0.1\,\mathrm{s}$. This value of Δt is also applied in the following equations:

$$\hat{v}_{t+1} = \hat{v}_t + \Delta t\, A(\hat{v}_t, \hat{g}_t, V_t - \hat{v}_t), \tag{6}$$

$$\hat{g}_{t+1} = \hat{g}_t + \frac{\Delta t}{2}\,(V_{t+1} + V_t - \hat{v}_{t+1} - \hat{v}_t). \tag{7}$$

During the simulation, all the models are forced to stay within speed bounds $v \in [0, v_{\max}]$ and they are forced to stay crash-free, i. e. $g \geq 0$ is ensured. This is important, since the method below can produce parameter values that are completely nonphysical, and consequently there is no guarantee that a given model behaves correctly.

2.3 Car-following Models Used

The following models have been used in the research presented here. Due to a lack of space only the simple ones are displayed with an equation.

CA The cellular automaton model after [23]. Here, a slightly generalized version has been used that allows for an adaptation of the cell-size of the cellular automaton and uses it as an adoptable parameter. This is needed, since the original model has been designed for a time-step-size of 1 second. The modified version reads:

$$\hat{v}_{t+1} = \min\left\{\hat{v}_t + 1,\ \frac{\hat{g}_t}{\tau},\ v_{\max}\right\}. \tag{8}$$

All variables are divided by the cell-size λ and then rounded toward the nearest integer. In addition, with a probability p one speed-unit is subtracted from the speed in eq. (8) to make this model stochastic. Interestingly, the average value of λ for the data-set below turns out to be around $0.25\,\mathrm{m}$.

Fritzsche This is psycho-physical model introduced in [9]. It is similar (but slightly simpler) to the Wiedemann model used in VISSIM.

Gipps This is the model as defined in [10]. It reads:

$$\hat{v}_{t+1} = \min\left\{\hat{v}_t + 2.5\,a\,\tau\,\Delta t\,\left(1 - \frac{\hat{v}}{v_{\max}}\right)\sqrt{0.025 + \frac{\hat{v}}{v_{\max}}},\right.$$
$$\left. -b\left(\theta + \frac{\tau}{2}\right) + \sqrt{\left(b\left(\theta + \frac{\tau}{2}\right)\right)^2 + 2b\,(\hat{g} - s_0) - \hat{v}\,\tau + \frac{V^2}{\hat{b}}}\right\}. \tag{9}$$

In this equation b is the (most severe) braking deceleration whereas s_0 refers to the average effective size of an vehicle, i. e. the physical length plus the average distance between vehicles at standstill ($v = 0\,\mathrm{km/h}$).

GLM The general linear model introduced in [35]. It is defined directly in the discretized version:

$$\hat{v}_{t+1} = \alpha\,\hat{v}_t + \beta\,\hat{g}_t + \gamma V_t + \delta. \tag{10}$$

IDM The intelligent driver model [31]. Implemented here is the simplest variant:

$$\frac{d}{dt}\hat{v} = a\left(1 - \left(\frac{\hat{v}}{v_{\max}}\right)^4 - \left(\frac{g_0 + \hat{v}T - \frac{\hat{v}(V-\hat{v})}{2\sqrt{ab}}}{\hat{g}}\right)^2\right). \tag{11}$$

Here, the parameter g_0 refers to the average distance between vehicles at standstill ($v = 0\,\mathrm{km/h}$).

KKW The Kerner Klenov Wolf CA-model which simulates three-phase traffic [16].

MITSIM This is the model introduced in [1] and used in the MITSIM simulator. Here, we have used the code in the open source variant of MITSIM; however, during the conversion in our own source code, some modifications have been made so that it cannot completely ruled out that the actually implemented model is slightly different from the real MITSIM code.

LinOVM An optimal velocity model with a linear optimal velocity function is already given in eq. (3). This model is close to the model invented in [24] and it is the continuous analogue [20] of the CA model [23].

OVM3 / OVM4 This is the optimum velocity model, originally described in [2] but in the variant defined in [22]:

$$\frac{d}{dt}\hat{v} = \frac{1}{T}\left(v_{\max}\frac{\hat{g}^2}{g_0^2 + \hat{g}^2} - \hat{v}\right). \tag{12}$$

This model comes in two variants: The one of eq. (12) and one with an additional parameter: By using $\hat{g} \to \hat{g} + T_a(V - \hat{v})$ the model is extended by an anticipation term and depends in addition on Δv.

SUMO This model has been introduced in [19], it bears a strong similarity with the Gipps model but is a stochastic model. Here, it is implemented in the variant used in the open source simulation SUMO [18]:

$$\hat{v}_{t+1} = \min\left\{\hat{v}_t + a\,\Delta t,\, \text{-}b\,\tau + \sqrt{(b\,\tau)^2 + V_t^2 + 2\,b\,\hat{g}_t},\, v_{\max}\right\}. \tag{13}$$

2.4 Model Analysis And Ranking Of The Parameters

In the example in the introduction the parameter space of the Newell model had been scanned more or less completely. In principle, it is possible to fix beforehand a number of parameter sets and then generate randomly parameter vectors in the p-dimensional parameter space (where p is the number of parameters of a given model). However, doing this randomly leaves too large holes in the p-dimensional space. Thus, so called quasi-random sequences are used to minimize these holes. A couple of different approaches are available for the computation of quasi-random sequences. In this paper, Halton sequences [11, 12, 17] are utilized. This work uses the code as it has been published in [6], with small modifications.

After generating e. g. 10^4 quasi-random parameter sets together with the associated (e_v, e_g) values for a certain car-following episode, the parameter values as function of rank can be analyzed further. If the minimum value obtained in this manner is close to a true minimum of the objective in the parameter space, then also the neighboring values should be close: the parameter values themselves as well as the values of the objective function. So, by looking at the first 10 ranks, a "good" parameter has a small variation, while a "bad" parameter has a large variation. This has been checked for by introducing a dummy parameter in a model that is known to have no influence on the outcome whatsoever: It displays the maximum possible variation. To catch this behavior in a certain number, for each parameter in the ranked representation first the so called contrast:

$$c_i = \frac{p_i^{(\max)} - p_i^{(\min)}}{|p_i^{(\max)}| + |p_i^{(\min)}|} \tag{14}$$

and from this, the so called importance $w_i = 1 - c_i$ can be computed. The superscripts denote the minimum and maximum of the parameter values of the first ten ranks of the parameter i. The contrast of the dummy parameter turns out to be 1, leading to an importance of 0, while normal parameters display importance values in the range between $[0,1]$. In Figure 3, this is exemplarily displayed for the OVM3 model with its three parameters T, $v_{\max}$ and g_0. Here, it seems in most of the car-following episodes that the relaxation time T of the model is the least important parameter, while the other two ($v_{\max}$ and g_0) are "good" parameters with respect to their contribution to the model output in the specific traffic situation.

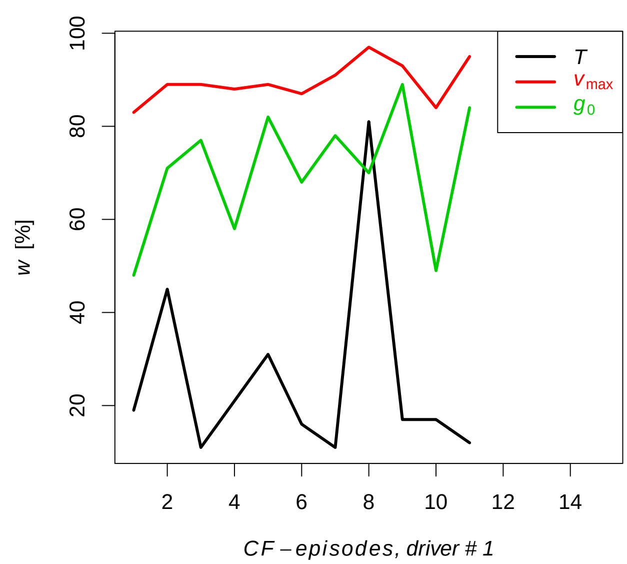


FIGURE 3 Importance of the three parameters of the OVM3 model as function of the car-following episode for driver #1.

3 RESULTS

All 10 models described in section 2.3 have been run with 10^4 quasi-random parameter sets. On average, 10 million updates/s on a three years old 3 GHz processor have been achieved for the different models, which is much faster than any commercial or open source general purpose microscopic simulation program. Interestingly, even in this code there is a certain amount of overhead involved (e. g. parameters will be handed over to the acceleration function in any time-step etc.), which can be seen best by the fact that the difference in computation time needed between the simplest and the most complex model is just a factor of three – the MITSIM model has the longest execution time, while the GLM was the fastest to compute.

3.1 Convex Hulls And Pareto Fronts

The most interesting result of this analysis is that in most cases a simultaneous minimization of speed and distance error is not possible. Obviously, there is a correlation between the two errors, indicated by the fact that all the error values roughly form an ellipse. In addition, the correlation coefficients between the speed-error and the gap-error range between 0.1 and 0.7. A typical example is displayed in Figure 4 for the GLM. No systematic variations between drivers have been found.

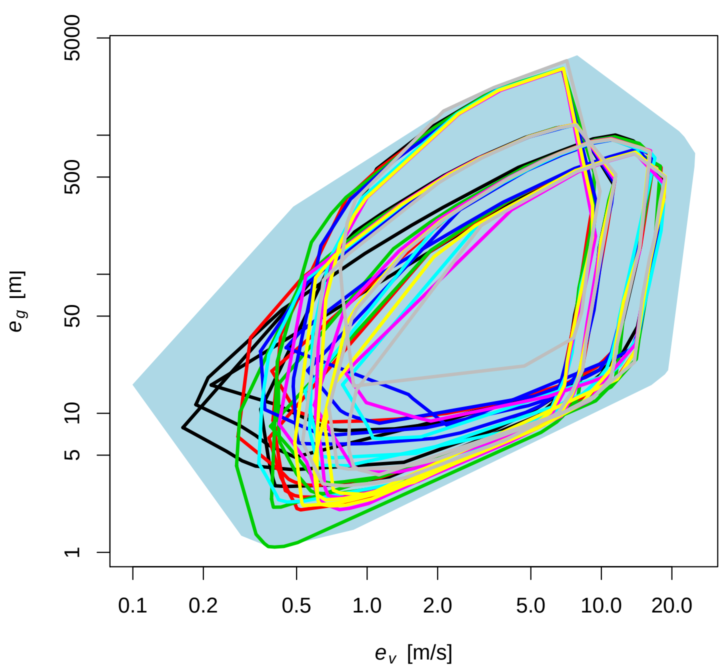


FIGURE 4 Convex hull in the error-space for the GLM model. The polygon in the background is the set of all the error values of all the car-following episodes analyzed, overlaid are a few examples of the convex hulls of individual episodes and drivers.

Interestingly, the homogeneous structure of the parameter space that has been sampled with quasi-random sequences transform into a well-defined structure of the space of objectives. This can be seen exemplarily by the bag-plot of the KKW model in Figure 5: Some of the error-values appear much more often than others, obviously there is a certain clustering visible in this space. One possible interpretation of this fact is that microscopic traffic flow models do not only rely on a good adaptation of their parameters. In addition to that it seems that all tested car-following models capture the underlying true mechanism that shape driving to a certain extent. Even a set of parameters that is not very well adapted to the reality at hand yields already an at least a good approximation of the speed-curve.

Figure 6 super-imposes all the Pareto surfaces of all the models together. There are apparent differences, e. g. some models (CA and OVM3) cover relatively small areas in the error space. Others models have quite large areas such as the Fritzsche, Gipps, and GLM model. In a certain sense, models with smaller areas are

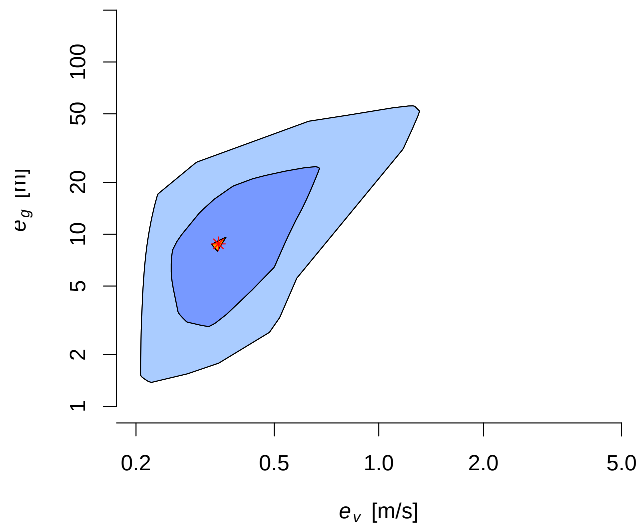


FIGURE 5 Bag-plot of all the error values (e_v, e_g) for all of the quasi-random points of the KKW model for the third driver in the platoon. The red point in the center is the bivariate median with a value of $(e_v = 0.34\,\mathrm{m/s},\ e_g = 8.85\,\mathrm{m})$, while the darker-blue "bag" is the 50-percentile area of the data, i. e. 50 % of the points lie inside the bag. The light blue area finally is again the convex hull of all the data.

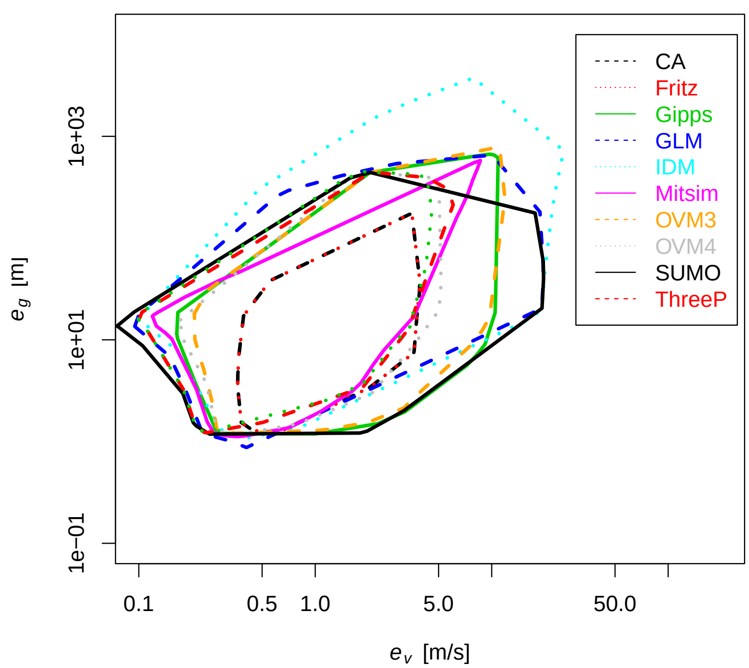


FIGURE 6 The convex hull for all the models and all car-following episodes. Note the logarithmic scaling of the axes.

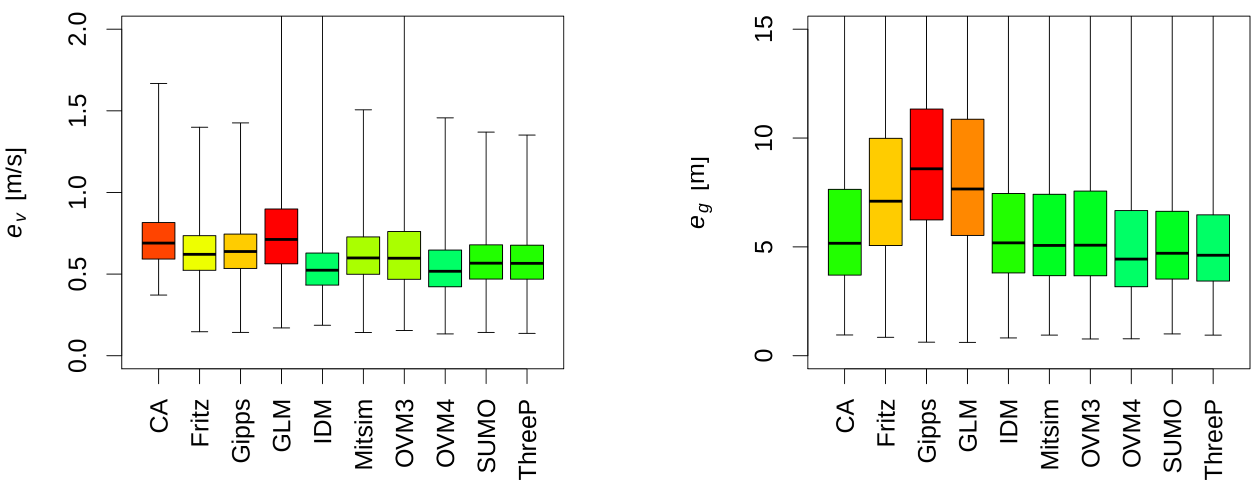
better than the ones with the larger areas, since in this case it is more likely to find sufficiently good solutions for a real situation even with a moderate amount of effort.

However, this has to be taken with a grain of salt. For some parameters, such as e. g. maximum acceleration, good constraints have been known and used by the authors. Especially the many parameters of the Fritzsche and the MITSIM model do not have such good constraints for many of their parameters, or at least they have not been known to the authors. In addition, some parameters may have areas where they lead to a truly bad model outcome. This has not been entirely checked for all models. Clearly, this includes some subjectivity into the area covered in the error-space.

Finally, it could be stated that there is a large amount of similarity between the different models, especially concerning the most interesting lower left corner, where the "good" values reside. Furthermore, one may also notice the factor of 10 between the speed-error and the gap-error.

3.2 Error Distributions

By looking at the 5 % best values obtained in this manner, a comparison between the different models can be performed. Since this is a large amount of error values, a statistical approach is used here: For each model and each objective, the quantiles are computed and compared with each other in a box-whisker plot. In accordance with especially the results in [4], there are only small differences between the various car-following models, with the exception of the Newell model in the version implemented here. Figure 7 presents these results graphically.



(a) Speed error e_v **(b) Gap error e_g**

FIGURE 7 Comparison of the ten models presented here in terms of the speed-error (left) and the gap-error (right). The colors are chosen in accordance with the median of the respective error with green indicating a small error and red indicating a large error. The maximum error values have been cut-off from the plots because of their large magnitude that makes the boxes appear too small.

While the speed-errors are usually already fairly small, the gap-errors are still of considerable quantities. Car-following models have difficulties to describe the width of the gap-distribution correctly. This is certainly due to the fact that all models assume an invariant preferred headway for a driver, at least over a certain period of time. However, the results presented in [34, 14] indicate that this is not the case. Therefore, especially for the gaps, other models are needed that take into account the volatility of the preferred headway.

In addition to this, also other error measures are needed: An example would be the Kolmogorov-Smirnov statistics for the simulated and the empirical gap distribution.

3.3 Parameters Needed

As it has been already demonstrated in the introduction, for some models only a part of their parameters are actually needed in a specific traffic situation. The parameter T of Newell's model, which didn't have any influence on the gap-error, is such an example. Furthermore, such a parameter is a challenge for any minimization algorithm. And in fact, for this example it was not possible to compute the minimum with the minimization routines provided in the nls() function of "gnu R" [26].

To further quantify this effect, the average importance of all the parameters of a given model is computed. As displayed in Figure 8, it turns out that in the situation considered here, only a fraction of the parameters used are actually needed, i. e. carry importance.

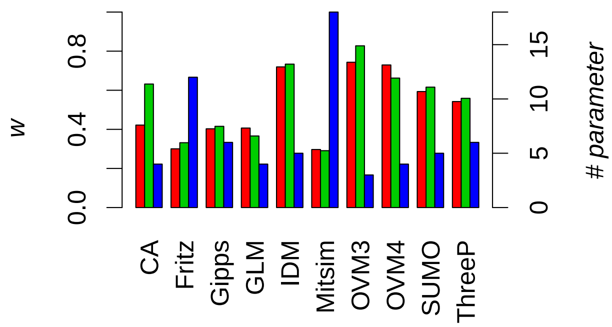
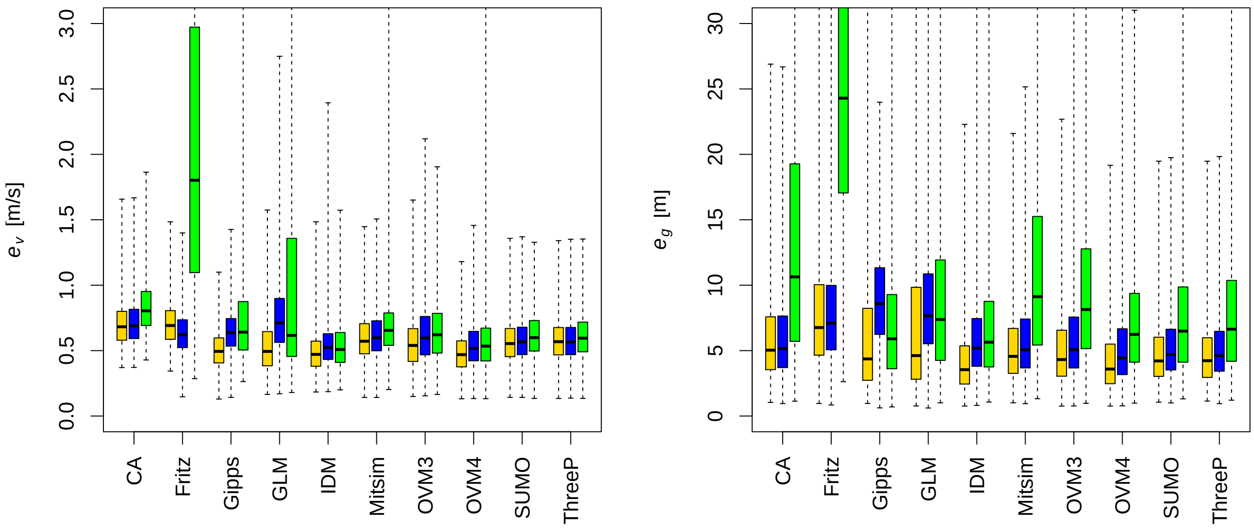


FIGURE 8 Importance w of the parameter sets of the models. The red bars indicate the importance with respect to the speed-error, while the green bars stand for the importance of the gap-error. The blue bars code the number of parameters of the models. There is a certain correlation visible, the more parameter a model has, the smaller is their importance.

That does not mean a weak parameter can be completely neglected. As an alternative, setting this parameter to a constant value and therefore not fitting it may help. When a certain parameter in a model turns out to be unimportant in any situation, then setting it to zero might in fact improve the model. However, this requires a much broader approach than the one carried out here.

3.4 Is It Minimum?

In addition to the brute force approach above, the true minima for the different models and episodes have been tried to compute, too. The open source library NLOPT [15] has been used for this purpose, since it provides several different algorithms to find minima numerically. No systematic test of the different methods have been performed, however, a short pre-test indicated that the BOBYQA algorithm (Bound Optimization BY Quadratic Approximation) [25] needed on average the shortest amount of iterations to find something that the algorithm claims to be a minimum. Again, this has not been tested thoroughly, since these results are only meant to be a consistency check. Typically, even BOBYQA needs about 100 iterations to find a minimum. The optimization was started just in the middle of the range of the parameters as it has been used for the quasi-random sequences. In addition, the boundaries as defined for the quasi-random sequences have been set as boundaries for the minimization procedure as well. In a handful of cases, the algorithm did not converge. The results are displayed in Figure 9. It is shown that there is little difference between the exact minimization and minimum obtained from the quasi-random approach. In most cases, the exact



(a) Speed error e_v **(b) Gap error e_g**

FIGURE 9 Comparison of the results of the numerical minimization with the results from the quasi-random scanning of the parameter space. Left: for the speed-error, right: for the gap-error. In gold, the error-distribution of the minimization. In blue, the result from the quasi-random sequences, and in green the distribution of the non-optimized variable at the minimum of the optimized variable for the numerical minimization.

minimization yields slightly smaller values than the quasi-random approach. Moreover, the error of the non-optimized variable is often much bigger but not always.

4 CONCLUSIONS

This paper has introduced a new technique to explore the parameter space of a given microscopic car-following model. The approach scans brute force the parameter space of such a model by filling the p-dimensional parameter space with quasi-random numbers. From this, a couple of statistics can be derived, which has been done exemplarily for two objectives: the gap-error e_v and the speed-error e_g of ten different models. Furthermore, a new characteristic called “importance” can be assigned to each model parameter, which may help to find and classify essential and negligible parameters and afterwards fix these parameters to ensure a better calibration / optimization to the empirical data / traffic situation under consideration. Clearly, this is a function of the data used for comparing the models with: When analyzing pure car-following episodes as has been done in this work, then the importance of the parameters for car-following is tested for.

This approach can be generalized for other models and traffic situations. It will work in the same manner for microscopic models simulating intersections, or for fluid-dynamical models simulating freeway stretches, or even for demand models computing origin-destination tables (however, it is much more difficult to have good empirical data in this case). It bears similarities with sensitivity analysis; however, it makes the whole approach much more general.

In addition to this technical innovation, also a few interesting results are found regarding the example used throughout this paper, i. e. microscopic models in car-following situations. These results can be summarized in a simple sentence: It seems that at the level of car-following, the difference between even so fairly different models as the MITSIM and the Newell model is almost negligible. That might be different in other situations, e. g. when looking at the behavior of vehicles at traffic lights or for lane-changing. But for car-following, little can be learned by comparing models with car-following data.

The results obtained here do not mean that car-following models do not need calibration. Quite the contrary is true! When looking at the Pareto areas for all the models one sees that the difference between the best and the worst parameter set is dramatic: There is a factor of roughly 100 between a good and a bad parameter set. So, calibration is needed to narrow this large range to a minimum.

Finally, the analysis put forward here points toward a weakness of all the models tested, and that is their difficulty in simulating the distance behavior of human drivers. The data simply display much greater variability than the models can support. This leads to the exceptionally large gap-error reported here. Therefore, still more work is needed to remedy this short-coming. In addition, this will also need different error measures which can cope with such a type of stochastic model since it cannot be expected that the seemingly random jumps in the preferred headway can be reproduced by a deterministic model.

ACKNOWLEDGMENT

Our special thanks go to Professor Rui Jiang from Beijing Jiaotong University, China, for kindly providing the experimental data.

REFERENCES

[1] K. I. Ahmed. *Modelling Drivers' Acceleration and Lane-Changing Behavior*. PhD thesis, MIT, 1999.

[2] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. Dynamical model of traffic congestion and numerical simulation. *Physical Review E*, 51:1035 – 1042, 2 1995.

[3] E. Brockfeld, R. D. Kühne, A. Skabardonis, and P. Wagner. Towards a benchmarking of microscopic traffic flow models. *Transportation Research Records*, 1852:124 – 129, 2003.

[4] E. Brockfeld, R. D. Kühne, and P. Wagner. Calibration and validation of microscopic traffic flow models. *Transportation Research Records*, 1876:62 – 70, 2004.

[5] E. Brockfeld, R. D. Kühne, and P. Wagner. Calibration and validation of microscopic traffic flow models. *Transportation Research Records*, 1934:179 – 187, 2005.

[6] John Burkardt. Source code to produce halton sequences. electronically (URL), 2003.

[7] Biagio Ciuffo and Vincenzo Punzo. Sensitivity analysis of car-following models: methodology and application. *Transportation Research Records: Journal of Transportation Research Board, TRB, National Research Council, Washington D.C.*, 2011.

[8] R. Dowling, A. Skabardonis, J. Halkias, G. McHale, and G. Zammit. Guidelines for calibration of microsimulation models: framework and applications. *Transportation Research Record*, 1876:1 – 9, 2004.

[9] H.-T. Fritzsche. A model for traffic simulation. *Transportation Engineering And Control*, (5):317 – 321, 1994.

[10] P.G. Gipps. A behavioural car-following model for computer simulation. *Transportation Research Part B*, 15(2):105 – 111, 1981.

[11] John Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2:84 – 90, 1960.

[12] John Halton and G.B. Smith. Algorithm 247: Radical-inverse quasi-random point sequence. *Communications of the ACM*, 7:701 – 702, 1964.

[13] J.C. Helton, J.D. Johnson, C.J. Sallaberry, and C.B. Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, 91(10 – 11):1175 – 1209, 2006.

[14] Rui Jiang, Mao-Bin Hu, HM Zhang, Zi-You Gao, Bin Jia, Qing-Song Wu, Bing Wang, and Ming Yang. Traffic experiment reveals the nature of car-following. *PloS one*, 9(4):e94351, 2014.

[15] Steven G. Johnson. The nlopt nonlinear-optimization package. web site, March 2012.

[16] B. S. Kerner, S. L. Klenov, and D. E. Wolf. Cellular automata approach to the three phase traffic flow theory. *preprint*, 2002. http://arXiv.org/abs/cond-mat/0206370.

[17] Ladislav Kocis and William Whiten. Computational investigations of low-discrepancy sequences. *ACM Transactions on Mathematical Software*, 23(2):266 – 294, 1997.

[18] D. Krajzewicz, G. Hertkorn, C. Rössel, and P. Wagner. SUMO (Simulation of Urban MObility); an open-source traffic simulation. In A. Al-Akaidi, editor, *Proceedings of the 4th Middle East Symposium on Simulation and Modelling (MESM 2002)*. SCS European Publishing House, 2002.

[19] Stefan Krauss, Peter Wagner, and Christian Gawron. Metastable states in a microscopic model of traffic flow. *Physical Review E*, 55(5):5597 – 5602, 1997.

[20] Sefan Krauß, Peter Wagner, and Christian Gawron. Continuous limit of the nagel-schreckenberg model. *Physical Review E*, 54:3707 – 3712, Oct 1996.

[21] R. Liu and J. Wang. A general framework for the calibration and validation of car-following models along an uninterrupted open highway. In B. Heydecker, editor, *Mathematics in Transport*, page 111 – 124. Elsevier, 2007.

[22] R. Mahnke and N. Pieret. Stochastic master-equation approach to aggregation in freeway traffic. *Physical Review E*, 56:2666 – 2671, Sep 1997.

[23] Kai Nagel and Michael Schreckenberg. A cellular automaton model for freeway traffic. *Journal de Physique I France*, 2:2221 – 2229, 1992. According to the web page of J Phys. I FRance, this paper has so far (12-9-2012) 829 cross references.

[24] G.F. Newell. A simplified car-following theory: a lower order model. *Transportation Research Part B*, 36(3):195 – 205, 2002.

[25] M. J. D. Powell. The bobyqa algorithm for bound constrained optimization without derivatives. Technical Report NA2009/06 (2009), Department of Applied Mathematics and Theoretical Physics, Cambridge England, 2009.

[26] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.

[27] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, v181 n2 (201002):259 – 270, 2010.

[28] A. Saltelli, M. Ratto, T. Anres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. Global sensitivity analysis - the primer, 2008.

[29] R. E. Steuer. *Multiple Criteria Optimization: Theory, Computation, and Application*. Wiley, New York, 1986.

[30] T. Toledo, H. N. Koutsopoulos, A. Davol, M. Ben-Akiva, W. Burghout, I. Andréasson, T. Johansson, and C. Lundin. Calibration and validation of microscopic traffic simulation tools: Stockholm case study. *Transportation Research Records*, 1831:66 – 75, 2003.

[31] Martin Treiber, Ansgar Hennecke, and Dirk Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, 62(2):1805, 2000.

[32] T.G. Trucano, L.P. Swiler, T. Igusa, W.L. Oberkampf, and M. Pilch. Calibration, validation, and sensitivity analysis: What's what. *Reliability Engineering & System Safety*, 91(10 – 11):1331 – 1357, 2006.

[33] P. Wagner, R. Nippold, and T. Toledo. Calibration by accleration or by trajectory (run it!). In *Proceeding of the Traffic Flow Theory Conference 2010*, 2010.

[34] Peter Wagner. Analyzing fluctuations in car-following. *Transportation Research Part B*, 46:1384 – 1392, October 2012.

[35] Peter Wagner, Gunnar Flötteröd, Ronald Nippold, and Yun-Pang Wang. Simplified car-following models. In *Proceedings of the Transportation Research Board*, number 91, page 1 – 11. Transportation Research Board, January 2012. only on CD.