18th International Conference on the Application of Computer
Science and Mathematics in Architecture and Civil Engineering
K. Gürlebeck and C. Könke (eds.)
Weimar, Germany, 07–09 July 2009

PARAMETER IDENTIFICATION OF MESOSCALE MODELS FROM MACROSCOPIC TESTS USING BAYESIAN NEURAL NETWORKS

Jörg F. Unger* **and Carsten Könke***

**Institute of Structural Mechanics*
Bauhaus University Weimar
Marienstrasse 15
99423 Weimar, Germany
E -mail: joerg.unger@uni-weimar.de

Keywords: Multiscale, Parameter identification, Bayesian neural networks.

Abstract. *In this paper, a parameter identification procedure using Bayesian neural networks is proposed. Based on a training set of numerical simulations, where the material parameters are simulated in a predefined range using Latin Hypercube sampling, a Bayesian neural network, which has been extended to describe the noise of multiple outputs using a full covariance matrix, is trained to approximate the inverse relation from the experiment (displacements, forces etc.) to the material parameters. The method offers not only the possibility to determine the parameters itself, but also the accuracy of the estimate and the correlation between these parameters. As a result, a set of experiments can be designed to calibrate a numerical model.*

1 INTRODUCTION

Multiscale simulations offer the possibility to simulate phenomena present on lower scales. In order to obtain realistic simulations, the numerical models have to be calibrated. Often, there are only macroscopic experiments available to determine these parameters. Unfortunately, it is often difficult to fit the parameters of the numerical model manually, because, due to the interaction between these parameters, a modification of a single parameter will influence the full response. Consequently, an automated procedure is required.

Although especially suited for multiscale simulations, where often the material models and the experimental validations are on different scales, the methodology presented here is generally applicable to any parameter fitting procedure of numerical models using experimental data. The procedure is illustrate for the parameter identification of a mesoscale model for concrete.

2 BAYESIAN NEURAL NETWORKS

The multilayer perceptron as one kind of neural networks is a metamodel consisting of an input layer, an output layer and a certain number of hidden layers. The number of neurons in the input/output layer corresponds to the dimension of the input/output vector. The free parameters $\boldsymbol{w}$ of the network, which are the weights and biases, are calculated from given associations $\mathcal{D} = \left\{\boldsymbol{x}^{(i)}, \boldsymbol{t}^{(i)}\right\}$, which are called training samples.

Within the Bayesian framework of neural network approximation [1], the output of the neural network $\boldsymbol{y}$ given the input $\boldsymbol{x}$ is assumed to be superposed by Gaussian noise:

$$p(\boldsymbol{t}|\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\Sigma}_\beta) = \mathcal{N}(\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{w}), \boldsymbol{\Sigma}_\beta), \tag{1}$$

where $\mathcal{N}(\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{w}), \boldsymbol{\Sigma}_\beta)$ is a standard joint Gaussian distribution with mean value $\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{w})$ and covariance $\boldsymbol{\Sigma}_\beta$. It is to be noted that, in contrast to standard Bayesian neural networks, the authors use a full covariance matrix $\boldsymbol{\Sigma}_\beta$, with a separate noise variance for each output and their correlation.

In a similar way, the prior distribution of the free parameters (weights and biases), which corresponds to the distribution of the free parameters without any knowledge of the training data $\mathcal{D} = (\boldsymbol{X}, \boldsymbol{t})$, is assumed to be

$$p(w_j|\alpha_j) = \mathcal{N}(w_j|0, \alpha_j^{-1}). \tag{2}$$

Theoretically, a precision parameter α_j can be defined for all free parameters. However, the network possesses a symmetry with respect to the neurons (e.g. if neuron 1 and 2 in the hidden layer with their associated weights are exchanged, the input-output relation will remain unchanged). In order to account for this symmetry property, the free parameters are placed into groups with a single precision parameter for each group.

Assuming for now that the hyperparameters α_i and the noise covariance $\boldsymbol{\Sigma}_\beta$ are known, and using the assumption that the samples are independent and identically distributed (i.i.d.), it can be shown [2] that

$$\begin{aligned}\ln p(\boldsymbol{w}|\mathcal{D}, \boldsymbol{\alpha}, \boldsymbol{\Sigma}_\beta) &= \ln p(\mathcal{D}|\boldsymbol{w}, \boldsymbol{\Sigma}_\beta) + \ln p(\boldsymbol{w}|\boldsymbol{\alpha}) + \text{const} &(3)\\ &= -\frac{1}{2}\left[\sum_{m=1}^{M}\left(\boldsymbol{t}^{(m)} - \boldsymbol{y}^{(m)}\right)^T \boldsymbol{\Sigma}_\beta^{-1}\left(\boldsymbol{t}^{(m)} - \boldsymbol{y}^{(m)}\right) + \sum_{j=1}^{N^w}\alpha_j w_j^2\right] + \text{const.} &(4)\end{aligned}$$

N_o is the number of output parameters, N_w is the number of free parameters and M the number of training samples. The terms independent of $\boldsymbol{w}$ are summed up in the constant term, since they have no influence in the optimization procedure. In this paper, a preconditioned conjugate gradient approach has been used to solve this optimization problem.

The predictive distribution of the output $\boldsymbol{t}$ for a new input vector $\boldsymbol{x}$ is obtained by marginalizing over the free parameters and performing a Taylor series expansion of the output. It can be approximated by a normal distribution

$$p(\boldsymbol{t}|\boldsymbol{x}, \mathcal{D}, \boldsymbol{\alpha}, \boldsymbol{\Sigma}_\beta) \approx \mathcal{N}\left(\boldsymbol{t}|\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{w}_{MAP}), \boldsymbol{\Sigma}(\boldsymbol{x})\right) \tag{5}$$

$$\boldsymbol{\Sigma}(\boldsymbol{x}) = \boldsymbol{\Sigma}_\beta + \boldsymbol{J}(\boldsymbol{x})^T \boldsymbol{A}^{-1} \boldsymbol{J}(\boldsymbol{x}) \tag{6}$$

where $\boldsymbol{J}$ is the sensitivity of the outputs with respect to the free parameters and $\boldsymbol{A}$ is the hessian of Eq. (4) at $\boldsymbol{w}_{MAP}$, which is the solution of the optimization problem in Eq. (4).

The hyperparameters α_c and the noise covariance $\boldsymbol{\Sigma}_\beta$ can be calculated by integration over the weigths and maximizing the probability that the training data has been generated by the model. Within a fixed point iteration scheme, these parameters are updated according to

$$\alpha_c = \frac{\gamma_c}{\sum_{j:\alpha_j \epsilon C} w_j^2} \qquad \text{with } \gamma_c = k_c - \alpha_c \sum_{j:\alpha_j \epsilon C} \boldsymbol{A}^{-1}_{(jj)} \tag{7}$$

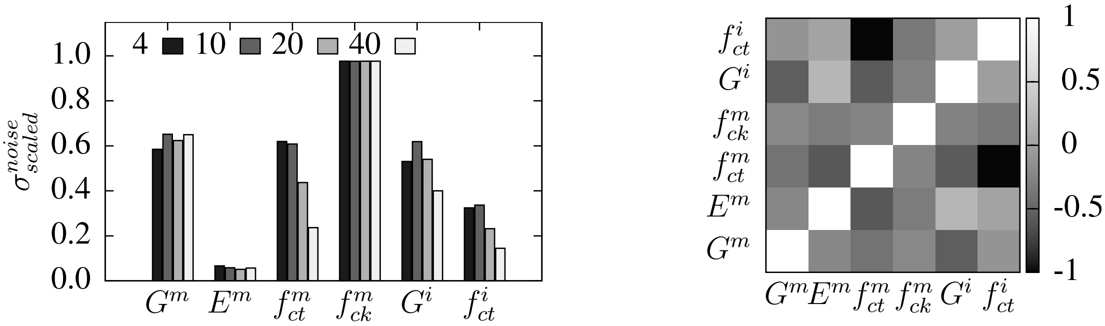
$$\boldsymbol{\Sigma}_\beta = \frac{1}{M} \left(\sum_{m=1}^{M} \left(\Delta^{(m)} \Delta^{(m)^T} \right) + \frac{\partial \ln |\boldsymbol{A}|}{\partial \boldsymbol{\Sigma}_\beta^{-1}} \right), \tag{8}$$

where k_c is the number of parameter in the same group c. After each update, the optimization problem in Eq. (4) has to be solved until convergence of both substeps.

3 DESIGN OF EXPERIMENTS AND PARAMETER IDENTIFICATION

The general approach for building numerical models can be divided into three phases. At first, the engineer builds a numerical model with a set of free parameters - e.g. constitutive parameters. Afterwards, simple experimental tests are used in order to calibrate the numerical model. Finally, the model can be used to predict the response of more complicated experiments. The difficulty within this procedure is the problem to determine the set of experiments that is required in order to calibrate a numerical model, especially, if the material parameters do not have a clear physical interpretation or their influence on the response is correlated.

In this context, a new procedure for parameter identification and design of experiments for the calibration of numerical models is proposed. The procedure can be outlined as follows. First, a candidate for the required calibration test is designed and a parameterized numerical model is build. From engineering experience, the admissible parameter space is defined and a latin hypercube sampling procedure for the set of parameters is performed using a uniform distribution and uncorrelated parameters. For each sample of material parameters, the response is calculated from the numerical model, and the Bayesian neural network is trained using the responses as inputs and the material parameters as output parameters. Finally, the uncertainty of the parameter estimates and their correlation can be estimated from the training data. If the quality of the parameter estimate is not sufficient, the procedure has to be restarted with a modified set of experiments or a different model, where e.g. some parameters are fixed or the constitutive formulation is changed. If the most probable set of material parameters for a



(a) Influence of number of the number of inputs

(b) Correlation matrix of noise

Figure 1: Scaled noise variances and correlation for the identified parameters

given experimental test is to be determined, the discretized input of the experimental curve is used as input of the Bayesian neural network, whose output is then an estimates of the material parameters including information related to their accuracy and correlation.

4 EXAMPLE

In the present example, the parameter identification of a mesoscale model for concrete is presented. In the numerical model, concrete is resolved on the mesoscale distinguishing between particle, mortar matrix and the interfacial transition zone (ITZ). The particles are assumed to be linear elastic, the ITZ is simulated using a mixed cohesive zone formulation, and a combined damage plasticity model is applied for the mortar matrix. A tensile test is used to exemplarily demonstrate the determination of material parameters - for the matrix (Young's modulus E^m, fracture energy G_f^m, tensile and compressive strength f_{ct}^m, f_c^m) and for the ITZ (tensile strength f_c^i and fracture energy G_f^i).

In Fig. 1(a), the scaled noise variance for each material parameter is plotted, which describes the accuracy of the parameter estimate. A value of 1 implies that the parameter cannot be identified at all, and a value of zero means that the parameter can be estimated almost exactly. It is observed that, as expected, the compressive strength cannot be identified using a tensile test. This implies that an additional test, e.g. a compression test, has to be included. It is furthermore observed that the Young's modulus of the matrix can almost exactly be determined. This can be explained by the fact that this parameter is the only one that influences the linear elastic part of the load-displacement curve. Consequently, there is a straightforward relation between load-displacement curve and material parameter. The other parameters simultaneously influence the peak and the post-peak branch of the response. As a consequence, their individual contributions can only be efficiently extracted when using more input points on the load-displacement curve.

From Fig. 1(b) it is additionally observed that the tensile strength of the matrix and the interface are strongly correlated with a correlation coefficient of almost -1. This implies that, when increasing one tensile strength, the other one has to be decreased on order to obtain the same response. One possible solution is to fix the ratio between the two parameters and use only a single parameter in the identification procedure.

5 CONCLUSION

In this paper, a parameter identification procedure based on Bayesian neural networks is developed that allows to obtain information about the accuracy of the parameter estimates.

Furthermore, the procedure allows to determine the experimental test required to calibrate a numerical model without the requirement to actually perform experiments.

REFERENCES

[1] D. J. C. Mackay. Probable Networks and Plausible Predictions — A Review of Practical Bayesian Methods for Supervised Neural Networks. *Network: Computation in Neural Systems*, **6**, 469–505, (1995).

[2] J.F. Unger. *Neural networks in a multiscale approach for concrete*, Bauhaus University Weimar, (2009).