České vysoké učení technické v Praze
Fakulta stavební

Artificial Neural Networks in Calibration of Nonlinear Models

Autor:

Tomáš Mareš

Vedoucí práce:

Ing. Anička Kučerová, Ph.D.

26. dubna 2012

Abstract

Last decades witness rapid development in numerical modelling of structures as well as materials and the complexity of models increases quickly together with their computational demands. Despite the growing performance of modern computers and clusters, a suitable approximation of an exhaustive simulation has still many applications in engineering problems. For example, the field of parameters identification may represent a large domain for very efficient applications. The layered neural networks are still considered as very general tools for approximation and they became popular especially for their simple implementation. This contribution presents different strategies for application of neural networks in calibration of nonlinear models and discusses their possible advantages and drawbacks.

1 INTRODUCTION

Development in numerical modelling provides the possibility to describe a lot of complex phenomena in material or structural behaviour. The resulting models are, however, often highly nonlinear and defined by many parameters, which have to be estimated so as to properly describe the investigated system and its behaviour. The aim of the model calibration is thus to rediscover unknown parameters knowing the experimentally obtained response of a system to the given excitations. The principal difficulty of model calibration is related to the fact that while the numerical model of an experiment represents a well-defined mapping from input (model, material, structural, or other) parameters to output (structural response), there is no guarantee that the inverse relation even exists.

The most broadly used approach to parameter identification is usually done by means of an error minimisation technique, where the distance between parametrised model predictions and observed data is minimised [23]. Since the inverse relation (mapping of model outputs to its inputs) is often ill-posed, the error minimisation technique leads to a difficult optimisation problem, which is highly nonlinear and multi-modal. Therefore, the choice of an appropriate identification strategy is not trivial.

Another approach intensively developed during the last decade is based on Bayesian updating of uncertainty in parameters' description [16, 15]. The uncertainty in observations is expressed by corresponding probability distribution and employed for estimation of the so-called posterior probabilistic description of identified parameters together with the prior expert knowledge about the parameter values [11, 24]. The unknown parameters are thus modelled as random variables originally endowed with prior expert-based probability density functions which are then updated using the observations to the posterior density functions. While the error minimisation techniques lead to a single point estimate of parameters' value, the result of Bayesian inference is a probability distribution that summarizes all available information about the parameters. Another very important advantage of Bayesian inference consists in treating the inverse problem as a well-posed problem in an expanded stochastic space.

Despite the progress in uncertainty quantification methods [17, 20], more information provided by Bayesian inference is generally related to more time-consuming computations. In many situations, the single point estimate approach remains the only feasible one and development of efficient tools suitable for this strategy is still an actual topic. Within the several last decades, a lot of attention was paid to the so-called intelligent methods of information processing and among them especially to soft computing methods such as artificial neural networks (ANNs), evolutionary strategies or fuzzy systems [9]. A review of soft computing methods for parameter identification can be found e.g. in [14]. In this paper, we focus on applications of ANNs in the single point approach to parameter identification.

2 ARTIFICIAL NEURAL NETWORK

Artificial neural networks (ANNs) [4, 5] are powerful computational systems consisting of many simple processing elements - so-called neurons - connected together to perform tasks analogously to biological brains. Their main feature is the ability to change their behaviour based on an external information that flows through the ANN during the learning (training) phase.

A particular type of ANN is the so-called feedforward neural network, which consists of neurons organized into layers where outputs from one layer are used as inputs into the following layer, see figure 1. There are no cycles or loops in the network, no feed-back connections. Most frequently used example

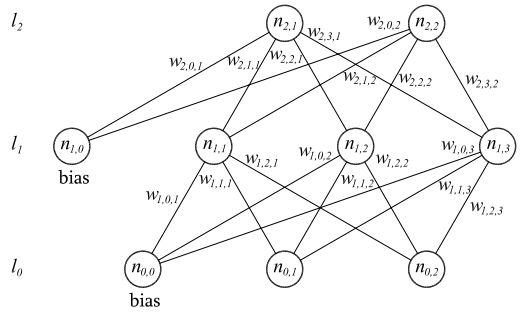

Figure 1: Architecture of multi-layer perceptron

is the multi-layer perceptron (MLP) with a sigmoid transfer function and a gradient descent method of training called the back-propagation learning algorithm. In practical usage, the MLPs are known for their ability to approximate non-linear relations and therefore, when speaking about an ANN, the MLP is considered in the following text.

The input layer represents a vector of input parameters which are directly the outputs of the input layer. The outputs of one layer are multiplied by a vector of constants, the so-called synaptic weights, summarized and used as inputs into the following layer. Elements in the hidden and output layers - neurons - are defined by an activation function, which is applied on the input and produces the output value of the neuron. The synaptic weights are parameters of an ANN to be determined during the training process. The type of the activation function is usually chosen in accordance with the type of a function to be approximated. In the case of continuous problems, the sigmoid activation function is the most common choice.

One bias neuron is also added into the input and hidden layers. It does not contain an activation function, but only a constant value. Its role is to enable to shift the value of a sum over the outputs of his neighbouring neurons before this sum enters as the input into the neurons in the following layer. The value of biases is determined by training process together with the synaptic weights.

Despite of ANN's popularity there are only few recommendations for the choice of ANN's architecture. The authors, e.g. in [8, 7]], show that the ANN with any of a wide variety of continuous nonlinear hidden-layer activation functions and one hidden layer with an arbitrarily large number of units suffices for the "universal approximation" property. Therefore, we limit our numerical experiments to such case. The number of units in the input and the output layer is usually given by the studied problem itself, but there is no theory yet specifying the number of units in the hidden layer.

To overcome this problem, we use two sets of data for ANN's preparation: training data are used for calibration of the synaptic weights of the ANN with a chosen number of hidden units and the resulting ANN is then evaluated on independent validation data. Then, one hidden neuron is added to the existing ANN, which is again trained, evaluated on validation data and the ratio between the obtained error to the error obtained for the previous ANN is computed. We count the situations, where the ratio is higher than 0.99. When these situations occur three times, the addition of hidden neurons is stopped. Then the ANN with the smallest error on validation data is employed for model calibration.

3 STRATEGIES FOR MODEL CALIBRATION

In overall, there are two main philosophies for application of ANN in identification problems. In a forward mode/direction, the ANN is applied to approximate the model response. The error minimisation technique then becomes a minimisation of distance between the ANN's predictions and experimental data. The efficiency of this strategy relies on the evaluation of the trained ANN to be significantly much faster than the full model simulation. The advantage of this strategy is that the ANN is used to approximate a known mapping which certainly exists and is well-posed. Computational costs of this strategy are separated in two parts of a similar size: (i) the ANN training - optimisation of synaptic weights and (ii) the minimisation of error in ANN prediction for experimental data - optimisation of ANN

inputs (i.e. determination of investigated model parameters). The latter part concerns optimisation of an error function which is often multi-modal, non-differentiable and some robust optimisation method has to be applied to solve this problem. An important shortcoming of this method is that this ill-posed optimisation problem needs to be solved repeatedly for any new experimental measurement. This way of ANN application to the parameter identification was presented e.g. in [1], where an ANN is used for predicting load-deflection curves and the conjugate directions algorithm is then applied for optimisation of ductile damage and fracture parameters. Authors in [19] train an ANN to approximate the results of FE simulations of jet-grouted columns and optimise the column radius and a cement content of the columns by a genetic algorithm. Principally same methods are used for identification of elasto-plastic parameters in [2].

The second philosophy, an inverse mode, assumes the existence of an inverse relationship between the outputs and the inputs of the calibrated model. If such a relationship exists at least on a specified domain of parameters' value, it can be approximated by an ANN. Then the retrieval of desired inputs is a matter of seconds and could be easily executed repeatedly for any new experiment and no other optimisation process is necessary. Here the ANN training represents the whole computational costs and a solution of the ill-posed problem. This way of ANN application to parameter identification was presented e.g. in [18] for identification of mechanical material parameters, in [27] for estimation of elastic modulus of the interface tissue on dental implants surfaces, in [28] for identification of interfacial heat transfer coefficient or in [13] for determination of geometrical parameters of circular arches.

In computational mechanics, there is often a disproportion between the number of inputs and outputs of a numerical model. While the model has usually only several parameters (inputs), their response is mostly described by a load-deflection curve, stress or strain fields. In other words, the response is usually a quantity defined in discredised spatial, time and/or pseudo-time domain. Generally, these domains can be easily parametrised. In the forward mode of identification, the problem of many outputs can be handled e.g. by including the domain parameters among the ANN inputs and thus reducing the number of outputs. An approximation of the complete model response is then obtained by repeated evaluation of the ANN with varying values of domain parameters. In the inverse mode, the situation is more difficult, because usage of high number of inputs leads to excessive complexity of the ANN architecture and the training process. Fortunately, particular components of a model response are usually highly correlated and thus the principal component analysis (PCA) [12] can be easily applied to transform them into a smaller number of uncorrelated quantities ordered according to their variance (i.e. significance). Then only a selected number of most important principal components can be used as ANN's inputs.

Since the ANN training needs a preparation of a set of training data, it is also worthy to use these data for a sampling-based sensitivity analysis [6, 21] and obtain some information about importance of particular observations or significance of each parameter for a system behaviour. To achieve some reliable information from sensitivity analysis as well as a good approximation by ANN, one has to choose the training data carefully according to a suitable design of experiments, see e.g. [10] for a competitive comparison of several experimental designs.

4 NUMERICAL RESULTS

In this section we would like to present applications of the ANN to an identification of affinity hydration model parameters as an illustration to the forward and the inverse mode of a model calibration. We will compare the results achieved by the both approaches on fully independent simulated testing data and validate them together with the direct optimisation method on experimental measurements. The affinity hydration model is very simple with a fast evaluation and the ANN application is not necessary for its parameters identification, but we use it here for an illustrative purpose. Multi-objective identification of the model parameters can be found in [25].

4.1 Affinity hydration model

The affinity model provides a simple framework describing all stages of cement hydration. The rate of hydration can be expressed by the temperature-independent normalized chemical affinity $\tilde{A}(\alpha)$ [3]

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \tilde{A}(\alpha) \exp\left(-\frac{E_a}{RT}\right) , \tag{1}$$

where α stands for the degree of hydration, T is an arbitrary constant temperature of hydration, R is the universal gas constant (8.314 $\mathrm{Jmol^{-1}K^{-1}}$) and E_α is the apparent activation energy.

For the hydration heat prediction, an analytical form presented in [26] is used:

$$\tilde{A}(\alpha) = B_1 \left(\frac{B_2}{\alpha_\infty} + \alpha\right) (\alpha_\infty - \alpha) \exp\left(\bar{\eta}\frac{\alpha}{\alpha_\infty}\right) , \tag{2}$$

where B_1 and B_2 are coefficients related to chemical composition, α_∞ is the ultimate hydration degree and $\bar{\eta}$ represents microdiffusion of free water through formed hydrates. Then a curve of the degree of hydration development α can be obtained by the numerical integration of equation (2).

4.2 Data preparation and sensitivity analysis

Since the bounds for model parameters vary in orders, one can employ the expert knowledge about the parameter meanings and before preparation of training data transform them into standardised parameters $p_i \in [0; 1]$. The bounds for affinity model parameters together with relations to the standardised

Table 1: Bounds for affinity model parameters.

Parameter	Minimum	Maximum	Relation
B_1	10^6	10^7	$p_1 = \log B_1 - 6$
B_2	10^{-6}	10^{-3}	$p_2 = (\log B_2 + 6)/3$
$\bar{\eta}$	-12	-2	$p_3 = (-\bar{\eta} - 2)/10$
α_∞	0.7	1.0	$p_4 = (\alpha_\infty - 0.7)/0.3$

parameters p_i are given in table 1.

In the space of standardised parameters we prepare a design of experiments having 100 samples based on Latin Hypercube Sampling optimised with respect to modified L_2 discrepancy. In [10] it is shown that such an experimental design has a good space-filling property and is nearly orthogonal. For each design point we perform a model simulation to obtain a bundle of 100 curves for the degree of hydration $\alpha(t)$, see figure **??**.

Since the model response is represented by the degree of hydration being a function of the time, the time domain is discretised into 199 steps uniformly distributed with the logarithm of the time. Hence, the model input vector consists of 4 parameters and the output vector consists of 199 components. For each input-output pair with 100 simulations we evaluate the Spearman's rank correlation coefficient ρ in order to compute the sensitivity of output to input parameters [6]. The results of such a sampling-based sensitivity analysis are plotted in figure **??**.

In the inverse mode, the model output vector consisting of 199 components is too large for usage as an input vector for the ANN. Hence, we performed the principal component analysis in order to reduce this number to 100 components with non-zero variance (this number is related to 100 simulations involved in PCA) and only six components with relative variance higher then 0.5 %, see figure **??**. Resulting principal components are technically new quantities obtained by a linear combination of original model outputs. This transformation has of course an influence to sensitivity analysis and thus we computed correlations between the model inputs and principal components, see figure **??**. Results of the described 100 simulations are also used as training data for ANNs. Then, the last preparatory step concerns the generation of validation data for evaluation of particular ANN's architecture (i.e. number of hidden nodes)

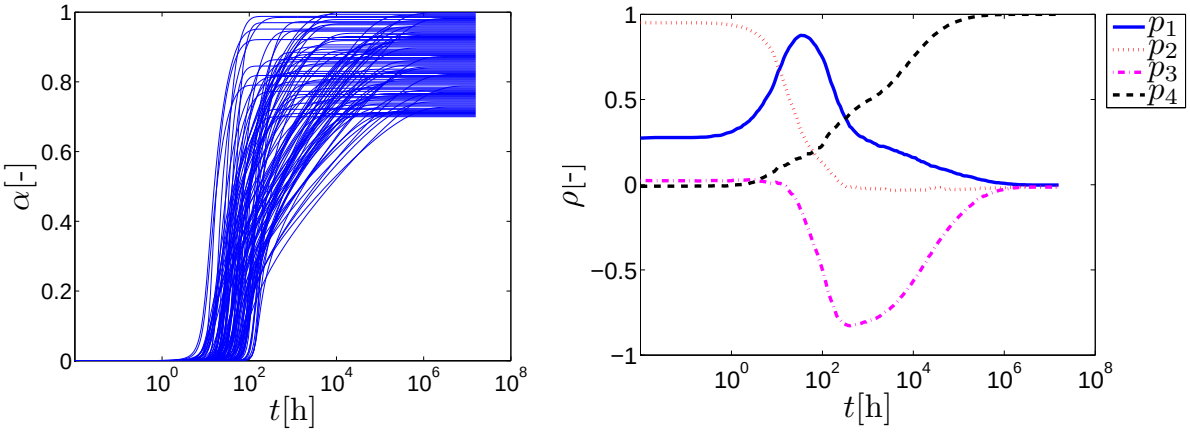

Figure 2: Model output with plotted parameter sensitivities.

and testing data for final assessment of resulting ANNs. For each of these data sets we run 50 simulations for randomly generated sets of input parameters.

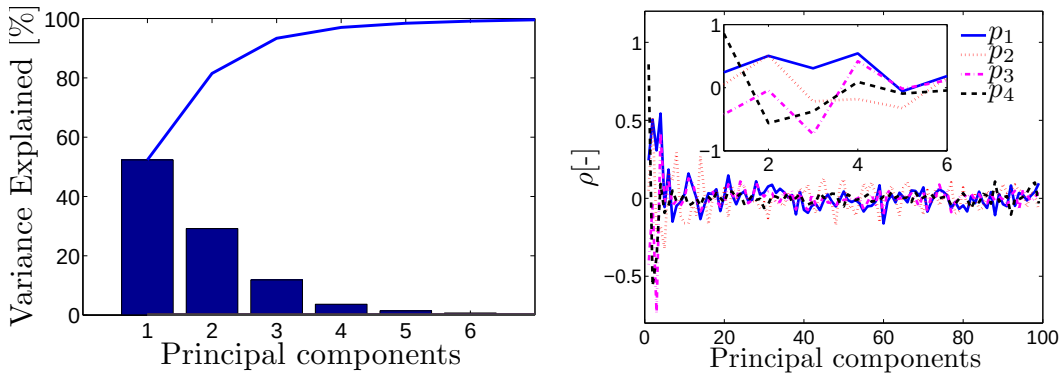

Figure 3: model output principal components and responding parameter sensitivities.

4.3 Neural network training

Prepared training and validation data were used for development of ANN models. In forward mode, the aim was to approximate the curve of hydration degree discretised in 199 time steps. Instead of developing an ANN with 199 outputs we add the value of the time as an additional input of ANN which then have only one output corresponding to the degree of hydration at a given time step. In the inverse mode, only selected principal components were used as ANN inputs and one ANN was trained independently for each parameter p_i for increasing the simplicity of the searched relationship. The ANNs were trained for two different choices of inputs. In the first one, six first principal components were used as inputs for all ANNs in the inverse mode, while in the second one, particular principal components were selected according to their sensitivity to the predicted parameter. The particular choice of ANN inputs, outputs and number of hidden neurons achieving best results on validation data are presented in table 2.

The conjugate gradient-based method [22] was employed as a training algorithm for the ANNs. The synaptic weights optimisation process was stopped either when the number of iterations achieved 5000 or if the ratio of the average error on training data during last 100 iterations to the error obtained for previous 100 iterations was higher than 0.999.

In order to evaluate the quality of particular ANNs, the relative errors ε are computed for training,

Table 2: Architecture of particular ANNs and their errors on training, validation and testing data

ANN	Input	Hidden	Output	$\varepsilon_{\text{training}}[\%]$	$\varepsilon_{\text{validation}}[\%]$	$\varepsilon_{\text{testing}}[\%]$
Forward	p_1,p_2,p_3,p_4,t	6	α	2.34	2.37	2.24
Inverse	$1^{\text{st}} - 6^{\text{th}}$ principal component	7	p_1	3.82	3.75	5.08
	$1^{\text{st}} - 6^{\text{th}}$ principal component	2	p_2	4.54	3.34	4.70
	$1^{\text{st}} - 6^{\text{th}}$ principal component	5	p_3	1.32	1.88	1.72
	$1^{\text{st}} - 6^{\text{th}}$ principal component	7	p_4	0.20	0.23	0.29
Inverse2	$1^{\text{st}} - 4^{\text{th}}$ principal component	6	p_1	3.46	3.53	5.56
with	$2^{\text{nd}} - 5^{\text{th}}$ principal component	5	p_2	6.33	4.60	4.87
selected	$1^{\text{st}}, 3^{\text{rd}}, 4^{\text{th}}$ principal component	5	p_3	4.91	3.92	5.73
inputs	$1^{\text{st}} - 3^{\text{rd}}$ principal component	4	p_4	2.03	1.76	2.35

Table 3: Results of identification procedure in relative errors on independent set of model data ε [%]

	p_1	p_2	p_3	p_4	α
Forward	7.39	14.83	3.61	1.86	0.52
Inverse	5.08	4.70	1.72	0.29	0.47
Inverse2	5.56	4.87	5.73	2.35	1.26

validation and testing data according to

$$\varepsilon = \frac{\sum_{i=1}^{I} |O_i - T_i|}{I(T_{\max} - T_{\min})}\,, \tag{3}$$

where O_i are the ANN outputs, T_i are the target values (i.e. α in case of the forward mode and p_i in case of inverse mode), I is the number of samples in a given data set (training, validation or testing) and $T_{\max}$, $T_{\min}$ are the maximal and minimal target values in training data set, respectively. Note that in forward mode, the number of data in training data set is actually 100×199, because each discrete point of the model response is considered as one training sample. Similarly, the validation and testing data sets consist of 50×199 samples. The resulting errors on training, validation and testing data for all the created ANNs are listed in table 2.

4.4 Verification of model calibration

The errors in table 2 represent the quality of constructed ANNs. Next we have to compare the quality of identification procedures. While in the inverse mode the ANNs predict directly the values of model parameters, in the forward mode we have to perform an optimisation of the parameters minimising the distance between the ANN prediction and target model response:

$$\sum_{m=1}^{M} (O(t_m) - \alpha(t_m))^2\,, \tag{4}$$

where $M = 199$ is the number of model response components. The optimisation process is governed by the GRADE evolutionary algorithm, see [14] for details about this method[1].

The optimisation process was performed for all training, validation and testing data and the relative errors ε according to equation (3) for parameter predictions were then computed for all the identification modes. The obtained results are listed in table 3.

[1]The parameters of GRADE algorithm were set to pool_rate = 4, radioactivity = 0.33 and cross_limit = 0.1. The algorithm was stopped after 10000 cost function evaluations.

To assess the quality of identification procedure in terms of model response, the model simulations were performed for all sets of identified parameters and the relative error ε was computed between the obtained responses and the original target responses. These results are also written in table 3. One can see that the forward mode of identification leads to worse errors in parameters prediction, but to smaller errors in terms of model response in comparison with the inverse mode with selected ANN inputs. The best results were achieved by the inverse mode with six principal components as ANNs inputs. It is also worthy to mention that the approximation of model response in forward mode leads to worse errors than the final errors in identified model responses (compare the results in tables 2 and 3 for forward mode). It can be possibly explained by the fact that the results in table 2 correspond to the ANN approximation of model response, while the errors in table 3 are obtained for identified parameters but using again the exact model simulation. So in other words, the latter errors cover only the error in parameters but not in model simulation itself.

In order to assess the distribution of errors in prediction of hydration degree in specified time steps, we computed the absolute distance between the response prediction $\tilde{\alpha}_n(t)$ and target response $\alpha_n(t)$

$$\varepsilon_{\alpha,n}(t) = |\tilde{\alpha}_n(t) - \alpha_n(t)| \tag{5}$$

at each time step t for all simulations n in the training, validation and testing sets. Then the minimal, mean and maximal distances were found for each time step and plotted in figure 5.

4.5 Validation of model calibration

The previous section was focused on mutual comparison of the presented identification strategies on simulated data. However, a complete comparison has to include their validation on experimental data. To that purpose we used the curve of hydration degree obtained by isothermal calorimetry for a cement "Mokra" CEM I 42.5 R taken directly from Heidelberg cement group's kiln in Mokrá, Czech Republic [26].

In general, validation does not allow for a comparison in terms of parameters value, because these are not known a priori. Nevertheless, the simplicity and the fast simulation of affinity hydration model permit a direct optimisation of model parameters without any incorporated approximation and the resulting optimal parameter values can be compared with the results obtained using the ANN approximations.

In direct optimisation, we have to formulate the cost function specifying how good is the model prediction. In our numerical study, we decided to test two widely used cost functions:

$$F_1 = \sum_{m=1}^{M} (\alpha(t_m) - \alpha^{\text{Mokra}}(t_m))^2 , \tag{6}$$

$$F_2 = \sum_{m=1}^{M} |\alpha(t_m) - \alpha^{\text{Mokra}}(t_m)| , \tag{7}$$

where $M = 199$ stands again for number of discretised values of hydration degree, $\alpha(t_m)$ is the model response and $\alpha^{\text{Mokra}}(t_m)$ are the interpolated experimental data. We applied again the GRADE algorithm with the same setting as in the previous section to minimise the both cost functions (6) and (7). The obtained parameter and cost functions values are written in table 4 and the resulting degree of hydration curves can be compared with experimental data in figure 4.

Subsequently, the forward and both inverse modes of identification were applied to the experimental data using the prepared ANNs. The identified parameters are again written in table 4 and corresponding simulated degrees of hydration are plotted in figure **??**.

The results show that the inverse modes of identification on experimental data completely failed. The reason is that ANNs are good in approximation and interpolation, but very bad in extrapolation. Due to this, the inverse approach is very sensitive to the noise and measurement errors in experimental data. Measurement errors can easily lead to ANN's input values which are out of the convex hull constructed on the training examples and the ANN is then supposed to extrapolate. This situation exactly happened here. While the experimental curve seems to be inside the range of the training curves (see figure **??**), the obtained principal components are beyond the convex hull computed for principal components of training samples. This is probably caused by the errors in degree of hydration at early stages, where it achieves

Table 4: Parameter values identified on experimental data obtained for "Mokra" cement together with the corresponding values of cost functions. The methods Direct1 and Direct2 correspond to the optimisation of cost functions F_1 and F_2, respectively.

Method	p_1	p_2	p_3	p_4	F_1	F_2
Direct1	0.856	1.000	0.208	0.053	0.002	0.336
Direct2	0.858	1.000	0.208	0.050	0.002	0.271
Forward	0.922	0.911	0.339	0.051	0.135	2.333
Inverse	1.292	0.788	0.041	-0.045	2.256	11.097
Inverse2	1.033	0.863	0.756	-0.089	3.156	17.203

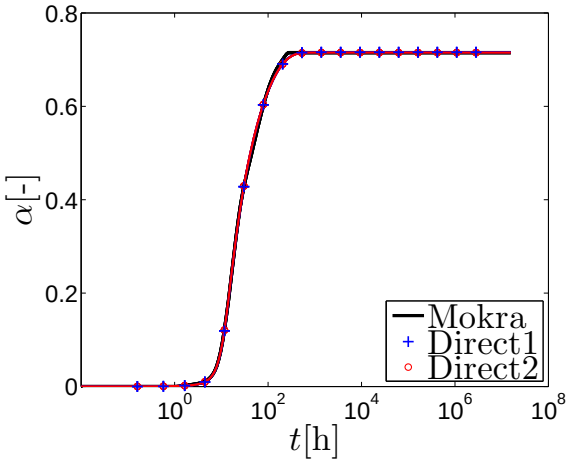

Figure 4: Comparison of experimentally obtained degree of hydration for "Mokra" cement with simulations for directly optimised model parameters.

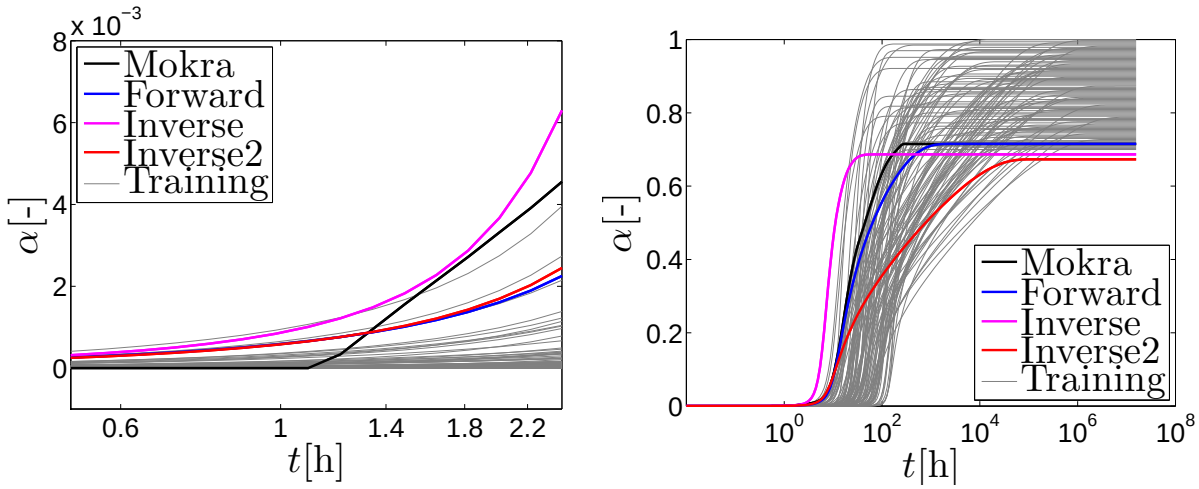

Figure 5: Comparison of experimentally obtained degree of hydration for "Mokra" cement with simulations for model parameters identified using prepared ANNs.

values close to zero in simulations, but values equal to zero in experimental data (see figure **??**). Moreover,

after this starting constant period, the experimental degree of hydration starts grow very quickly and very soon exceeds the values of all the training simulations. The direct optimisations also lead to limit value of p_2 parameter, which was not exceeded, because the optimisation constrained.

In such a situation, the forward mode is more efficient. The training data are prepared to well describe the whole domain of model parameters and hence, any vector of parameters generated inside this domain is automatically inside the convex hull constructed on training sets of model parameters. Therefore, in forward mode of identification the ANN never has to extrapolate and during the optimisation process it simply tries to find a solution which is as close as possible to experimental data. In this way it gives us at least some useful information.

5 CONCLUSIONS

The presented paper tries to review possible applications of artificial neural networks in calibration of numerical models. The basic identification modes are described in details: the forward and the inverse. Their advantages and drawbacks are illustrated on calibration of affinity hydration model. Verification of the procedures is performed using 50 independent testing data obtained from numerical simulations and experimental data obtained for cement "Mokra" CEM I 42.5 R are subsequently employed for the validation of the presented methods. The main advantage of the inverse approach is an easy application to new measurements. The computationally consuming and difficult part concerns the ANN development which has to be done only once. An important shortcoming of this procedure is related to its high sensitivity to experimental noise, measurement error or data lying far from the training simulations. The forward mode is on the other hand able to deal with these situations relatively well, but its principal drawbacks involve a discomfort when employed repeatedly for new experimental data. Each calibration includes an optimisation process which does not have to be trivial.

6 ACKNOWLEDGEMENT

The financial supports of this work by the Czech Science Foundation (project No. 105/11/P370) and the Ministry of Industry and Trade of the Czech Republic (project No. FR-TI1/612) are gratefully acknowledged.

References

[1] M. Abendroth and M. Kuna. Identification of ductile damage and fracture parameters from the small punch test using neural networks. Engineering Fracture Mechanics, 73:710–725, 2006.

[2] H. Aguir, H. BelHadjSalah, and R. Hambli. Parameter identification of an elasto-plastic behaviour using artificial neural networks–genetic algorithm method. Materials and Design, 32:48–53, 2011.

[3] D. Gawin, F. Pesavento, and B. Schrefler. Hygro-thermo-chemo-mechanical modelling of concrete at early ages and beyond. part i: Hydration and hygro-thermal phenomena. International Journal for Numerical Methods in Engineering, 67(3):299–331, 2006.

[4] K. N. Gurney. An introduction to neural networks. UCL Press, London, 2002.

[5] S. S. Haykin. Neural networks and learning machines. Prentice Hall/Pearson, New York, 3rd edition, 2009.

[6] J. Helton, J. Johnson, C. Sallaberry, and C. Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliability Engineering & System Safety, 91:1175–1209, 2006.

[7] K. Hornik. Some new results on neural network approximation. Neural Networks, 6:1069–1072, 1993.

[8] K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. Neural Networks, 2:359–366, 1989.

[9] Jyh-Shing Roger Jang and Chuen-Tsai Sun. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1997.

[10] Eliška Janouchová and Anna Kučerová. Competitive comparison of optimal designs of experiments for sampling-based sensitivity analysis. Computers & Structures, 2011.

[11] E. T. Jaynes. Probability Theory, The Logic of Science. Cambridge University Press, Cambridge, 2003.

[12] I. T. Jolliffe. Principal Component Analysis. Springer–Verlag, 2^{nd} edition, 2002.

[13] M. Klos and Z. Waszczyszyn. Modal analysis and modified cascade neural networks in identification of geometrical parameters of circular arches. Computers and Structures, 89:581–589, 2011.

[14] A. Kučerová. Identification of nonlinear mechanical model parameters based on softcomputing methods. PhD thesis, Ecole Normale Supérieure de Cachan, Laboratoire de Mécanique et Technologie, 2007.

[15] A. Kučerová, J. Sýkora, B. Rosić, and H. G. Matthies. Acceleration of uncertainty updating in the description of transport processes in heterogeneous materials. Journal of Computational and Applied Mathematics, 2012. In press, e-print: arXiv:1201.0942.

[16] Y.M. Marzouk, H.N. Najm, and L. Rahn. Stochastic spectral methods for efficient Bayesian solution of inverse problems. Journal of Computational Physics, 224(2):560–586, 2007.

[17] H. G. Matthies. Encyclopedia of Computational Mechanics, chapter Uncertainty Quantification with Stochastic Finite Elements. John Wiley & Sons, Ltd., 2007.

[18] D. Novák and D. Lehký. ANN inverse analysis based on stochastic small-sample training set simulation. Engineering Applications of Artificial Intelligence, 19:731–740, 2006.

[19] B. Pichler, R. Lackner, and H. A. Mang. Back analysis of model parameters in geotechnical engineering by means of soft computing. International journal for numerical methods in engineering, 57:1943–1978, 2003.

[20] B. Rosić, A. Litvinenko, O. Pajonk, and H. G. Matthies. Direct bayesian update of polynomial chaos representations. Journal of Computational Physics, 2011. submitted for publication.

[21] A. Saltelli, K. Chan, and E. M. Scott. Sensitivity analysis. NY:Wiley, New York, 2000.

[22] J. R. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain. Technical report, Pittsburgh, PA, USA, 1994.

[23] G.E. Stavroulakis, G. Bolzon, Z. Waszczyszyn, and L. Ziemianski. 3.13 - inverse analysis. In Editors in Chief: I. Milne, R. O. Ritchie, , and B. Karihaloo, editors, Comprehensive Structural Integrity, pages 685 – 718. Pergamon, Oxford, 2003.

[24] Albert Tarantola. Inverse Problem Theory and Methods for Model Parameter Estimation. Society for Industrial and Applied Mathematics, 2005.

[25] Z. Vitingerová. Evolutionary algorithms for multi-objective parameter estimation. PhD thesis, Czech Technical University in Prague, 2010.

[26] V. Šmilauer. Multiscale Modeling of Hydrating Concrete. Saxe-Coburg Publications, Stirling, 2012.

[27] K. Zaw, G. R. Liu, B. Deng, and K. B. C. Tan. Rapid identification of elastic modulus of the interface tissue on dental implants surfaces using reduced-basis method and a neural network. Journal of Biomechanics, 42:634–641, 2009.

[28] L. Zhang, L. Li, H. Ju, and B. Zhu. Inverse identification of interfacial heat transfer coefficient between the casting and metal mold using neural network. Energy Conversion and Management, 51:1898–1904, 2010.