

Calibration and validation of microscopic traffic flow models

Elmar Brockfeld*, Peter Wagner

Institute of Transport Research, German Aerospace Center

Rutherfordstrasse 2, 12489 Berlin, Germany

phone: +49 30 67055 231, fax: +49 30 67055 202

email: elmar.brockfeld@dlr.de

For Presentation on the 5th conference

Traffic and Granular Flow (TGF) 2003

October 2003

Delft

July 23, 2003

* corresponding author

TWO-PAGE EXTENDED ABSTRACT

Microscopic simulation models are becoming increasingly important tools in modeling transport systems. There are a large number of available models used in many countries. The most difficult stage in the development and use of such models is the calibration and validation of the microscopic sub-models describing the traffic flow, such as the car following, lane changing and gap acceptance models. This difficulty is due to the lack of suitable methods for adapting models to empirical data. The aim of this paper is to present recent progress in calibrating a number of microscopic traffic flow models. By calibrating and validating various models using the same data sets, the models are directly comparable to each other. This sets the basis for a transparent benchmarking of those models. Furthermore, the advantages and disadvantages of each model can be analyzed better to develop a more realistic behavior of the simulated vehicles.

In this work various microscopic traffic flow models have been tested from a very microscopic point of view concerning the car-following behavior and gap-acceptance. The data used for calibration and validation is from car-following experiments performed in Japan in October 2001. 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. So one gets the positions and speeds of each car in time intervals of 0.1 seconds. The experiment was repeated eight times letting the leading driver perform various driving patterns as there are constant speeds of 20, 40, 60 and 80 km/h for some time, driving in waves and emulating many accelerations/decelerations as they are typical at intersections. To minimize driver-dependent correlations between the data sets, the drivers were exchanged between the cars regularly after each experiment.

In this paper we present analyses concerning four of the experiments, namely the patterns mostly with intervals of constant speeds and wave-performing. For each of the four experiments one gets the ten trajectories of the cars in form of the DGPS-positions and speeds. From these the accelerations and distances/gaps between the cars have been calculated, which are used then for the simulation runs.

The study was done analyzing the time-development of the gaps between the cars. For the simulation setup only two cars are considered at a time. The leading car is updated as the speeds in the recorded data sets tell and the following car is updated as defined by the equations and rules of the used model, respectively. The absolute error a model produces is calculated via the simple quadratic distance between the recorded gaps and the simulated gaps. To get a percentage error it is additionally related to the mean average gap in each data set. 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 optimization method was used and started several

times with different initialization values for each “model - data set” pair. The variation in initialization 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 optimization algorithms. Subsequently, the validation was performed by determining the error of a given model on all the data sets which have not been used to calibrate the model.

By now, ten microscopic models of a very different kind using 3 to 14 parameters have been tested. The most basic parameters used by the models are the car length, a maximum velocity, an acceleration and mostly a deceleration rate. The acceleration and deceleration rates are specified in more detail in some models depending on the recent speed or traffic states (indicated by density for example). Furthermore, some models use a parameter for random braking or another kind of stochastic parameter describing individual driver behavior. Finally, few models use much more parameters to describe the driver’s behavior, which will be briefly described in the final paper. 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 automata - have been modified to adopt for an arbitrarily small time-step. So far the models tested are as follows (more will be added):

- CA (cellular automaton model by K. Nagel, M. Schreckenberg),
- SK-model (model by S. Krauss),
- OVM (“Optimal Velocity Model”, Bando, Hasebe),
- IDM (“Intelligent Driver Model”, Helbing),
- IDMM (“Intelligent Driver Model with Memory”, Helbing, Treiber),
- CATauT (CA model with more variable acceleration and deceleration, own development),
- GIPPSLIKE (basic model by P.G. Gipps),
- Aerde (model used in the simulation package INTEGRATION),
- FRITZSCHE (model used in the british software PARAMICS; it is similar to what is used in the german software VISSIM by PTV),
- MitSim (model by Yang, Koutsopoulos, used in the software MitSim).

The error rates of the models in comparison to the data sets during the calibration for each model reach from 9 to 24 %. But no model appears to be significantly the best one since every model has the same problems with distinct data sets and other data sets can be simulated quite good with each model. Interestingly, it can be stated that models with more parameters than others do not necessarily reproduce the real data better. The results of the validation process draw a similar picture. The produced errors in these cases are about 12 to 30 %, sometimes up to 40 or 60%, which is of course much bigger than in the simple calibration cases. All in all the results after the calibration agree with some results that have been obtained before. 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 12-30 % can probably not be suppressed no matter what a model is used. This would be due to the different behavior of each driver.