CALIBRATING WALKER MODELS: A METHODOLOGY AND APPLICATIONS

Mario C. Campanella, Delft University of Technology, Transport & Planning Department, Stevinweg 1, 2628CN Delft, the Netherlands, m.c.campanella@tudelft.nl

Serge P. Hoogendoorn, Delft University of Technology, Transport & Planning Department, Stevinweg 1, 2628CN Delft, the Netherlands, s.p.hoogendoorn@tudelft.nl

Winnie Daamen, Delft University of Technology, Transport & Planning Department, Stevinweg 1, 2628CN Delft, the Netherlands, w.daamen@tudelft.nl

ABSTRACT

In this paper we present a generic methodology for calibrating walking models. Applying the methodology to the Nomad walker model improved substantially the quality and significance of the parameters. This methodology aims to describe all aspects of the parameter estimation process of walker models, while making it applicable to the calibration of any walker model. The base of the methodology is the simultaneous calibration of several aspects of pedestrian traffic to improve the robustness and generic application of the model. We investigated the impact of combining different flow configurations, estimating parameters for groups of pedestrians, using errors of pedestrian state variables in the optimisation and introducing prior information about the parameters.

Keywords: calibration methodology, pedestrian modelling, walker models

INTRODUCTION

Walker models are widely used for predicting purposes. Many commercial and free application tools are available attesting the demand for such predictive models. The basic premise for the use of such models is the confidence that their results are within an accepted deviation of reality. These models are applied in a variety of tasks: calculating egress times in evacuations, determining level-of-service in walking facilities, crowd control and comparing the performance of different layouts of walking facilities. This wide range of usage indicates that these models should perform well in many tasks (generality). A general model is able to represent correctly different pedestrians (heterogeneity) walking in different configurations of flow.
Each pedestrian has its own, unique, walking behaviour reflecting this inter-pedestrian heterogeneity. One way to account for heterogeneity is to use a single walker model for the total population, and to assign different parameter sets for different pedestrians. Most models would allow for this approach and some authors estimated parameter distributions (Hoogendoorn \textit{et al.}, 2005, Johansson \textit{et al.}, 2007, Moussaid \textit{et al.}, 2009) rather than average parameter values (Campanella \textit{et al.}, 2009b, Kretz \textit{et al.}, 2008, Robin \textit{et al.}, 2009, Steiner \textit{et al.}, 2007, Teknomo and Gerilla, 2005, Osaragi, 2004) to account for heterogeneity. Campanella \textit{et al.}, 2009a showed that heterogeneity in the population has a large effect in the flow characteristics. Comparing flows generated by a homogenous population with flows with heterogeneous population they observed that heterogeneous populations in general produced flows with higher breakdown probabilities (the probability that flows will get jammed). These facts indicate that a calibration of a single model should always aim to estimate how the parameters distribute in the population.

Another important aspect of calibrating walker models is the risk of specialisation. A model that is calibrated using a specific flow configuration (e.g. bidirectional flows) may not perform well in other flow configurations (e.g. one directional flows or crossing flows). Ideally, walker models should perform well in all situations, but such a generic model is not available yet.

Literature on calibration showed that no generally accepted methodology and guidelines for calibration of traffic models and in particular pedestrian models exists. The main contribution of this paper is a methodology that focuses on (but is not limited to) estimating parameters reflecting the inter-pedestrian heterogeneity. Furthermore, it also incorporates the ability of using simultaneously different flow configurations and different aspects of the pedestrian traffic in calibration. Another added value of this methodology is that it enables to compare the quality of different parameter sets.

**GENERALISED CALIBRATION METHODOLOGY**

Figure 1 shows the scheme of the calibration methodology. It is presented as an iterative process of searching optimal parameters that are significant (the outcomes of the model are influenced by their variation). The clockwise loop is the optimisation process in which parameters are input in the walker model that is used to run simulations. Errors arising from differences between the simulation and the reference data are used by the optimisation algorithm to compare sets of parameters. While the error is not minimal a new set of parameters is generated and a new iteration is performed. When the minimum error is found the parameters are consider optimal. These parameters are then analysed to assess their significance. If they are found not to be significant enough, a new calibration must be performed. Otherwise, the parameters are ready to be submitted to a validation process.
A methodological investigation on calibration of walker models
CAMPANELLA, Mario; HOOGENDOORN, Serge; DAAMEN, Winnie

The basic components of the methodology are the scenarios. They represent those aspects of the traffic that the model is going to be applied for. Pedestrian state variables such as positions, velocities or accelerations are usually the most important outcomes of walker models. Another possible aspect would be the differences between fundamental diagram relations. Any aspect that represents meaningful pedestrian behaviour can be used to build a scenario. A calibration scenario is a computable algorithm that inputs a set of values (the parameters $\theta$ necessary to simulate the walker model) and calculates a value that is an expression of how different (or how alike) the model predictions are from the reference data. A calibration scenario comprehends the reference data, all the boundary conditions necessary to run the simulations and an objective function.

The objective function is a mathematical function that takes in values that represent the aspects of the simulated traffic and compare with the same aspects in the reference data. If we are interested in comparing how different the outcomes are then the objective-function calculates the error $\varepsilon$ (that will be minimised). Several forms of objective function are used and the simplest is the difference function.

To obtain a model for generic application, several scenarios with different trajectories and/or different aspects can be optimised simultaneously. The resulting parameter set will be the best compromise between the different flow configurations and traffic aspects. Figure 1 shows the multiple-objective function $\psi$ that is responsible to calculate a unique error from $n$ scenario errors. The common form of $\psi$ is a weighted addition of the errors: $\psi = \lambda_1 \varepsilon_1 + \ldots + \lambda_n \varepsilon_n$. If one scenario must have more importance in the calibration then its $\lambda$ must be set to a larger value. One problem with using errors in multiple-objectives is that there is no way to find the values of $\lambda$ other then setting them arbitrarily.

Figure 1 – Calibration methodology scheme

12th WCTR, July 11-15, 2010 – Lisbon, Portugal
An alternative for the objective function is to determine a likelihood function that will express the probability that the model outcomes are equal to the reference data. In this case the calibration will try to maximize the likelihood. The advantage of using probabilities is that by assuming that they are independent they can be multiplied to give the combined probability. Furthermore, by calculating the logarithm of the likelihood we obtain the log-likelihood that can be added without any need of weights to express the joint probability of two different scenarios. In the next section we will present likelihood functions that were applied in the calibrations performed in this paper.

Estimation with likelihood functions for trajectory errors

The model estimation in this paper follows the calibration framework developed in (Hoogendoorn and Daamen, 2010) and in (Hoogendoorn and Hoogendoorn, 2010). In these papers the authors proposed likelihood objective functions that allowed individual and multiple trajectories to be optimised both for walker and car-following models. The authors propose extensions of the likelihood functions that allow the inclusion of information from other data sources such as prior-information on the distribution of the parameters.

The basic idea of the likelihood framework is to calculate estimation errors from a single pedestrian along his simulated trajectory and use it as the likelihood probability. To calculate these errors one pedestrian from the reference data is randomly chosen to be simulated (pedestrian \( p \)). The simulation begins with the first time-step \( (t_i) \) in which \( p \) is present in the reference data. All pedestrians \( (p \text{ included}) \) that are present in the walking area at \( t_i \) have their state variables \( z(t) = (\vec{r}(t), \vec{v}(t), \vec{a}(t)) = \text{(position, velocity, acceleration)} \) set exactly the same as in the reference trajectories. At every subsequent time-step until the time step \( t_K \) in which the pedestrian \( p \) exits all pedestrians apart of \( p \) will have their states updated according to the reference data. Pedestrian \( p \) however will have his state estimated by the model. The difference (error) between the estimated state variable and the correct value as stated in the reference data is then a measure of the likelihood that the outcomes of the model are equal to the reference data.

\[
\varepsilon(t|\theta^p) = \left\| \vec{z}^{ref}(t_i) - \vec{z}^{E}(t, \theta^p) \right\|
\]  

If we assume that these errors are independent and normally distributed with parameters \( N(0,\sigma_p) \) we can apply the probability density for the normal distribution obtaining the likelihood function for one time step. It can be shown (Hoogendoorn and Hoogendoorn, 2010) that the log-likelihood function for the errors of the entire trajectory is then (2):

\[
\hat{L}(\theta^p) = -\frac{n}{2} \ln \left( \sum_{k=1}^{K} c(t_k|\theta^p)^2 \right) - \frac{n}{2}
\]  

This equation is the log-likelihood objective-function that must be maximised by finding the optimal parameter set \( \theta^* \) (3):

\[
\theta^* = \text{argmax } \hat{L}(\theta)
\]
Estimation with prior information

Imagine that we know (from previous calibrations) the mean $\bar{\theta}_i$ and the deviation $\sigma_i$ of the parameter $i$, part of the parameter set $\theta$. If we assume that $\bar{\theta}_i$ follows a normal distribution we can obtain the likelihood with prior information:

$$L_{\text{prior}} = -\frac{n}{2} \ln \left(2\pi \sigma_i^2 \right) - \frac{\varepsilon(\theta_i, \bar{\theta}_i)}{2\sigma_i^2} + L(\theta) \quad (4)$$

CALIBRATION EXPERIMENTS

In the following sections we apply the methodology to investigate four aspects that influence the quality of the estimations. These four aspects deal with important parts of the methodology but are not the only that should be investigated (others are mentioned in the subsection future work in the conclusion). The first two aspects are elements of the methodology that affect the precision of estimated parameters.

1. We want to know how flow configurations influence the values of estimated parameters. Do unidirectional flows in which pedestrians only interact with pedestrians walking in the same direction produce significant and well estimated parameters responsible for all interaction behaviours? How do other configurations such as bidirectional and crossing flows perform as well?

2. Pedestrian states can be characterised by three variables: positions, velocities and accelerations. These states are commonly used to calculate errors in calibrations because they express dynamical properties of the walking behaviour. Therefore, we investigate what is the quality of the estimations when using errors from these variables.

The two other investigations aim at improving the significance of estimated parameters by using special objective-functions.

3. Imagine a set of trajectories in which only a subset of the pedestrians are walking close enough to obstacles for a significant period of time. This will cause the parameters responsible to the obstacle interaction not to be significant for the population. However, if we would estimate one parameter set for a group of pedestrians then we would increase the probability of having situations where pedestrians are close to obstacles. The question of how this grouping affects the estimated parameters is investigated.

4. Another way of dealing with parameters that are found to be not significant is to find another source for the mean values (and standard deviations). Having such
information we investigate how much is a new estimation improved by applying these prior information in the objective function.

**Experimental design**

To be able to assess the quality of the estimations we need to have complete knowledge about the pedestrians that created the reference trajectories. Such knowledge enables us to compare the results of the calibrations quantitatively with a ground truth. For this purpose we created synthetic trajectories with the Nomad model (Hoogendoorn and Bovy, 2003), which were to be used as reference trajectories.

Since we performed a large amount of calibrations, limiting of the amount of parameters that were to be estimated was necessary. The population was created with parameters previously estimated with real data (Campanella et al., 2009b). Furthermore, we varied six of the parameters among the population to create a heterogeneity: $A_0$ and $R_0$ are the parameters responsible for the interaction behaviour, $T$ is the parameter responsible for the pedestrians to stay along their intended path, the free speed is $V_0$, there is the pedestrian radius and a noise parameter that accounts for behaviours that are not modelled (Hoogendoorn and Bovy, 2003). These parameters were assigned according to a normal distribution $N(\mu, \sigma)$. The values of the standard deviations presented in table 1 were based on calibrations presented in Hoogendoorn et al., 2005.

Table I – Distribution of means and deviations of the parameters that were varied in the reference trajectories

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Mean $\mu$</th>
<th>Deviation $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_0$</td>
<td>10.0</td>
<td>0.7</td>
</tr>
<tr>
<td>$R_0$</td>
<td>0.16</td>
<td>0.02</td>
</tr>
<tr>
<td>$T$</td>
<td>0.25</td>
<td>0.04</td>
</tr>
<tr>
<td>$A_w$</td>
<td>20.0</td>
<td>0</td>
</tr>
<tr>
<td>$V_0$</td>
<td>1.45</td>
<td>0.26</td>
</tr>
<tr>
<td>radius</td>
<td>0.22</td>
<td>0.2</td>
</tr>
<tr>
<td>noise</td>
<td>0</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 1 shows a seventh parameter: $A_w$, that was not varied. $A_w$ is responsible for interactions with obstacles. The first four parameters in the table 1 ($A_0$, $R_0$, $A_w$ and $T$) were chosen to be estimated in the investigations. Together with the $V_0$ and the radius they are the most important in the Nomad model and social force models. They are present in all different versions of these models (Hoogendoorn et al., 2005).

Each calibration consisted of 25 estimations of the four parameters to determine the distribution of these. This number was shown to be enough to create a sample with a distribution of reference parameters within the chosen sampling errors with 95% of confidence. Each simulation estimated the parameters of one different pedestrian chosen.
randomly (one trajectory). The details of the estimation with trajectories are written in the section (Estimation with likelihood functions for trajectory errors).

Along the investigations three different flow configurations were simulated: a bidirectional corridor, a unidirectional flow with a narrow bottleneck and a 90° crossing flow. The respective dimensions of the walking areas are for the bidirectional flow: 10m x 4m, for the narrow bottleneck flow: 10m x 4m with a narrow corridor of 1m in the middle and 8m x 8m for the crossing flows (Figure 2).

![Figure 2 – The three walking configurations (from left): bidirectional corridor, unidirectional flow with a narrow bottleneck and a crossing flow.](image)

These flows were chosen because these represent the flows most frequently used in calibration and validation procedures. Furthermore, they present a wide variety of traffic situations so that they create a large amount of behaviours that need to be properly covered by the model. In these flows pedestrians need to avoid opposing pedestrians, to follow or to overtake leading pedestrians, to interact with pedestrians coming from the sides and to deal with conflicts near bottlenecks. The input flows were created in a stepwise ascending manner to assure that both free flow and congestion could occur in all flows.

The simulations were run with very small time steps (Δt=0.02s) to reduce the influence of errors arising from numeric approximations (Campanella et al., 2009c).

As an optimisation algorithm we use a procedure combining a genetic algorithm (GA) and a Simplex optimiser. Initially a GA with a population of 20 individuals is applied. The best individual after a significant number of generations is considered the closest to the optimal solution and is passed to the Simplex optimiser. The experiments have shown that this individual was always close enough to the correct parameter value to allow the Simplex to search for the optimum value. The combined procedure therefore has a much better accuracy than when using only the Simplex method. Tests with special trajectories without the noise and no distributed parameters have shown that the Simplex very often could not estimate the correct parameters. This happened because the objective function outcome is not smooth and present several local minima trapping the search results. Using the pure GA turned out to be at least two times slower in estimating the correct parameters.

**THE INFLUENCE OF FLOW CONFIGURATIONS**

In this section we investigate the influence of flow configurations for estimating correct parameters of the Nomad model. In this section we estimated all parameters using the errors
of the accelerations in the trajectories presented in equation 1. The three flows shown in figure 2 were used for calibrations. A fourth estimation named *multiple-flows* was performed adding errors arising from the three flows. The results of the estimations are presented in table 2.

Table 2 – The results per flow configuration of four parameters estimated with the methodology and the mean of the maximum log-likelihoods for the 25 calibrations.

<table>
<thead>
<tr>
<th>Parameters $\hat{\theta}$</th>
<th>Log-Likelihood $L_{\text{Mean}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
</tr>
<tr>
<td>Correct values</td>
<td>A0</td>
</tr>
<tr>
<td>Bidirectional</td>
<td>10.0</td>
</tr>
<tr>
<td>Crossing</td>
<td>8.4</td>
</tr>
<tr>
<td>Bottleneck</td>
<td>7.2</td>
</tr>
<tr>
<td>Multiple-flows</td>
<td>8.5</td>
</tr>
</tbody>
</table>

**Results for the single flow estimation**

The mean and deviations of the parameters A0 and T were better estimated using the bidirectional corridor and the crossing flows than the unidirectional flow. Furthermore, table 2 shows that the mean of the optimal log-likelihoods for these flows was significantly bigger in comparison with the bottleneck flow indicating that the trajectories were better estimated. However, the unidirectional bottleneck flow permitted a better estimation of AW. For the bidirectional corridor this parameter was tested and shown to be significant for only 24% of the estimations and 0% for the crossing flows against 72% of the unidirectional bottleneck. This happened because the flow configurations presented insufficient situations in which pedestrians were close to obstacles.

**Combining flow configurations**

To improve this situation 25 new calibrations were performed with a multiple-objective function that simulated all three configurations simultaneously and added the errors. The table 2 also shows the results for this multiple-flows estimation. In the multiple flow calibration the parameters A0 and R0 were as well calibrated as with the crossing and bidirectional flows. T was a bit worse estimated but still much better in comparison with the bottleneck flow. This shows that the negative effect of the large errors produced by the unidirectional bottleneck flow was compensated by the smaller errors of the other two flows. The Aw mean was identical for the multiple and unidirectional bottleneck flow. However, the deviation obtained with the multiple-flow was significantly smaller. This improvement was also observed with the increase of significant estimations (92%). These results indicate that the richness of the individual flows was captured by the multiple-objective estimation.
COMPARING PEDESTRIAN STATES

In the log-likelihood presented in equation 2 any of the three state variables (or a combination of them) could be used to calculate the errors shown in the equation 1. In this investigation we compared how well the three state variables are able to estimate the distribution of the parameters. In table 3 we present the means and deviations from the estimated parameters for each state variable and a fourth error that was a simple addition of all three errors at each moment of the trajectory. All calibrations were performed with the bidirectional corridor. This configuration was chosen because it presented the best overall results apart from the multiple-flow that is more complex and thus a more computational demanding process.

Table 3 – The results per state variable of four parameters estimated with the methodology and the mean of the minimum errors for the 25 calibrations.

<table>
<thead>
<tr>
<th>Parameters ( \hat{\dot{\theta}} )</th>
<th>Log-Likelihood ( \hat{L} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>deviation</td>
</tr>
<tr>
<td>A0</td>
<td>0.16</td>
</tr>
<tr>
<td>acceleration</td>
<td>8.5</td>
</tr>
<tr>
<td>velocity</td>
<td>8.9</td>
</tr>
<tr>
<td>position</td>
<td>6.5</td>
</tr>
<tr>
<td>combined</td>
<td>8.5</td>
</tr>
</tbody>
</table>

Comparing the use of single states

The acceleration and combined errors produced significantly better results than the position and velocity. The means of the estimations using the acceleration and velocity did not differ so much but the deviations for the velocity were larger for the A0 parameter. This caused a significant difference in the maximum log-likelihood mean values indicating that the acceleration errors were significantly smaller.

The reason for the poor performance of the position and velocity errors in comparison to the acceleration can be explained by numeric errors in the integration of the movement equations. The Nomad model calculates the acceleration of the pedestrian, the velocity is obtained by one time integration and the position is obtained by two time integrations at each time step. The errors originating from the discrete integration approximations must get significant when compared with the error arising from the model causing these differences. We observed that during the calibrations with the position errors the GA did not find good solutions and the parameter sets it passed to the Simplex optimiser were too far from the optimal for it to improve further.
Combining state errors

An error adding the three individual errors was also created and showed equivalent results to the acceleration error indicating that the acceleration error is preferable.

APPLYING MULTIPLE-TRAJECTORY SCENARIOS

In this section we investigate the effect of combining multiple-trajectories to improve the significance of parameters. We have seen that using the bidirectional flow data to estimate the parameters of individual trajectories gave poor results for the Aw parameter due to insufficient information in the data. To improve this situation we modified the single trajectory optimisation expressed by equation (2) by including errors of several trajectories. We investigated the effects of adding the errors of different combinations of 10 and 20 trajectories to the outcome of the parameters.

Table 4 shows that combining the trajectories of pedestrians in the estimation did make the distributions of the T get much worse, A0 did not change much with the means for the combined slightly worse but the deviations better, R0 was estimated similarly but Aw improved considerably. The significance of Aw was also much higher: 92% of significant estimations for 10 trajectories and 96% for 20 trajectories. This indicates that aggregating the trajectories improved the estimation when data that is not rich for all pedestrians. The two levels of aggregation did not present significantly different results

Table 4 – The results per aggregation level of four parameters estimated with the methodology and the mean of the minimum errors for 25 calibrations.

<table>
<thead>
<tr>
<th>Parameters $\hat{\theta}$</th>
<th>Log-Likelihood $\hat{L}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
</tr>
<tr>
<td>A0</td>
<td>R0</td>
</tr>
<tr>
<td>Correct values</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>8.5</td>
</tr>
<tr>
<td>10</td>
<td>7.4</td>
</tr>
<tr>
<td>20</td>
<td>7.1</td>
</tr>
</tbody>
</table>

USING PRIOR INFORMATION OF PARAMETERS

In this section we investigated how the inclusion of prior information in the objective-function affects the quality of the predictions. We recalibrated the bidirectional flow the acceleration error, but this time we included the information gathered in the multiple-trajectories about the mean and deviation of the Aw parameter. We used the values $Aw = 18$ and $s = 5.8$ and put them in the equation (4). The estimation of A0, R0 did not improve and T got significantly...
A methodological investigation on calibration of walker models
CAMPANELLA, Mario; HOOGENDOORN, Serge; DAAMEN, Winnie

worse as table 5 shows. However, the objective with prior estimated Aw much better indicating that the proposed method of including prior information can be used to calibrate a model with poor data.

Table 5 – The results with prior information of four parameters estimated with the methodology and the mean of the minimum errors for the 25 calibrations.

<table>
<thead>
<tr>
<th>Parameters $\hat{\theta}$</th>
<th>Log-Likelihood $L$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
</tr>
<tr>
<td>A0</td>
<td>R0</td>
</tr>
<tr>
<td>Correct values</td>
<td>10</td>
</tr>
<tr>
<td>bidirectional</td>
<td>8.5</td>
</tr>
<tr>
<td>bidi. with prior</td>
<td>8.0</td>
</tr>
</tbody>
</table>

CONCLUSION

In this paper we presented a generalised calibration methodology composed of calibration scenarios that represent aspects of pedestrian traffic. By defining calibration components these can easily be investigated and modified when parameters are not significant. An application of the methodology on Nomad showed that unidirectional flows produce less optimal parameters. Furthermore, non significant parameters can be estimated by combining the results of several flow configurations in multiple-objective estimations. Using acceleration errors proved to produce more optimal parameters when compared with velocity and position errors. These results are valid for models such as the Nomad and social force that predict the acceleration. For CA and other type of models that predict the velocities similar results between velocities and positions are expected. Aggregating several trajectories in the objective function improved the significance of those parameters that were not very satisfactory. The presented methodology greatly improves the confidence in the results of pedestrian walker models. Models calibrated according to it can be made more precise, general and robust.

Future work

More aspects influencing calibration of walker models can be investigated such as the effect of noise and upper-body lateral movements (swaying) in the trajectories. The tracking process that produces reference data from empiric observations may introduce measurement errors in the position of the pedestrians that can be alleviated with a smoothing algorithm. However, smoothing may introduce errors. The effect of smoothing in data will be investigated in the future using empiric data.

The methodology applied to real trajectories can also be used to study the heterogeneity and variations of behaviours in different flow configurations.

12th WCTR, July 11-15, 2010 – Lisbon, Portugal
A methodological investigation on calibration of walker models

CAMPANELLA, Mario; HOOGENDOORN, Serge; DAAMEN, Winnie

REFERENCES


12th WCTR, July 11-15, 2010 – Lisbon, Portugal