German Aerospace Center

Institute of Transport Research

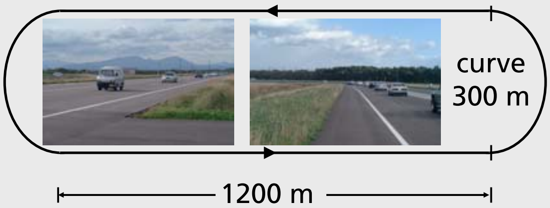
Contact persons:
Elmar Brockfeld
Peter Wagner

Calibration and Validation of microscopic traffic flow models

Ideas

» Compare microscopic traffic flow models by calibrating & validating them with the same recorded data sets
» Calibration and validation in a very microscopic way analysing the time-series of the headways

Data sets



Sketch of the test track

» Data from ten GPS-equipped cars driving on a 3km test track in Hokkaido, Japan (*1*)
» GPS positions in intervals of 0.1 second
» First car performed certain "driving patterns" (see next point) on the straight sections
» Driving patterns are: constant speeds of 20, 40, 60 and 80 km/h; driving in waves varying from about 25 to 65 km/h (half, single, and double waves on a straight 1200 meter long section)
» The data of four experiments, each with ten cars driving 15-30 minutes, are used for the analyses

Simulation set-up

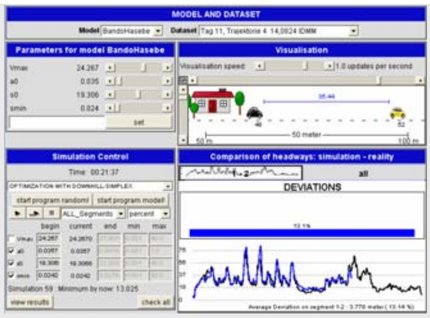
» For each simulation run one vehicle pair is considered
» The first car is updated with the speeds as recorded in the data
» The second car is updated following the rules of the traffic model under consideration
» For error measurement the percentage error e is calculated by taking the absolute differences of the observed gaps *g_obs* to the simulated gaps *g_sim* and relating it to the mean gap in each of the observed data sets, respectively, over the complete simulation time *T*.

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

» To find the optimal parameters for the models a gradient free (direct search) optimisation technique is used named downhill simplex (*2*)
» The optimisation process is started with random initialisation several times to avoid local minima.

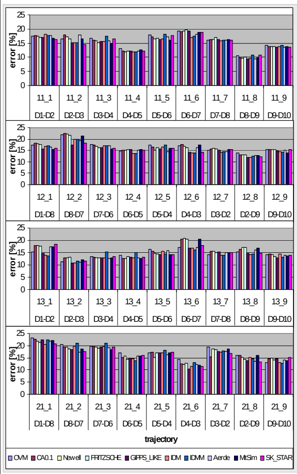
Tested Models

» 4 parameters, CA0.1 („cellular automaton model") (*3*)
» 7 p, SK_STAR (model based on the SK-model by S. Krauss) (*4*)
» 4 p, OVM ("Optimal Velocity Model") (*5*)
» 6 p, IDM ("Intelligent Driver Model") (*6*)
» 7 p, IDMM ("Intelligent Driver Model with Memory") (*7*)
» 7 p, Newell (CA-variant of the model with more variable acceleration and deceleration by G. Newell) (*8, 9*)
» 6 p, GIPPSLIKE (basic model by P.G. Gipps) (*10*)
» 6 p, Aerde (model used in the software INTEGRATION) (*11*)
» 13 p, FRITZSCHE (model used in the british software PARAMICS; similar to what is used in the german software VISSIM by PTV) (*12*)
» 15 p, MitSim (model used in the software MitSim) (*13*)



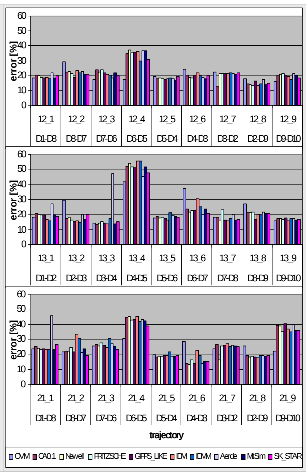
Java Applet for testing the models

Calibration results



Calibration of 10 models with 36 data sets (four experiments „11", „12", „13" and „21", each with 9 driver pairs; „D" = Driver; leading driver is always „D1"):
» Error rates from 9% to 24 %, mostly between 12% and 17 %
» No model appears to outperform the others regularly; the average error the best model produces is 15.14 %, that of the worst model 16.20 %
» Diversity in the driver behaviour (5%-15%) is bigger than diversity of the models (average differences per data set is 2.5 %)
» All models share the same problems with the same data sets
» For these data sets models with more parameters do not produce better results

Validation results



The resulting 9 parameter sets for each model of experiment „11" are validated with the other three „12", „13" and „21". (11_1 -> 12_1, 13_1, 21_1; 11_2 -> 12_2, 13_2, 21_2 ...):
» Error rates are mostly between 17% and 22%
» Additional error in comparison to calibration is about 3.2% to 5.5 % for most models
» Small problems with the models OVM (6.6% worse than calibration) and Aerde (7.6% worse) in some cases
» "Overfitting" of data set 11_4 produces high errors up to 55% for data sets 12_4, 13_4 & 21_4
» Special driver behaviour (driver pairs 11_4 and 21_9) because of exceptional bad validation results

Conclusions

» Models with more parameters must not necessarily produce better results than simpler models
» Using only few data for calibration can cause "overfitting", thus other data sets can not be reproduced well
» From this microscopic point of view the cal/val errors of 15-25% seem hard to be reduced further, no matter what model is used

References

(1) G. S. Gurusinghe, T. Nakatsuji, Y. Azuta, P. Ranjitkar, Y. Tanaboriboon, TRB2003-004137.
(2) 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).
(3) K. Nagel and M. Schreckenberg, J. Physique I, 2221, **2**, 1992.
(4) S. Krauss, P. Wagner, and C. Gawron, Phys. Rev. E **55**, 5597 – 5605 (1997).
(5) M. Bando, K. Hasebe, A. Nakayama, A. Shibata, Y. Sugiyama, Physical Review E, 239-253, **51**, 1999.
(6) M Treiber, A Hennecke and D Helbing, Physical Review E, 1805-1824, **62**, 2000.
(7) M. Treiber, D. Helbing, "*Memory effects…*", see at http://arxiv.org/abs/cond-mat/0304337.
(8) G. Newell, J. Op. Soc. Japan **5**, 9 – 54 (1962).
(9) G. Newell, Transportation Research B **36**, 195 – 205 (2002).
(10) P.G. Gipps, Transp. Res. B **15**, 105 – 111 (1981).
(11) Brent C. Crowther, "*A Comparison of CORSIM and INTEGRATION …*", Master thesis, Virginia Polytechnic Institute and State University (2001).
(12) H.-T. Fritzsche, Transp. Engin. Contr. **5**, 317 – 321 (1994).
(13) Kazi Iftekhar Ahmed, Ph.D. thesis, MIT (1999).

DLR - IVF
Rutherfordstraße 2
D-12489 Berlin
Tel: (+49)30-67055-231
(+49)30-67055-237
Fax:(+49)30-67055-202
E-Mail:
Elmar.Brockfeld@dlr.de
Peter.Wagner@dlr.de