Performance of car following behaviour in microscopic traffic flow models

Elmar Brockfeld, Rene Kelpin, Peter Wagner
Institute of Transport Research, German Aerospace Center
Rutherfordstraße 2, 12489 Berlin, Germany
E-mail: elmar.brockfeld@dlr.de, rene.kelpin@dlr.de, peter.wagner@dlr.de

Abstract. In this paper ten microscopic traffic flow models of very different kind are analysed concerning the correct reproduction of the car-following behaviour on single lane roads. The models are calibrated and validated with data collected via DGPS-equipped cars (Differential Global Positioning System) on a test track in Japan. The positions of the cars are delivered every 0.1 second with very high accuracy, which is perfect for analysing the car following behaviour. To calibrate the models, in each case one driver pair is under consideration. The measured data of a leading car are fed into the model under consideration and the model is used to compute the behaviour of a following car. In the analysis the resulting simulated time series of headways are carried out and the deviations to the measured headways are calculated to calibrate the models. For validation purposes a driver-independent and a driver-specific approach have been conducted. The results show that calibration errors of about 13 % to 19 % are hard to undercut. The driver-independent validation results give additional errors of about 6 percentage points, which can be slightly reduced by about 1 to 3 percentage points with a driver-special validation. Most interestingly, no model could be denoted to be the best, even very sophisticated models do not perform better than simple ones.

Key Words: modelling, simulation, traffic flow, calibration, validation, DGPS

Introduction

Microscopic simulation models are becoming increasingly important tools in modelling transport systems. They are applied in simulation programs for transport planning, traffic forecasting and advanced vehicle control and safety systems (AVCSS). An important part of the models are the microscopic sub-models which describe the interaction between adjacent vehicles. For that purpose rules and equations are defined which describe the car-following and lane changing behaviour of the vehicles. An essential problem is the calibration and validation of the parameters used in these rules.
Out of a vast amount of existing models (see [1] and [2] for an overview on publicly available models) ten microscopic traffic flow models of very different kind are analysed concerning the correct reproduction of the car-following behaviour on single lane roads. In contrast to typical macroscopic analyses, which compare aggregated data on links (see [3] for example), this means an analysis from a very microscopic point of view.
The data used for calibration and validation are from car-following experiments conducted in Japan in October 2001 [4]. The data have been collected by letting nine

DGPS-equipped cars follow a lead car driving along a 3 km test track for about 15-30 minutes.
At first the experiments on the test track and the recorded data sets are briefly described and the simulation setup for testing the models is defined. In the following the measurement procedure for calculating the error differences between the recorded data and the data produced by the models is specified. After the tested models are listed and basically described, the calibration procedure and its results are presented. At the end the comprehensive design of the validation is described followed by the results obtained and leading to some conclusions.

1 Data and error measurement

1.1 The data and the simulation set-up

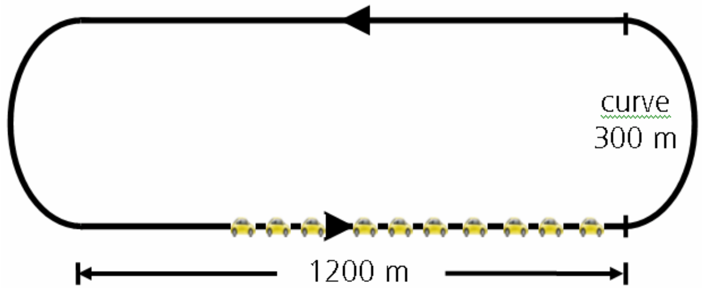

FIGURE 1: Sketch of the test track with ten cars driving on the course.

The data sets used for the analyses have been recorded on a test track in Hokkaido, Japan in October 2001 [4]. Eight experiments have been conducted, where nine cars drove on a 3 km test track (2 x 1.2 km straight segments and 2 x 0.3 km curves; see figure 1) for about 15-30 minutes in each experiment following a lead car, which performed some driving patterns. These are for example driving with constant speeds of 20, 40, 60 and 80 km/h for some time, varying speeds (regularly increasing/decreasing speed) and emulating many accelerations/decelerations as they are typical at intersections. The regulary increase/decrease of speed is done with different frequencies, the velocity cycles from 20 to 60 km/h being performed one to four times on the straight segments.
To reduce driver-dependent correlations between the data sets, the drivers were exchanged between the cars after each experiment. Having all cars equipped with DGPS (Differential Global Positioning System), the position of each car is stored in 0.1 second intervals throughout each experiment. From these position data other important variables like the speed, the acceleration and the headway between the cars were extracted for simulation purposes. The accuracy of the DGPS is about 1 cm and the appointment of the speeds has got an error of less than 0.2 km/h as described in [4]. Thus, the data sets have got such a high resolution that they are adequate for the analysis of car-following behaviour and calibration of car-following models.

experiment	Duration [min]	Full loops	Driver succession (leader „D1“)
„11“	26	6	D1 D2 D3 D4 D5 D6 D7 D8 D9 D10
„12“	25	7	D1 D8 D7 D6 D5 D4 D3 D2 D9 D10
„13“	18	6	D1 D2 D3 D4 D5 D6 D7 D8 D9 D10
„21“	14	4	D1 D8 D7 D6 D5 D4 D3 D2 D9 D10

TABLE 1: Coverage of the data taken from the four experiments.

In this paper we present analyses concerning four of the eight experiments, namely the patterns mostly with intervals of constant speeds and wave-performing. The duration of the experiments are about 15 to 30 minutes as can be seen in table 1.
For the simulation set-up only two cars are considered at a time (see figure 2): the leading car is updated according to the speeds and positions in the recorded data sets and the following car is updated as defined by the equations of the used model.

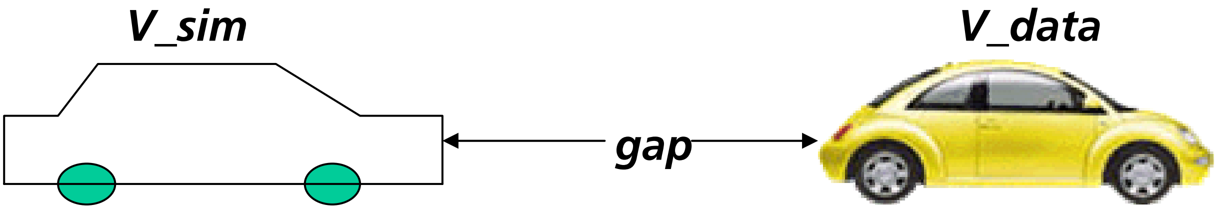

FIGURE 2: Simulation set-up.

1.2 Error measurement

The absolute error a model produces in comparison to a measured data set is calculated via the simple distance between a recorded time series and a simulated time series of headways. To get a percentage error it is additionally related to the average value of the time series in each particular data set:

$$e = \frac{\frac{1}{T}\sum_{t=0}^{T}\left|x^{(sim)}(t) - x^{(obs)}(t)\right|}{\frac{1}{T}\sum_{t=0}^{T} x^{(obs)}(t)}, \qquad (1)$$

where $x^{(sim)}$ and $x^{(obs)}$ are a simulated and an observed traffic flow variable, which is in this case the headway between two cars. T is the time series over the total time of each experiment. As an example figure 3 shows an error measurement for the time series of the headway between a driver pair for one particular model.

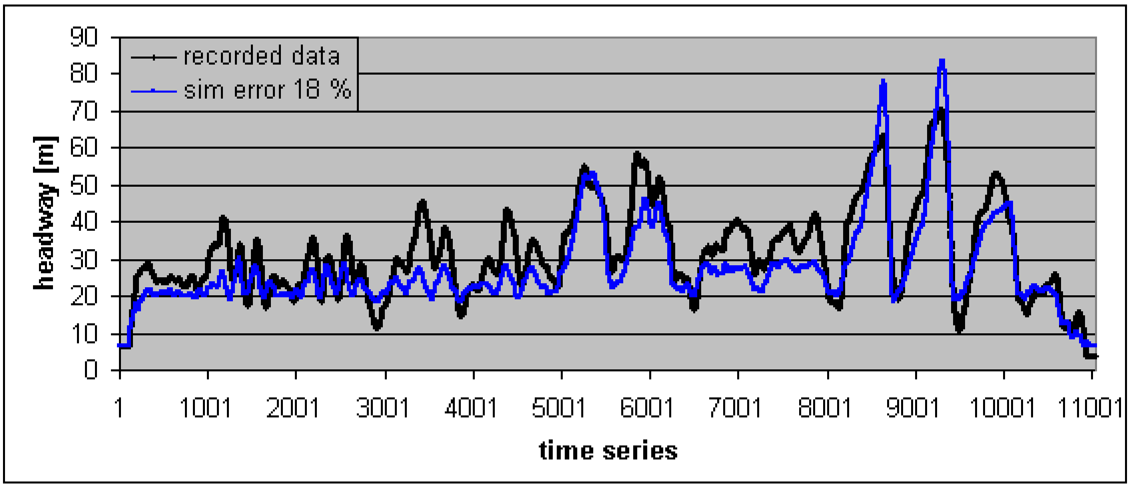

FIGURE 3: Example for error measurement using the time series of headways.

2 The models

The models used for the simulations are all microscopic traffic flow models, which describe the behaviour of a following car in relation to a leading car. For the vehicle movement, typically equations like the following were used, defining the new speed of a vehicle at time $t+\Delta t$, depending on the values of some variables at time t:

$$v(t+\Delta t) = f(g(t), v(t), V(t), \{p\})$$
$$g(t+\Delta t) = V(t) - v(t), \qquad (2)$$

where v is the speed of the following and V that of the leading car, respectively, and g is the headway between the cars. The symbol $\{p\}$ denotes a set of parameters of the model under consideration.

In the calibration approach the following microscopic traffic flow models of very different kind with 3 to 15 parameters have been tested. Some models are used in commercial simulation programs, which are popular in European countries, the USA and Japan, and some are scientific simulation approaches.

Abbreviation	**Description**	**params**
CA0.1	cellular automaton model [5]	4
SK_STAR	model based on the SK-model by S. Krauss [6]	7
OVM	“Optimal Velocity Model”, Bando, Hasebe [7]	4
IDM	“Intelligent Driver Model” [8]	7
IDMM	“Intelligent Driver Model with Memory” [9]	7
Newell	can be understood as a continuous CA with more variable acceleration and deceleration [10,11]	7
GIPPSLIKE	basic model by P.G. Gipps [12]	6
Aerde	Used in the simulation package INTEGRATION [13]	6
FRITZSCHE	used in the British software PARAMICS; similar to what is used in the German software VISSIM [14]	13
MitSim	model by Yang and Koutsopulus, used in the software MitSim [15]	15

TABLE 2: List of tested models.

The most basic parameters used by the models are the car length, the maximum speed, an acceleration rate (except for the CA0.1-model) and a deceleration rate (for most models). The acceleration and deceleration rates are specified in more detail in some models depending on the current speed or the current headway to the leading vehicle. Furthermore, some models (CA0.1, SK_STAR and MitSim) use some kind of stochastic parameters describing individual driver behaviour. Most models use something like a reaction time of the drivers to the behaviour of the leading car.

With these kinds of parameters a lot of the models are covered, except for some models with a more detailed conceptual design. The IDMM has as a special feature a memory effect. Depending on the density ahead, the cars try to keep their speeds according to a rolling horizon. The MitSim model defines two thresholds concerning the headway, which cause a switching between three different driving modes. Especially if a driver is very close to the leader the calculations become very sophisticated, depending on the headway, own speed, speed-difference and the current density. In addition to the basic simulation update equation (2) the model needs the speed of the leader one time step before as a special feature. The FRITZSCHE model provides switching to various driving modes, too. For this model the switching depends not only on the headway (g), but also on the speed-difference (ΔV) between the follower and the leader. Thus, a $(\Delta V, g)$-car following plane is divided into different regions of free driving, approaching, emergency brake and two other driving behaviours. As a specific, differing to equation (2), the model needs the

acceleration of the follower and the leader one time step before and uses some kind of “brake light” of the leader to react on its deceleration.
As the time step for the models is 0.1 seconds according to the recorded data, some models with a traditional time step of 1 second - as for example used for simple cellular automatons - have been modified to adopt for an arbitrarily small time-step.

3 Calibration

3.1 Calibration procedure

Altogether 36 vehicle pairs (4 experiments * 9 vehicle pairs) were used as data sets for the analyses. Each model has been calibrated with each of the 36 different constellations separately gaining optimal parameter sets for each “model - data set” combination. To find the optimal parameter constellations a gradient-free optimisation method known as the “downhill simplex method” [16] was used and started several times with different initialisation values for each “model - data set” pair. The variation in initialisation is done to avoid sticking with a local minimum, which of course can occur because getting a global minimum can not be guaranteed by those type of optimisation algorithms.

3.2 Calibration results

As can be seen in figure 5 the error rates of the models in comparison to the data sets during the calibration for each model reach from 9 to 24 % over all models and all 36 driver pairs. But no model appears to be significantly the best one since every model has the same problems with distinct data sets (21_1 for example) and other data sets can be simulated quite good with each model (11_8 for example). The average differences between the models reproducing single driver pairs is about 2.5 percentage points. Interestingly, it can be stated that models with more parameters than others do not necessarily reproduce the real data better. It is noteworthy that this diversity of the models is much smaller than the differences in the driver behaviour (mainly ranges from 13 % to 18 %, thus about 5 percentage points).
Looking at the average errors each model produces with the 36 data sets, it can be seen in figure 4, that, again, the differences of the models are not very big. The best model produces an error of 15.14 %, the worst one of 16.20 %. Thus, no model can be denoted to be the best and especially complex models do not produce better results than simple models.

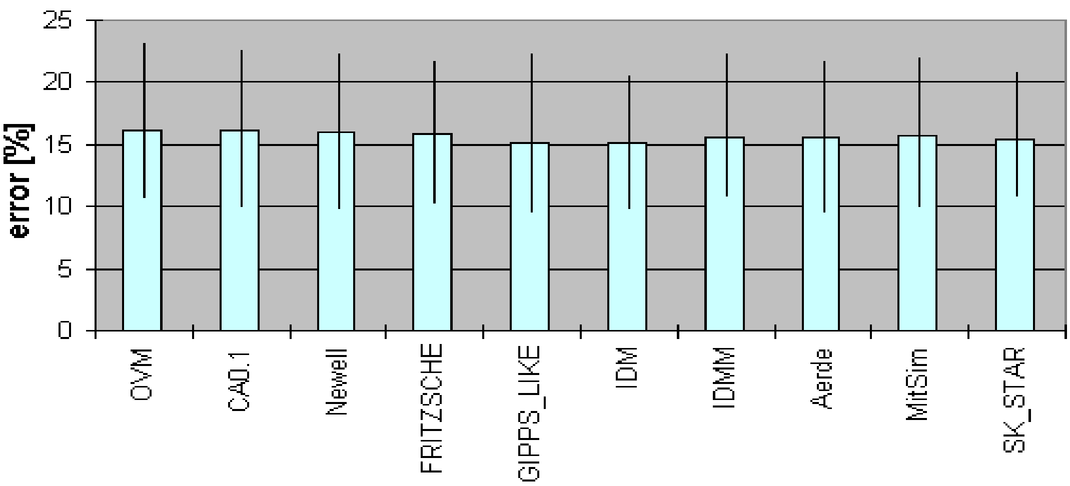

FIGURE 4: Mean calibration results for all models including the total result range.

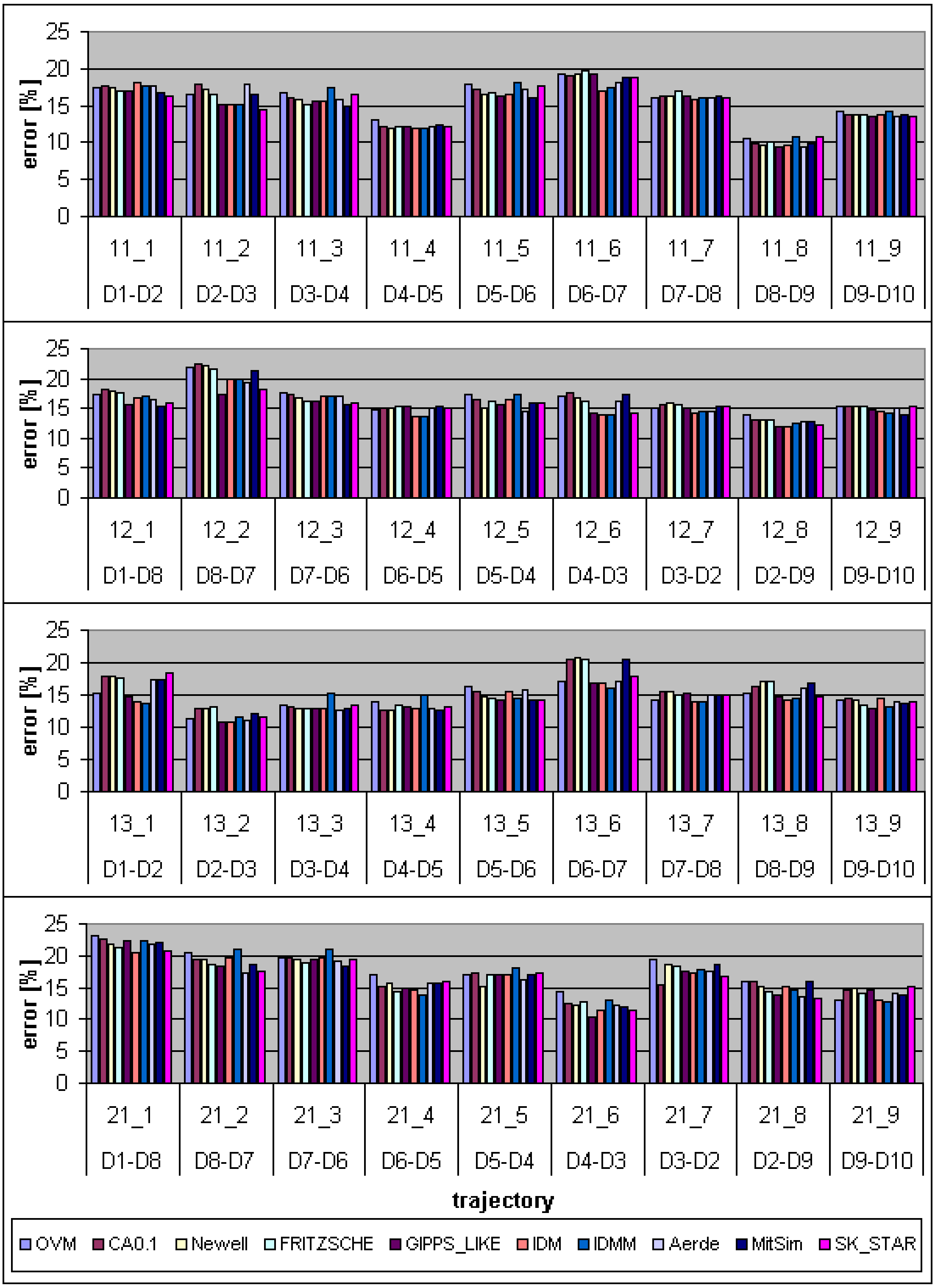

FIGURE 5: Errors of the models after the calibration procedure (data sets of four experiments 11_x, 12_x, 13_x, 21_x with 9 driver pairs; D1…D10: drivers).

4 Validation

4.1 Design of the two validation approaches VAL1 and VAL2

The design of the validations is as defined in table 3. The data sets of the driver pairs are denoted in the left part of the table. In general, a validation is understood in that way that the optimal parameter results obtained from a model after the calibration of a data set are taken to reproduce another data set. Two approaches have been conducted then, which are a driver-independent validation (VAL1) and a driver-special validation (VAL2).

	driver pair in experiment				driver independent validation (VAL1)		Driver-special validation (VAL2)				
Denotation (x = "11" || x = "12" || x = "13" || x = "21")	"11"	"12"	"13"	"21"	to validate (v = "11" || v = "12" || v = "13" || v = "21")	validate with (x != v)	driver to validate	validate cross over each row			
x_1	D1-D2	D1-D8	D1-D2	D1-D8	v_1	x_6	D2	11_1	12_7	13_1	14_7
x_2	D2-D3	D8-D7	D2-D3	D8-D7	v_2	x_5	D3	11_2	12_6	13_2	21_6
x_3	D3-D4	D7-D6	D3-D4	D7-D6	v_3	x_3	D4	11_3	12_5	13_3	21_5
x_4	D4-D5	D6-D5	D4-D5	D6-D5	v_4	x_4	D5	11_4	12_4	13_4	21_4
x_5	D5-D6	D5-D4	D5-D6	D5-D4	v_5	x_5	D6	11_5	12_3	13_5	21_3
x_6	D6-D7	D4-D3	D6-D7	D4-D3	v_6	x_6	D7	11_6	12_2	13_6	21_2
x_7	D7-D8	D3-D2	D7-D8	D3-D2	v_7	x_7	D8	11_7	12_1	13_7	21_1
x_8	D8-D9	D2-D9	D8-D9	D2-D9	v_8	x_8	D9	11_8	12_8	13_8	21_8
x_9	D9-D10	D9-D10	D9-D10	D9-D10	v_9	y_9	D10	11_9	12_9	13_9	21_9

TABLE 3: Denotation of driver pairs and validation design. (Example for a driver pair in the left part: data set 11_1 is D1-D2, thus driver D1 followed by driver D2)

The idea for VAL1 is to apply the optimal parameter sets for each model from one driver pair to another to check them for transferability. Each driver pair in each experiment is validated with one driver pair of the other experiments, thus three times. The total number of validations for each model is 9*4*3=108. For this validation a relation is defined in which no validation is performed with a data set the two drivers are part of in the other experiments (see left and center part of table 3).
The idea of the driver-special validation VAL2 is to validate the behaviour of each singular driver with the data sets obtained from him/her from the other experiments, respectively. Thus each driver is validated twelve times (in each experiment with the data sets of the other three experiments as defined in the right part of table 3) and thus the total number of driver-special validations is 9*12=108 for each model.

4.2 Validation results (VAL1 and VAL2)

Some of the results of the validation process VAL1 are shown in figure 6 and they draw a similar picture as the results obtained from the calibration. Except for singular problems of some models with special data sets all models share the same problems with the same data sets. The errors produced in these cases are mainly about 17 % to 27 %, which is about 5 to 10 percentage points more than in the calibration case.

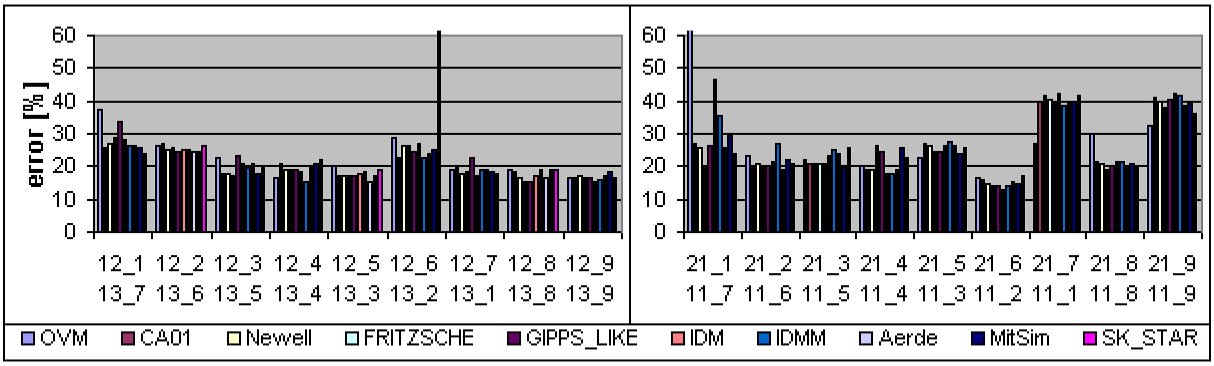

FIGURE 6: Exemplary results of the driver independent validation. Parameter sets obtained after calibration of the data sets in the bottom rows (13_x and 11_x) are taken to reproduce the data sets in the top rows (12_x and 21_x).

In some cases the errors are around 40 %, singularly much higher. In these cases the calibration seems to have reached what is known as “overfitting”. Because of the adaptation to a particular data set the obtained parameter sets after the calibration are not suitable to reproduce other data sets well. An example is dataset 11_1, whose calibrated parameters produce high errors when validated with dataset 21_7. The driver-special validation draws a similar picture as can be seen in figure 7. Again, cases of “overfitting” can be recognized (parameters of 11_4 taken to reproduce 13_4). The most values seem to be a bit lower than in the case of VAL1.

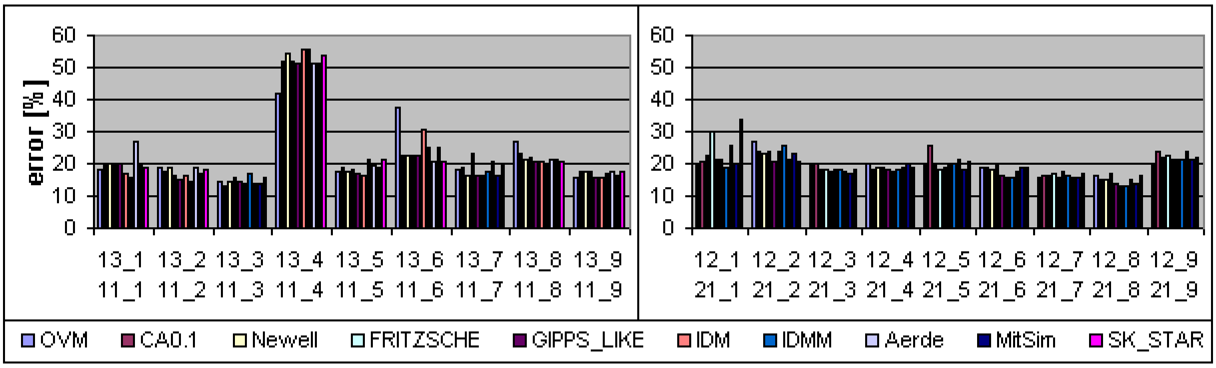

FIGURE 7: Exemplary results of the driver specific validation. Parameter sets obtained after calibration of the data sets in the bottom rows (11_x and 21_x) are taken to reproduce the data sets in the top rows (13_x and 12_x).

To roughly compare the results of the two different validation approaches, the distribution of all error results obtained (1080 for each validation strategy) are shown in figure 8 and compared to the distribution in the calibration case. There is an obvious difference between the three distributions. While the calibration errors are at most between 13 % and 19 % with a maximum frequency at 15-17 %, the validations produce errors which are much more spread. Most errors in VAL2 lie in the area of 15 % to 25 % with a maximum at 19-21 %, those of VAL1 are between 17 % and 27% with a maximum in the area of 21-23 %.

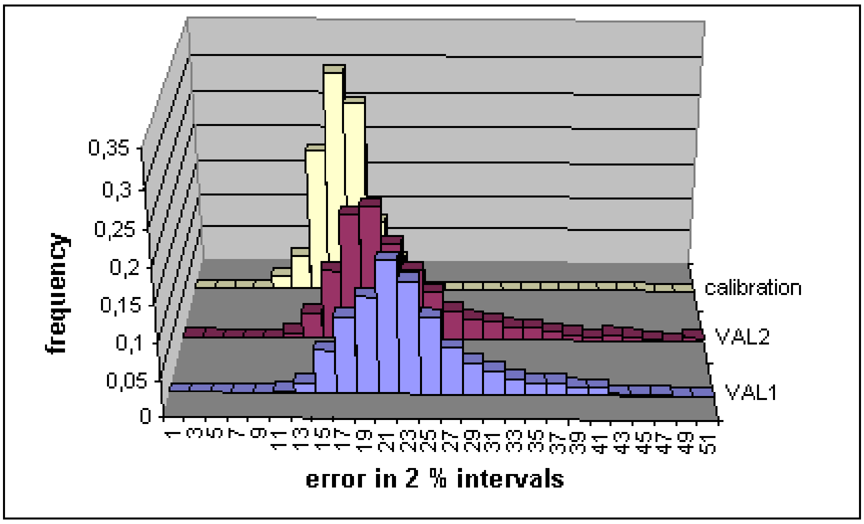

FIGURE 8: Distribution of errors produced by the calibration and the two validations.

To get a more precise insight into the benefit a driver-special validation has got in comparison to a driver-independent validation, the error values are sorted and plotted in two curves in figure 9 together with a subtracted curve. It can be seen that the driver-specific validation gives a benefit of at least 1 percentage points for about 85% of the datasets and about 2 percentage points for about 40 % of the data sets.

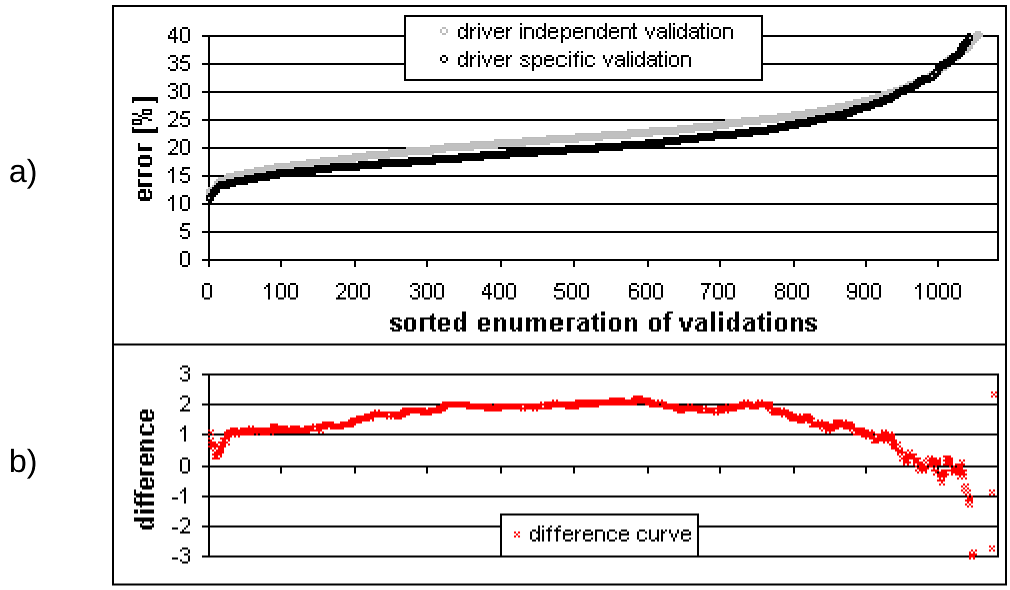

FIGURE 9: Comparison of driver specific and driver independent validation. All results (1080 simulations each) have been sorted and plotted into the top diagram a). Diagram b) shows the difference curve, which describes the benefit of VAL2.

Concerning the individual models table 4 gives an overview on the exact results obtained. The data related to the mean error values inherit distortions because of the peaks caused by “overfitting”. To avoid this influence, the median values have been calculated, which give a good overview on the results. The values after the calibration do not differ very much between the models with an amplitude of 1.25 percentage points around 15-16 %. The VAL1 gives medians of about 22 % for all models, which means an additional error of 6-7 percentage points. The VAL2 is able to slightly reduce this error for all models by 0.96 to 3.21 percentage points. But this still means total errors of 19-21 % for all models in average.

	OVM	**CA0.1**	**Newell**	**FRITZSCHE**	**GIPPS_LIKE**	**IDM**	**IDMM**	**Aerde**	**MitSim**	**SK_STAR**
CAL MEAN	16,20	16,18	16,03	15,92	15,14	15,16	15,58	15,50	15,71	15,39
VAL1 MEAN	25,44	23,16	22,54	22,38	23,14	23,89	23,13	23,48	22,49	24,74
VAL2 MEAN	24,90	22,37	21,47	21,76	20,75	21,87	21,86	21,71	21,06	22,78
VAL1 MEAN - CAL MEAN	9,24	6,97	6,51	6,47	8,00	8,73	7,55	7,98	6,78	9,35
VAL2 MEAN - CAL MEAN	8,70	6,19	5,44	5,84	5,61	6,71	6,28	6,20	5,35	7,38
VAL1 MEAN - VAL2 MEAN	0,54	0,79	1,07	0,62	2,39	2,01	1,27	1,77	1,43	1,96
CAL MEDIAN	16,09	16,04	15,71	15,90	15,13	14,84	14,92	15,73	15,63	15,35
VAL1 MEDIAN	22,09	21,70	21,66	21,91	22,05	21,78	21,88	22,29	21,60	22,58
VAL2 MEDIAN	20,25	20,74	19,63	19,76	18,84	19,51	19,47	20,35	19,51	20,75
VAL1 MEDIAN - CAL MEDIAN	6,00	5,66	5,95	6,00	6,92	6,94	6,96	6,56	5,97	7,23
VAL2 MEDIAN - CAL MEDIAN	4,16	4,70	3,92	3,86	3,71	4,67	4,55	4,61	3,88	5,40
VAL1 MEDIAN - VAL2 MEDIAN	1,84	0,96	2,03	2,15	3,21	2,27	2,41	1,95	2,09	1,83

TABLE 4 Summarised results of the calibration and the two validation approaches. (VAL1: driver independent; VAL2: driver specific)

5 CONCLUSIONS

The main results of the analyses are that all models produce nearly the same errors, thus sophisticated models with up to 15 parameters seem not to be better than simple models with only 4 or 6 parameters. In total it is found that the differences in the driver behaviour are much bigger than the diversity of the models. At last, the driver-special validation produces slightly better results than the driver-independent validation. Thus, the behaviour of individual drivers can be reproduced a bit more accurately than trying to transfer optimal parameter results from one driver to another.
But the results of the validation are in parts very bad which probably calls for the development of much better models. The other way to interpret the results is that – from this microscopic point of view – errors of about 12 % to 27 % can probably not be suppressed no matter what a model is used. This would be due to the different behaviour of each driver.

References

[1] A. Schadschneider, D. Chowdhury, L. Santen: Statistical Physics of Vehicular Traffic and Some Related Systems, Physics Reports, 2000, vol 329, 199.

[2] D. Helbing: Verkehrsdynamik: Neue Physikalische Modellierungskonzepte, Berlin, Germany: Springer, 1996.

[3] M. Cremer, M. Papageorgiou: Parameter identification for a traffic flow model, Automatica 17, 1981, pp. 837-843.

[4] G. S. Gurusinghe, T. Nakatsuji, Y. Azuta, P. Ranjitkar, Y. Tanaboriboon, *Multiple Car Following Data Using Real Time Kinematic Global Positioning System*, TRB2003-004137.

[5] K. Nagel and M. Schreckenberg, J. Physique I, 2221, 2, 1992.

[6] S. Krauss, P. Wagner, and C. Gawron, *Metastable states in a microscopic model of traffic flow*, Phys. Rev. E 55, 5597 – 5605 (1997).

[7] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, Y. Sugiyama, Dynamical model of traffic congestion …, Physical Review E, 239-253, 51, 1999.

[8] M Treiber, A Hennecke and D Helbing, Congested traffic states in empirical observation und numerical simulations, Physical Review E,1805-1824,62, 2000.

[9] M. Treiber, D. Helbing, Memory effects in microscopic traffic flow models and wide scattering in flow-density data ,see http://arxiv.org/abs/cond-mat/0304337.

[10] G. Newell, Theories of instabilities in dense highway traffic, Journal of the Operations Reasearch Society of Japan, vol 5, 9 – 54 (1962).

[11] G. Newell, A simplified car-following theory – a lower order model, Transportation Research B 36, 195 - 205 (2002).

[12] P.G. Gipps, A behavioral car following model for computer simulation, Transp. Res. B 15, 105 – 111 (1981).

[13] Brent C. Crowther, A Comparison of CORSIM and INTEGRATION for the Modeling of Stationary Bottlenecks, Master thesis, Virginia Polytechnic Institute and State University (2001).

[14] H.-T. Fritzsche, A model for traffic simulation, Traffic Engineering and Control. Vol 35, No. 5, 317-321 (1994).

[15] Kazi Iftekhar Ahmed, Modeling Drivers' Acceleration and Lane-Changing Behavior, Ph.D. thesis, MIT (1999).

[16] William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian Flannery: Numerical Recipes in C: The art of Scientific Computing, 2nd Edition, Cambridge University Press, 1992.