

TOWARDS A BENCHMARKING OF MICROSCOPIC TRAFFIC FLOW MODELS

Elmar Brockfeld,

Institute of Transport Research, German Aerospace Centre
Rutherfordstrasse 2, 12489 Berlin, Germany
phone: +49 30 67055 233, fax: +49 30 67055 202
email: elmar.brockfeld@dlr.de

Reinhart D. Kühne

Institute of Transport Research, German Aerospace Centre
Rutherfordstrasse 2, 12489 Berlin, Germany
phone: +49 30 67055 204, fax: +49 30 67055 202
email: reinhart.kuehne@dlr.de

Alexander Skabardonis

Institute of Transportation Studies
University of California, Berkeley CA 94720-1720
phone: (510) 642-9166, fax: (510) 642-1246
email: skabardonis@uclink4.berkeley.edu

Peter Wagner *

Institute of Transport Research, German Aerospace Centre
Rutherfordstrasse 2, 12489 Berlin, Germany
phone: +49 30 67055 233, fax: +49 30 67055 202
email: peter.wagner@dlr.de

For Presentation and Publication
82nd Annual Meeting
Transportation Research Board
January 2003
Washington, D.C.

August 1, 2002

No WORDS: 4,071
Plus 4 Figures (1,000)
Plus 1 Table (250)
TOTAL: 5,321

**Corresponding Author*

ABSTRACT

Several microscopic traffic flow models have been tested with a publicly available data set. The task was to predict the travel times between several observers along a one-lane rural road, given as boundary conditions the flow into this road and the flow out of it. By using nonlinear optimization, for each of the models the best matching set of parameters have been estimated. For this particular data set, the models that performed best are the ones with the smallest number of parameters. The average error rate of the best models is about 16%, however, this value is not very reliable: the error rate fluctuates between 2.5 and 25% for different parts of the data set.

INTRODUCTION

Right now, according to recent counts, up to one hundred different microscopic simulation models are known, see e.g., (1,2,3,4) for reviews. The fact, that these models belong to different scientific communities, that seem barely take any notice of each other, makes the situation even more confusing. The most prominent contributors to (not only microscopic) models of traffic flow are of course the traffic engineers and the physicists. What is missing in our opinion is some common idea about the worth of this plethora of models. I.e. one would like to know which model is the best for my application. Therefore, a benchmarking of these models is called for. The work presented here is a first step of a long-term research project to provide such a benchmark for microscopic traffic flow models, with the ultimate goal to reduce the above-mentioned number considerably. For more details, see (5).

To develop a commonly accepted benchmark, three things are needed:

1. a computer-implementable public description of the models,
2. publicly available data sets so that other groups are able to reproduce the benchmarks,
3. different combinations of testing algorithms and data sets that finally add up to provide such a benchmark.

The work here is mainly about issues one and the first steps into three.

Section 2 of the paper describes the methodology for testing the simulation models. A brief description of the selected models is given in Section 3. The database and application of the models is described in Section 4. The results are presented in Section 5. The last section summarizes the study findings.

METHODOLOGY FOR MODELS TESTING

The models this text is concerned with can be classified roughly as belonging to one of the following groups: cellular automata (discrete space, discrete time), mathematical maps (continuous space, discrete time), ordinary or delay differential equations (anything continuous) and “mesoscopic” models. An example for a mesoscopic model is the queueing model, described in Section 3. What seems much more interesting, but in a certain sense is still missing, would be a classification according to behavior, i.e., according to the macroscopic features a certain model displays.

In general, any microscopic simulation model is defined by a set of equations (for step size h going to zero, a differential equation results):

$$\begin{aligned} v(t+h) &= f(g(t), v(t), \tilde{v}(t), \xi(t); p), \\ x(t+h) &= x(t) + v(t)h. \end{aligned} \quad (1)$$

Here, $x(t), v(t)$ are the position and velocity of a following car, and $\tilde{x}(t), \tilde{v}(t)$ are the position and velocity of the leading car, respectively. The variable $g(t) = \tilde{x}(t) - x(t) - \ell$ is the free space in front of the following car (ℓ is the length of a car). The noise term $\xi(t)$ need not be white noise, and p is a set of parameters, that allows adapting the model to varying circumstances. The equations (1) above are written very generally, the left-hand side is meant as the time update of the current system-state no matter to which class the model belongs.

Given a certain data set, the objective is to determine the set of parameters that best fit the data set. This can be done as follows:

a) choosing a certain error measure $e(p)$ for instance the mean absolute error for any system observable performance metric T (e.g., travel time on a highway section):

$$e(p) = \frac{\langle |T_{sim}(p) - T_{obs}| \rangle}{\langle T_{obs} \rangle}, \quad (2)$$

b) run a simulation of the model with a certain set of parameters, and

c) use an algorithm to improve e by changing the set of parameters p .

Usually, those models are very hard to analyze analytically, ruling out the possibility of computing the Jacobi-matrix with respect to the parameters, therefore a so-called direct search approach is needed (7,8,9,10). Direct-search methods work without the need to compute derivatives or the need of an explicit analytical formulation of the system to be optimized, a computer implementation will do. A detailed description of these methods is by far beyond the scope of this paper. For example, the method developed in (8) elaborates on the simple idea to compute a quadratic approximation to the function values found so far and using the minimum of this quadratic approximation as a guess for the next iteration. Differently from the more familiar gradient-based optimization algorithms, direct search methods initially need a simplex in the n-dimensional parameter space to get started.

The above described non-linear optimization algorithms are not guaranteed to yield anything useful, since they can get stuck into a local minimum. For the examples considered in this work, however, they seem to work surprisingly good. (The usual precautions have to be taken: restart the algorithm after settling to a minimum; start from different initial conditions; for low dimensional optimization problems (small set of parameters) the parameter space can be searched and even visualized more or less thoroughly etc.)

SELECTED MODELS

The evaluation methodology was applied to test the following microscopic models, using a real-life data set (described in Section 4):

- **CA**: Cellular automaton model (11)
- **CT**: Cell transmission model (12) as a reference model,
- **FRITZ**: Fritzsche model (13), which is the basis for the simulation software PARAMICS,
- **GIPPS/SK**: The Gipps model (14), and a variant of it from the physics community (15); this model is the basis of the model used in the simulator AIMSUN2,
- **IDM**: the intelligent driver model (16), again from the physics community,
- **OVM**: the optimal velocity model (17,18),
- **uQUEUE**: a queueing model, again for reference reasons,
- **MITSIM**: MITSim-Model as described in (6) which in parts can be understood as an implementation of the classic car following family of models (19),
- **INT**: the model used in the simulation package INTEGRATION (20),
- **VDR++/caSync**: two recent members of CA-family, the so called VDR-model (21) and a recent version (22), both of them claim to describe what is know as synchronized traffic flow.

Table 1 provides basic information about each model. Practically all the models have a two parameters in common, the maximum speed v_{\max} and the generalized length of a vehicle ℓ , that is the length of the vehicle plus the minimum distance a driver keeps to the car in front when standing in a jam (which defines the jam density.) Some other commonly used parameters are the maximum acceleration and deceleration rates a, b respectively, the reaction time τ and the strength ε of the noise for the stochastic models. The models that use a partitioning of space have as an additional parameter the cell size λ . Not exactly a parameter, the step-size h is usually needed. For the CASYNC, FRITZ, MitSim and VDR++ models, not all parameters are listed, which is indicated by dots in the corresponding entry of the list.

Note also that several of the models include so-called hidden parameters. For example, the reaction time of drivers is simply set to one second. Then the authors “forget” about this parameter, it does not enter the equations anymore. In the following, we tried to unearth and at least to mention those hidden parameters.

Additional information about each model can be found in (5). A short description of some models in given below:

Cell transmission (11)

Intended as an approximation to the Lighthill Whitham theory of traffic flow, this model divides a road into small cells of length $\lambda = h v_{\max}$. Then cars (or better occupancies, because it could be fractions of cars) are moved between the cells according to a very simple rule:

$$\delta n(i \rightarrow i+1) = \min\{n_i, \beta(N - n_{i+1})\}. \quad (3)$$

Here, $\delta n(i \rightarrow i+1)$ is the flow from cell i into cell $i+1$, and N is the maximum occupancy of a cell. In the implementation used in this work, all cells are alike. Of course, there is a relationship between N and the car length ℓ , leaving this model with the four parameters v_{\max} , $\beta = w/v_{\max}$, ℓ and the step size h . The parameter w is the speed of the backward running jam wave. The other parameters that have been obtained by this approach are given by $v_{\max} = 21 \text{ m/s}$, $w = 4.5 \text{ m/s}$, $h = 0.58 \text{ s}$, $\ell = 7.6 \text{ m}$, which are comparable with results found in the literature, especially the speed of the backward running jam wave. This value could be improved slightly (<1%), if one allows for a fundamental diagram that depends on the state of the cell in front, at the cost of an additional parameter.

Queueing model

Models of this type are especially interesting, since they are the fastest known simulation models that still have individual cars where the cell transmission model described above is about densities. Hence, they provide an excellent tool for a couple of applications (like dynamic traffic assignment). They can be described as a mesoscopic model, comparable to what is used in the simulation package DYNEMO. Again, the road is divided into cells of size λ , where λ is about 100 m. Any cell can hold at most N cars and is organized as a first-in first-out priority queue. When a car enters such a cell, it gets assigned the exit time, which is its exit time t plus the minimum travel time, which is simply λ/v_{\max} . After a car has left a cell, it has to be made sure, that the following car fulfills the flow constraint, i.e., whatever its exit time is, it is not allowed to leave before the time $t + \tau_{ff}$, where $\tau_{ff} = 1/c$ with c as the capacity of the cell. Furthermore, this waiting time may depend on the state of the next cell, it is set to the minimum of τ_{ff} and $\tilde{n} \tau_{jj}$. If the following cell is full, the car must stay in its cell. As mentioned above, this model is numerically very efficient, and its error rate is not much different from the model that performs best.

Fritzsche model (13)

Is included here as an example of a fairly complex model, featuring twelve parameters to fit. A variant of this model is being used with the simulation software PARAMICS, however, nothing is known about the difference between the published version and the version used in PARAMICS. The basic idea is to divide the $(\Delta v, g)$ -car following plane into different regions, with different behaviors. The regions are called following I and II, emergency (the distance to the car ahead is too small, so try to brake as hard as possible), approaching and driving freely. It could be seen easily, that any line in this plane is described by at least two parameters, so one readily ends up with twelve parameters.

Gipps model (14)/SK model (15)

Again, this model is the basis of a commercially available simulation package, AIMSUN2. There are actually two versions of it, the original version of Gipps, for which some new

results are available (25), and a version that is used by the physics community. The latter (named SK thereafter) has been designed rigorously for speed and simplicity, and has been investigated very thoroughly. Both models' formulation are based on the premise that the following condition must hold for safe car-following:

$$d(v) + v\tau \leq d(\tilde{v}) + g. \quad (4)$$

Here, $d(v)$ is the braking distance, for constant deceleration rate b it is just $d(v) = v^2 / (2b)$. This equation has only two parameters, deceleration rate b , and reaction time τ . To make a microscopic simulation model the acceleration rate a and maximum speed v_{\max} are needed. The Gipps model has a more complicated acceleration equation and an additional safety factor, while the SK-model adds a stochastic term with amplitude ε .

DATA-SET USED AND IMPLEMENTATION ISSUES

A unique data set recorded by Carlos Daganzo (23,24) and co-workers was used to evaluate the models. The data were obtained by observers along a four-mile (6.2 km) section of San Pablo Dam Road, a single lane highway in San Francisco Bay Area, California. Eight observers recorded the times each vehicle passed the observer location. Figure 1 shows the observer positions and segment numbering in the study section. At the end, 80 m behind the last observer, there was a traffic light. A special car provided the start of a sequence, thereby assigning a unique identification number to each vehicle. Provided the observers made no error, and there were no passing, then the cumulative N -curves of the eight observers contain all information needed, especially the travel times. There were two data sets for two different days, each recorded from about 7 a.m. to 9 a.m. and containing about 2,300 observations each. Unfortunately, no speeds have been recorded, however the models can cope with this omission surprisingly well.

Figure 2 shows a plot of speeds as a function of space and time. The generation and final dissolution of the congested area in this system could be observed. While the actual traffic demand changes only slightly, the changes in the timing of the traffic signal cause the congested area finally to cover nearly all the 6.2 km of the study section.

To facilitate the data processing and comparison with model predictions, we discarded the data of observer five on the first day, and of observer six on the second day, because there were big holes in those data sets. Figure 1 shows the simulation set up and segment numbering for each data set.

Each microscopic simulation models was fed with the cars observed by the first observer. However, for some of the models this caused problems, because the data set contained very short headways. Therefore, we tried to insert the cars with the maximum possible speed, allowing the car-following dynamics to take care of congested conditions inside the system. This means that the traces $x(t)$, $v(t)$, (if one could manage to compare them to the real ones),

are not correct for the first couple of meters. Still, some models needed more effort than others, and still there were some models where we could not manage to insert them smoothly into the system: they developed a jam right after the insertion, leading to large and unrealistic error measures.

The outflow condition was handled as follows: a virtual traffic light is put at the position of the eighth observer, which is switched to green at a certain time if more observed cars than simulated cars are out of the system. Other, softer constraints have been tested as well, but do not give better results, so the simplest scheme has been used to enforce that roughly the same number of cars are in the system as in reality. Since all the models need a car in front, two schemes for providing such a lead car have been used. Either the car that just left the system is updated with the cars still in the system, but is not allowed to accelerate, or the speed of the car that just the system is set to maximum speed. Different models perform differently i.e., some models perform better with the first rule, while others perform better with the second rule. Nevertheless, this procedure introduces errors, however, when looking at the errors of the models it could be seen, that this segment is not the worst of all the segments.

The actual implementation of all these models took much longer than originally expected, since all of them need some hand crafting until they could finally be managed to run. This was also due to the hidden parameters included in several models.

RESULTS

For all the models, the simulated travel times per segment were extracted, and the error measure was calculated per Equation (2). Since the data set contains two days of data, any model could be fitted with the data from the first day, and run it with those parameters with the data of the second day (and vice versa). Interestingly, all models performed worse on the second day data, even when fitting the parameter with the second day, and then running the model on the first day, the error for the first day is smaller.

Figure 3 shows the error in travel times on each segment. The distribution of the error along the six segments is similar for different models: they perform best for the first, fourth and last segment and worst for the second, third and fifth segment. Further analyses have shown that the error is distributed in time also very inhomogeneously, with errors between 2% and 30% for the same model.

Up to now, we do not have any idea for this very peculiar type of pattern; however, we could not rule out that there are recording errors in the data set. An indication of this is visible already in Figure 2. It could be seen that the speed of some vehicles is about 40 m/s (100 mph), which is hardly possible on this road. However, that is not as bad as it sounds, because in reality data are never clean, so a model should have a certain kind of robustness against buggy data. Nevertheless, it is not completely clear, that the error found so far is the error in

the data itself, and therefore it's not that surprising, that the models all seem to perform almost identically.

Figure 4 shows the average error across the study section for each model. The results are shown for each data set with the best set of parameters. The best models give an average best error rate (meaning: picking the best value out of the four) of 15.5%. It's the IDM and the cell transmission model, followed by a group of models around 16-17% (OVM, SK, queuing model, Gipps). There is a group of models that stay well below 20% error, and another group of models that are above 20%. The difference between the top five models is hardly significant. Interestingly, the more complicated models seem to have worse performance. Also, adding parameters to the base model, e.g., hidden parameters in the SK and Gipps models did not improve the model performance.

Of course it could not be ruled out that the weaker performance of the more complex models (e.g., the 12 parameter Fritzsche model with an average error of over 20%) is an artifact from the nonlinear optimization, which tends to get more difficult with higher dimensional spaces. Therefore, for some of the hard to fit models a simulated annealing technique has been tried, that has the merits of not getting stuck that easily in local minima. No significant differences in the outcome for the different optimization routines have been found.

CONCLUSIONS

These results are interesting. However, before final conclusions can be stated, a lot more work has to be done to confirm these results by using other data sets, which are currently performed. Unfortunately, since most data are from freeways, one has then to deal with lane changing matters, which easily could double the number of parameters that have to be fitted. This clearly suggests to strive for models with few parameters. We are very skeptical that models with 20 parameters can be fitted reliably to whatever data at hand, and it seems therefore interesting to think about models that have a chance of getting tested and falsified.

Nevertheless, it is tempting to say at least (and at last) this: it is simple to invent a new model, but it is hard to see how good it compares to reality. This may be the reason, why we still have more models than results about them.

ACKNOWLEDGEMENTS

Thanks are to Carlos Daganzo and his team for this data set, and for a lot of discussions during the stay of PW at the ITS of the UC Berkeley. Additionally, we would like to acknowledge discussions with Nils Eissfeldt, Boris Kerner, and Andreas Schadschneider.

REFERENCES

1. D. Chowdhury, L. Santen, and A. Schadschneider, *Statistical Physics of Vehicular Traffic and Some Related Systems*, Physics Reports **329**, 199 (2000). [See also at arxiv.org/abs/cond-mat/0070053]
2. D. Helbing, *Traffic and Related Self-Driven Many-Particle Systems*, Reviews of Modern Physics **73**(4), 1067-1141 (2001). [See also at arxiv.org/abs/cond-mat/0012229]
3. Smartest, see <http://www.its.leeds.ac.uk/projects/smartest/>
4. A. Skabardonis, internal report, ITS of the UCB (1999).
5. Clearinghouse for data and models, see <http://ivf.dlr.de/clearing>
6. Kazi Iftekhar Ahmed, *Modeling Drivers' Acceleration and Lane-Changing Behavior*, Ph.D. thesis, MIT (1999).
7. W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in C*, chapter 10, Cambridge University Press and references therein (2002).
8. M.J.D. Powell, UOBYQA: unconstrained optimization by quadratic approximation, Report 2000/14, <http://www.damtp.cam.ac.uk/user/na/reports.html> (2000).
9. M.J.D. Powell, *A direct search optimization method that models the objective and constraint functions by linear interpolation*, in S. Gomez and J.-P. Hennart (eds), *Advances in Optimization and Numerical Analysis*, Kluwer Academic 1994.
10. L. Ingber, *Simulated annealing: Practice versus theory*, Math. Comp. Modelling **18**, 29 – 57 (1993).
11. K. Nagel and M. Schreckenberg, J. Physique I, 2221, **2**, 1992.
12. C.F. Daganzo, *The Cell transmission model: a simple dynamical representation of highway traffic*, Transp. Res. B **28**, 269 – 287 (1994).
13. H.-T. Fritzsche, *A model for traffic simulation*, Transp. Engin. Contr. **5**, 317 – 321 (1994).
14. P.G. Gipps, *A behavioral car following model for computer simulation*, Transp. Res. B **15**, 105 – 111 (1981).
15. S. Krauss, P. Wagner, and C. Gawron, *Metastable states in a microscopic model of traffic flow*, Phys. Rev. E **55**, 5597 – 5605 (1997).
16. M. Treiber, A. Hennecke and D. Helbing, *Congested traffic states in empirical observation and numerical simulations*, Physical Review E, 1805-1824, **62**, 2000
17. G. Newell, *Theories of instabilities in dense highway traffic*, J. Op. Soc. Japan **5**, 9 – 54 (1962).
18. M. Bando, K. Hasebe, A. Nakayama, A. Shibata, Y. Sugiyama, *Dynamical model of traffic congestion and numerical simulation*, Physical Review E, 239-253, **51**, 1999
19. D.C. Gazis, R. Herman and E.W. Montroll, *Non-linear follow-the-leader models of traffic flow*, Oper. Res. **9**, 545—560 (1961).
20. Brent C. Crowther, *A Comparison of CORSIM and INTEGRATION for the Modeling of Stationary Bottlenecks*, Master thesis, Virginia Polytechnic Institute and State University (2001).
21. W. Knospe, L. Santen, A. Schadschneider and M. Schreckenberg, *Towards a realistic microscopic description of highway traffic*, J. Phys. A **33**, L477 (2000).
22. B.S. Kerner, S.L. Klenov and D.E. Wolf, *Cellular Automata approach to the three phase theory of traffic flow*, preprint, <http://arxiv.org/abs/cond-mat/0206370>
23. See <http://www.ce.berkeley.edu/~daganzo/>
24. Karen R. Smilowitz & Carlos F. Daganzo, *Experimental Verification of Time-Dependent Accumulation Predictions in Congested Traffic*, Transportation Research Record **1710**, 85 – 95 (2000)
25. R.E. Wilson, *An analysis of Gipps' model of highway traffic*, IMA J. Appl. Math. **66**, 509 – 537 (2001.)
26. G. Newell, *A simplified car-following theory – a lower order model*, Transportation Research B **36**, 195 – 205 (2002).

LIST OF FIGURES

Figure 1. The observer positions and the segment numbering used along San Pablo Dam Road. For the simulation, the real traffic light at the end has been replaced by a virtual traffic light at observer 8.

Figure 2. The velocity plotted as function of space and time for the data set of the first day. To reduce noise, the speed data are filtered with a median filter: in a moving window of size 13, the medians of the speeds in this window have been plotted.

Figure 3. Performance of the models per each study segment.
Data and fitted parameters are from the first day.

Figure 4. Comparison of the average errors in travel time for the models used in this study.

LIST OF TABLES

Table 1. Short description of the models used in this study.

FIGURE 1

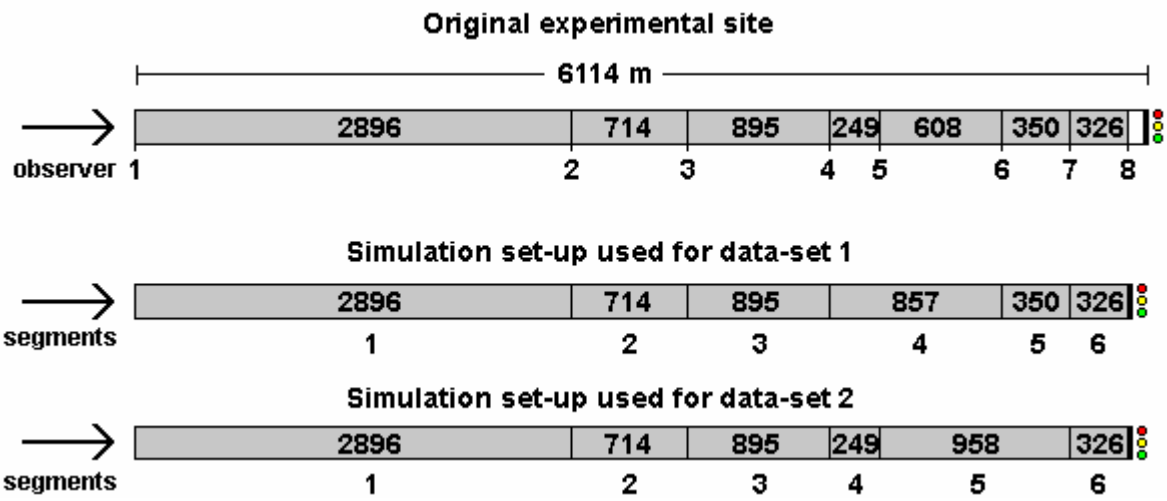


FIGURE 2

speeds v, day 1

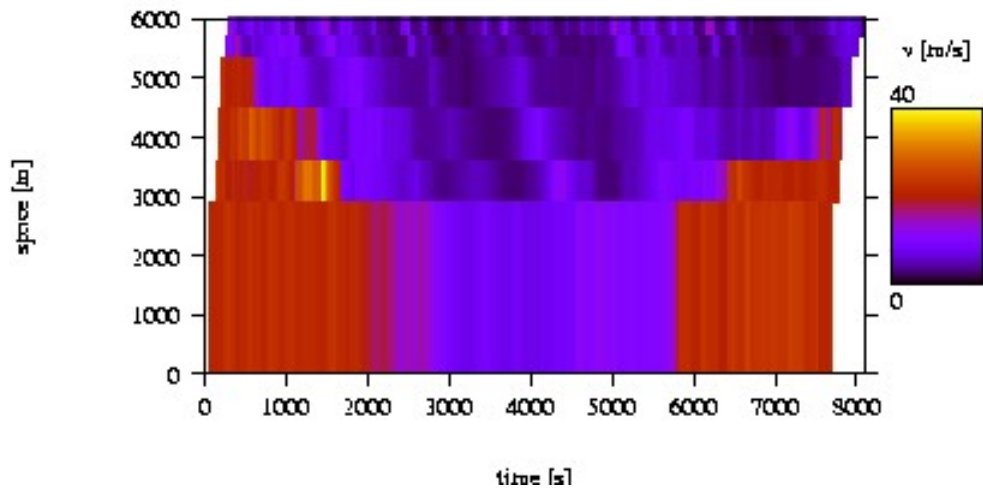


FIGURE 3

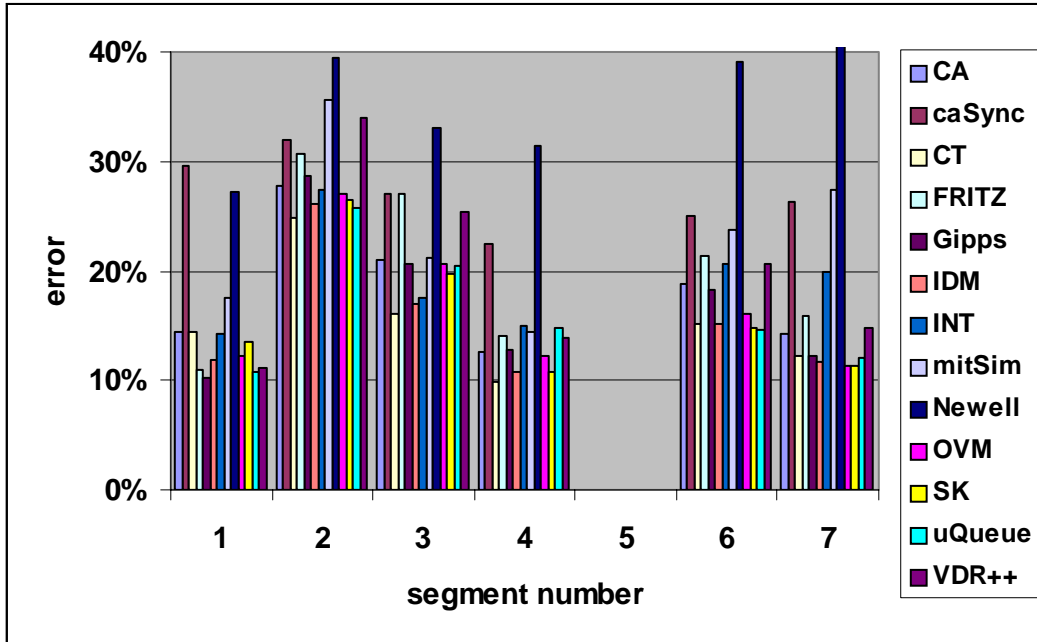


FIGURE 4

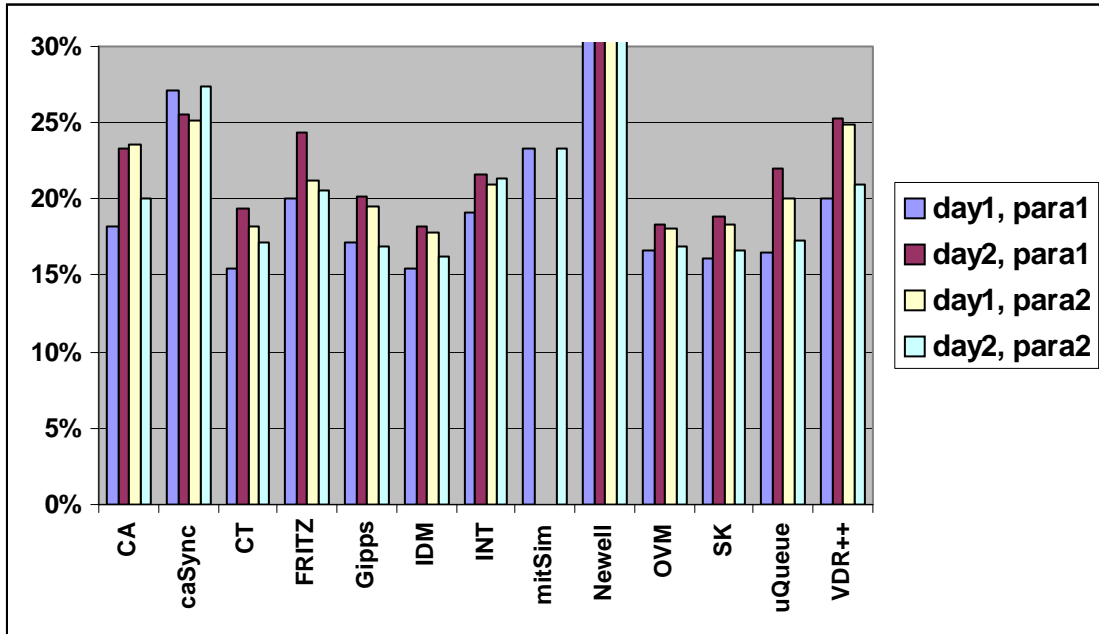


TABLE 1

Model (acronym)	Type	Equation (if not too complex)	Parameters (n)	Hidden parameters (h)	# params (n + h)
CA	Fully discrete $\lambda = 7.5m$ in the original formulation	$v' = \min\{v+1, g, v_{\max}\} - \xi$ where $\xi = \{0,1\}$ with ε	$v_{\max}, \ell, \varepsilon$	$\lambda = \ell$	2 + 1
CASync	Fully discrete, $\lambda = 0.5m$	Model has “following mode”. Cars in following mode keep their distance.	$v_{\max}, \varepsilon, \dots$	λ, ℓ	10 + 2
CT	Fluid	$\delta n(i \rightarrow i+1) = \min\{n_i, \beta(N - n_{i+1})\}$	v_{\max}, ℓ, w	λ, h	3 + 2
FRITZ	Differential equation/event-driven		v_{\max}, \dots	ℓ	12 + 1
Gipps	Time discrete	$v' = \min\{v_1, v_2, v_{\max}\}$ $v_1 = a\tau(1 - \frac{v}{v_{\max}})\sqrt{0.025 + v/v_{\max}}$ and $v_2 = -b(\tau/2 + \theta) + \sqrt{b^2(\tau/2 + \theta)^2 + b(\tilde{v}^2/\hat{b} + 2g - v\tau)}$	$v_{\max}, a, b, \tau, \theta, \hat{b}$	ℓ	6 + 1
IDM	Differential equation		$v_{\max}, a, b, \tau, \delta, g_0$	ℓ	6 + 1
INT	Differential equation	Cars try to reach the line $g = c_1 + c_3v + \frac{c_2}{v_{\max} - v}$	$v_{\max}, a, b, c_1, c_2, c_3$		6
MitSim	Differential equation		$v_{\max}, \varepsilon, \dots$	ℓ	13 + 1
Newell (26)	Differential equation	Following car drives a shifted Trajectory of lead car.	v_{\max}, a, d, τ, T		5
OVM	Differential equation	$\dot{v} = (V(g) - v)/T$ where $V(g) = v_{\max}(\tanh(a(g-s))) + f$	v_{\max}, a, s, T f is function of v_{\max}, a, s	ℓ	4 + 1
SK	Time discrete	$v' = \min\{v+a, v_1, v_{\max}\} - \varepsilon\xi$ where $\xi \in [0,1]$ and $v_1 = -b + \sqrt{b^2 + \tilde{v}^2 + 2bg}$	$v_{\max}, a, b, \varepsilon$	ℓ, τ	4 + 2
uQueue	Mesoscopic	$t_{wait} = \min\{\tau_{ff}, n_{i+1}\tau_{jj}\}$ where n_{i+1} is number of cars in next cell	$v_{\max}, \ell, \tau_{jj}, \tau_{ff}$	λ, h	4 + 2
VDR++	Fully discrete $\lambda = 1m$	Cars react to brake light of the next two cars ahead	$v_{\max}, \varepsilon, \dots$	λ, ℓ	6 + 2