Computer Assisted Mechanics and Engineering Sciences, 14: 183–196, 2007.

Identification of characteristic length of microstructure for second order continuum multiscale model by Bayesian neural networks

Łukasz Kaczmarczyk[1,2], Zenon Waszczyszyn[2]
[1] *Civil Engineering Department, University of Glasgow*
Rankine Building, Oakfield Avenue, Glasgow G12 8LT, UK
[2] *Institute for Computational Civil Engineering, Cracow University of Technology*
Warszawska 24, 31-155 Kraków, Poland

(Received in the final form April 23, 2007)

This paper deals with the second-order CH of a heterogeneous material undergoing small displacements. Typically, in this approach an RVE of a heterogeneous material is investigated. A given discretized microstructure is determined *a priori*, without focusing on details of specific discretization techniques. Application of BNN as a tool for identification of characteristic length of a microstructure is discussed. An indentation test was analyzed under plane strain constraints for generating pseudo-experimental patterns by means of FEM. A single input of BNN was formulated due to the application of PCA. The BNN of structure 1-16-1 with sigmoid hidden neurons was designed. The Bayesian inference approach was applied to obtain pdf of the characteristic length. Numerical efficiency of the proposed approach is demonstrated in the paper.

Keywords and acronyms: micro- and macrolevels, second order continuum, computational homogenization (CH), representative volume element (RVE), finite element method (FEM), Bayesian neural network (BNN), probability density function (pdf), principal component analysis (PCA), indentation test

1. Introduction

New industrial structures are built using highly heterogeneous materials. Typical examples of such materials are metal alloys, porous media, polycrystalline materials and composites. The size, shape, physical properties and spatial distribution of microstructural constituents determine the macroscopic, overall behaviour of these multi-phase materials.

The characteristic length of a microstructure has to be determined to take into account higher order effects, e.g. size effect. The characteristic length scale is an additional macroscopic parameter of the constitutive model and cannot be derived by any homogenization method, cf. [6].

In the present paper an indentation test is analyzed, cf. [3], where the characteristic length of heterogeneous material L is significant. The computational homogenization method (CH) is used to determine overall properties of equivalent material [8]. The Bayesian neural network (BNN) model is formulated, cf. [2], for the identification probability density function (pdf) of a characteristic length size of the microstructure. The usefulness of neural networks for inverse problems has been investigated by many authors, cf. references in [14].

In Section 2.1 a short description of computational homogenization method is given. Next, a description of the principal components analysis (PCA) and BNNs is presented. In Section 3 a numerical example of indentation test of a heterogeneous material is discussed and the advantages of this approach are demonstrated. In the end some final conclusions are expressed.

2. Short description of the applied methods

2.1. Computational homogenization method

In the paper the well-known framework of linking material properties at two levels of description is presented. The material is assumed to be heterogeneous in the microscale level, and homogeneous in the macroscale level of observation. There are a number of strategies which are used in the multiscale analysis. In the present paper a numerical approach, i.e. CH [8], is applied, see Fig. 1. This micro-macro model does not lead to a closed form of overall constitutive equations. Instead, the model determines the stress-strain relation at every point of interest at macroscale by detailed modelling of microstructure attached to that point.

Following [4], the multiscale models are constructed using three main ingredients (Fig. 2):

1. Modelling of mechanical behaviour at microscale, i.e. a RVE is introduced.

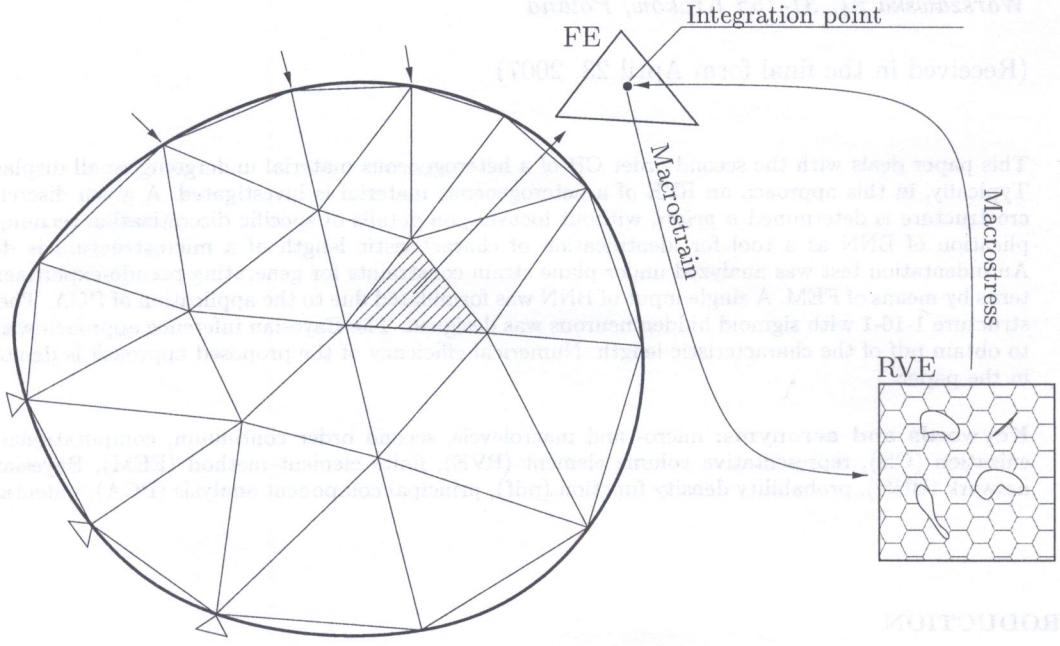


Fig. 1. Computational homogenisation

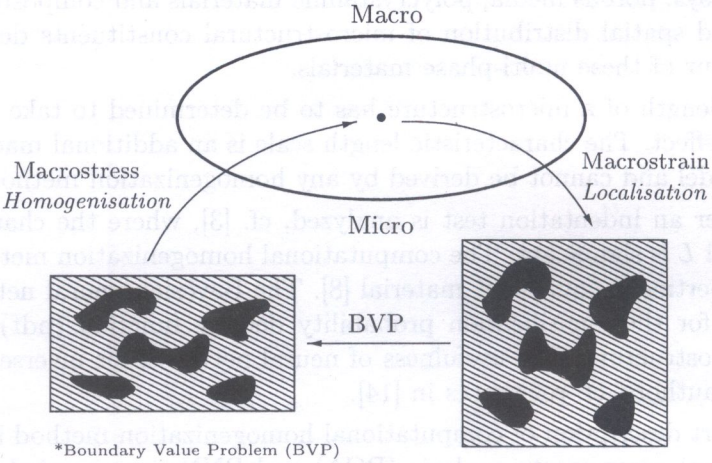


Fig. 2. Computing heterogeneous material response

2. A downscaling rule which determines the local solution inside the RVE, for given macroscopic deformation measures.

3. An upscaling rule giving the macroscopic stress measures, knowing the micromechanical stress state.

More detailed description of CH can be found in [7, 8].

2.2. Principal Component Analysis

The principal component analysis (PCA) is a statistical technique to simplify random data sets [5]. PCA is a linear transformation that transforms the data to a new coordinate system such that the projection of the data on the first axis of a new coordinate system (called the first principal component) has the greatest variance, the second greatest variance on the second axis, etc. PCA is used for dimensionality reduction of the input dataset.

In the considered case of identification of microstructure characteristic length L the data are related to points of the position vector $\mathbf{x}$ at an equilibrium path. Measured components of the input vector $\mathbf{x}$ are scaled by means of formula $\hat{a} = (a - E[a])/E[a]$. The correlation matrix is then defined,

$$\mathbf{R}_{xx} \approx \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}_n \mathbf{x}_n^{\mathrm{T}} = \frac{1}{N}\mathbf{X}\mathbf{X}^{\mathrm{T}}, \tag{1}$$

where N is the number of experiments in the matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$.

Then, the eigenpairs of the correlation matrix $(\lambda_i, \mathbf{v}_i)$ are computed and put in order according to $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$. The principal component vectors are defined as

$$\mathbf{Y}_i = \mathbf{v}_i^{\mathrm{T}} \mathbf{X}. \tag{2}$$

If the first PC vectors $\mathbf{Y}_i$ corresponding to $i = 1, \ldots, K$ for $K < N$ and $\lambda_K \gg \lambda_{K+1}$ are preserved then the input space, spread on $\mathbf{Y}_i$, can be reduced to $\mathbb{R}_K$, cf. [5].

2.3. Bayesian approach

Bayesian inference has been successfully applied to neural networks (NNs) for the analysis of not only classification but also regression problems [10]. This technique is based on the Bayes' theorem which relates posterior and prior probabilities of stochastic events, i.e. $P(A \mid B)$ and $P(B)$. Bayes' theorem plays the key role of continuous probability density distribution $p(B \mid A)$, called pdf for short, as probability density function. In the case of BNN, pdf takes the form

$$p(\mathbf{w} \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \mathbf{w}) p(\mathbf{w})}{p(\mathcal{D})} \tag{3}$$

where: $\mathbf{w} \in \mathbb{R}^W$ — weight vector, $\mathcal{D}$ — output data set. In the considered application of BNN the output set is composed of characteristic length of RVE, i.e. $\mathcal{D} = \{\hat{L}_1, \hat{L}_2, \ldots, \hat{L}_N\}$.

The denominator of Eq. (3),

$$p(\mathcal{D}) = \int_{\mathbb{R}^W} p(\mathcal{D} \mid \mathbf{w})\, p(\mathbf{w})\, \mathrm{d}\mathbf{w}, \tag{4}$$

plays the role of normalization factor. It can be computed by means of the marginalization principle [1, 5], i.e. by integration in the parameter space $\mathbb{R}^W$ over all the weight vector components w_i for $i = 1, \ldots, W$.

The application of Bayes' inference technique requires the above mentioned integration over weights, which needs the marginalization principle of as a new paradigm of BNN. This changes the

optimization paradigm of formulation of classical neural networks, mostly applied in engineering applications, cf. [5]. The predicted pdf of a target output value t_* is computed according to the formula

$$p(t_* \mid \mathbf{X}, \mathcal{D}) = \int_{(\mathbb{R}^W, \alpha, \beta)} p(t_* \mid \mathbf{w}, \beta)\, p(\mathbf{w}, \alpha, \beta \mid \mathbf{t})\, \mathrm{d}\mathbf{w}\mathrm{d}\alpha\mathrm{d}\beta, \tag{5}$$

where: α, β — hyperparameters of an assumed pdf of the random weight and target vectors, i.e. $\mathbf{w}(\alpha)$ and $\mathbf{t}(\beta)$, correspondingly.

Prediction of posterior distribution (5) needs the computation of integral over posterior pdf, conditioned on data sets $\{\mathbf{X}, \mathcal{D}\}$, and assumed pdf $p(\mathbf{w} \mid \alpha)$ and $p(\mathbf{t} \mid \beta)$. The integration is analytically intractable and needs special numerical methods. From among many methods listed in [9], the approach basing on Markov chains and hybrid Monte Carlo Method, cf. [1, 11], is frequently applied to reduce the cost of numerical integration.

Emphasizing the major advantage of the Bayesian approach, a statistical model is built which gives predictions of micro-parameters in terms of the posterior weights distribution. In case of identification of characteristic length,

$$\langle \hat{L} \rangle = \int \mathrm{BNN}(y, \mathbf{w})\, P(\mathbf{w} \mid \mathcal{D})\, \mathrm{d}\mathbf{w}. \tag{6}$$

$\langle \hat{L} \rangle$ are expected parameters, $\mathrm{BNN}(y, \mathbf{w})$ is the neural network model and $P(\mathbf{w} \mid \mathcal{D})$ is the posterior distribution of weights conditioned on the training data $\mathcal{D}$ (in what follows output data are only written since input data $\mathbf{x}$ can be omitted in Eq. (5)).

3. Numerical example

3.1. Multiscale finite element discretization of indentation test

Indentation experiments, which are used to determine the mechanical properties of small volumes of a material, are difficult to interpret at a microscopic level. The response of the system is very complicated due to large stresses generated in the vicinity of the indentor, which can create defects or cracks, or cause a phase transformation to occur in the material.

Here the qualitative analysis of a metal matrix composite (aluminium matrix with silicon inclusions) subject to indentation is discussed, see Fig. 3. Attention is focused on the size effect in because of a weak scale separation in the vicinity of the indentor application. The indentor is idealized as a rigid body while the target material is modelled as a deformable body. To simplify the problem, symmetry conditions are recognised and plane strain conditions subject to small displacements and strains are considered. The indentor is assumed to be wedge shaped with an angle of 15° ($\alpha = 75°$) and the contact between the indentor and the target material is assumed to be frictionless (effect of friction is considered negligible for indentors with angles larger than 12°, cf. [3]). At the macrolevel, QU34L4, finite elements [6, 12] are used since strain gradient effects will be captured. The finite element mesh for the macroscale is shown in Fig. 4. The element size for the finest part of the mesh is 0.0025 mm. The size of each element in the coarse domain is 0.16 mm.

The characteristic size of the RVE varies as $L = 0.001$, 0.002, and 0.004 mm, while the size of the inclusion relative to the size of the RVE is kept constant at $0.4L$. Discretization of the RVE at the microscale, shown in Fig. 5, is undertaken using four-nodes quadrilateral finite elements with reduced integration and hourglass control. For stabilization of the tangent stiffness matrix, the Galerkin/least-squares approach was used [6, 13]. The RVE is subject to displacement type boundary constraints. The matrix (aluminium) is assumed to be an elasto-plastic hardening material, whereas the inclusion (silicon) is modelled as a stiffer elastic material (see Fig. 5).

In Fig. 6 a typical deformation of the macroscopic mesh is shown, where h is indentation depth and s measures the contact surface. The distribution of macrostresses is shown in Fig. 7. Utilising

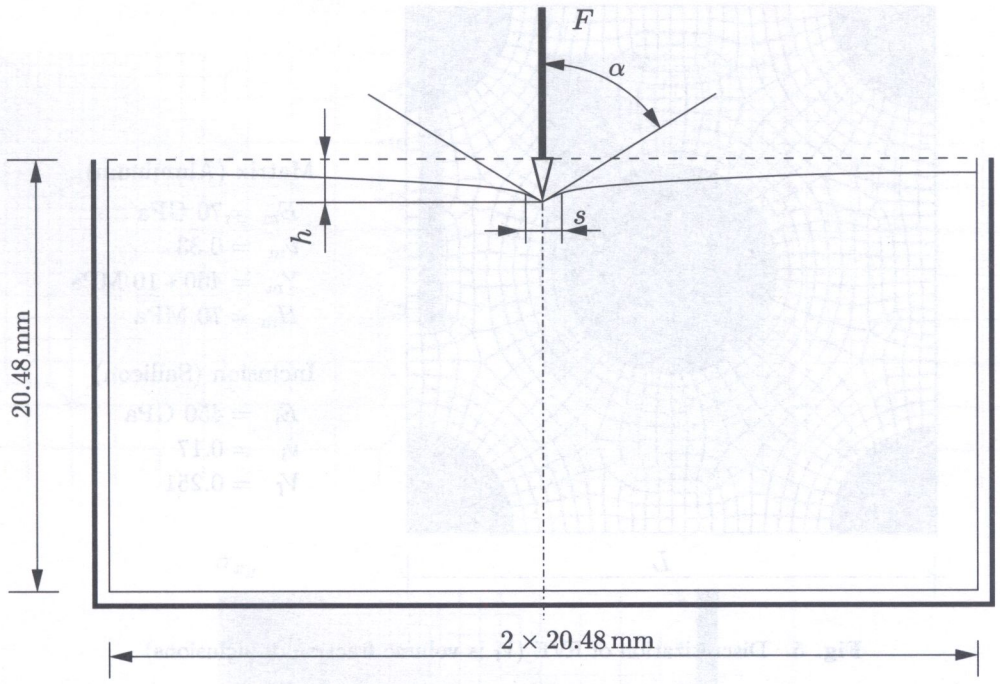


Fig. 3. Indentation test

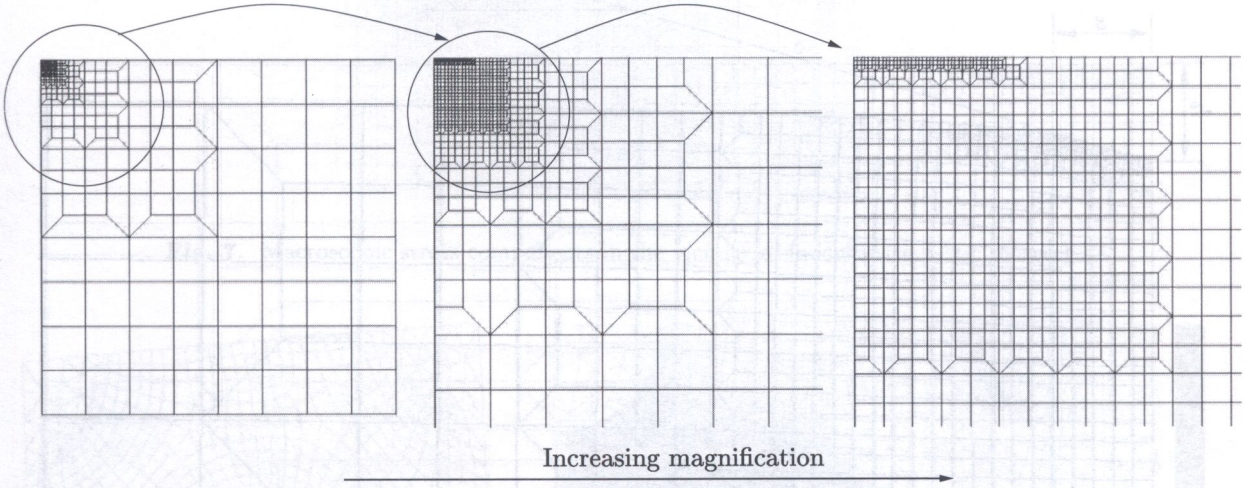
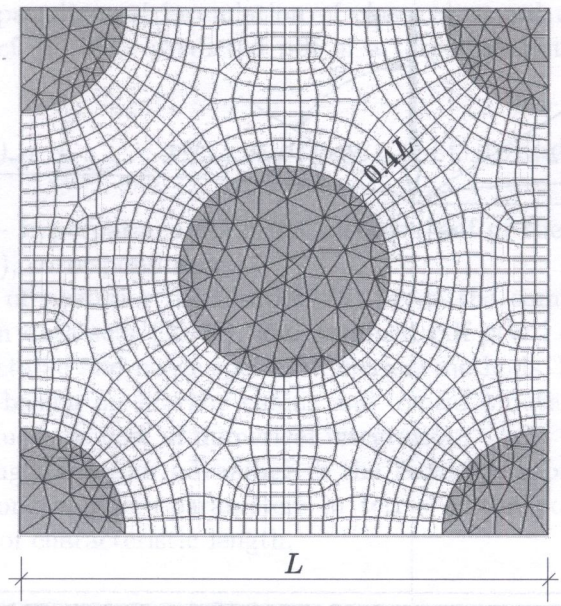


Fig. 4. Discretization for indentation test of a half of deformable body. The entire mesh is shown on the left with increasing magnification of the mesh in the vicinity of the indentor shown to the right

the second gradient of displacements regularizes the problem and so there are no singularities in the strain or stress fields. Since the material at the microscopic level is considered elasto-plastic, the equivalent material at the macroscopic level behaves inelastically.

The deformation of a typical RVE within the vicinity of the indentor is shown in Fig. 8(a). It can be seen that the solution within the RVE strongly depends on higher order effects and therefore the result is significantly dependent on the intrinsic size of the RVE. Fig. 8(b) shows the distribution of equivalent plastic microstrain which depends on the history of the macrostrains and the gradient of macrostrains at the integration point.

The indentation test is undertaken for three different sizes of the RVE and the matrix is modelled either as an elastic or an elasto-plastic material. In addition, for comparison purposes, the indentation test is also undertaken assuming (i) homogeneous aluminium (i.e. all matrix) and (ii) ho-



Matrix (Aluminum):

$E_m = 70$ GPa
$\nu_m = 0.33$
$Y_m = 460 * 10$ MPa
$H_m = 70$ MPa

Inclusion (Sailicon)

$E_i = 450$ GPa
$\nu_i = 0.17$
$V_f = 0.251$

Fig. 5. Discretization of RVE (V_f is volume fraction of inclusions)

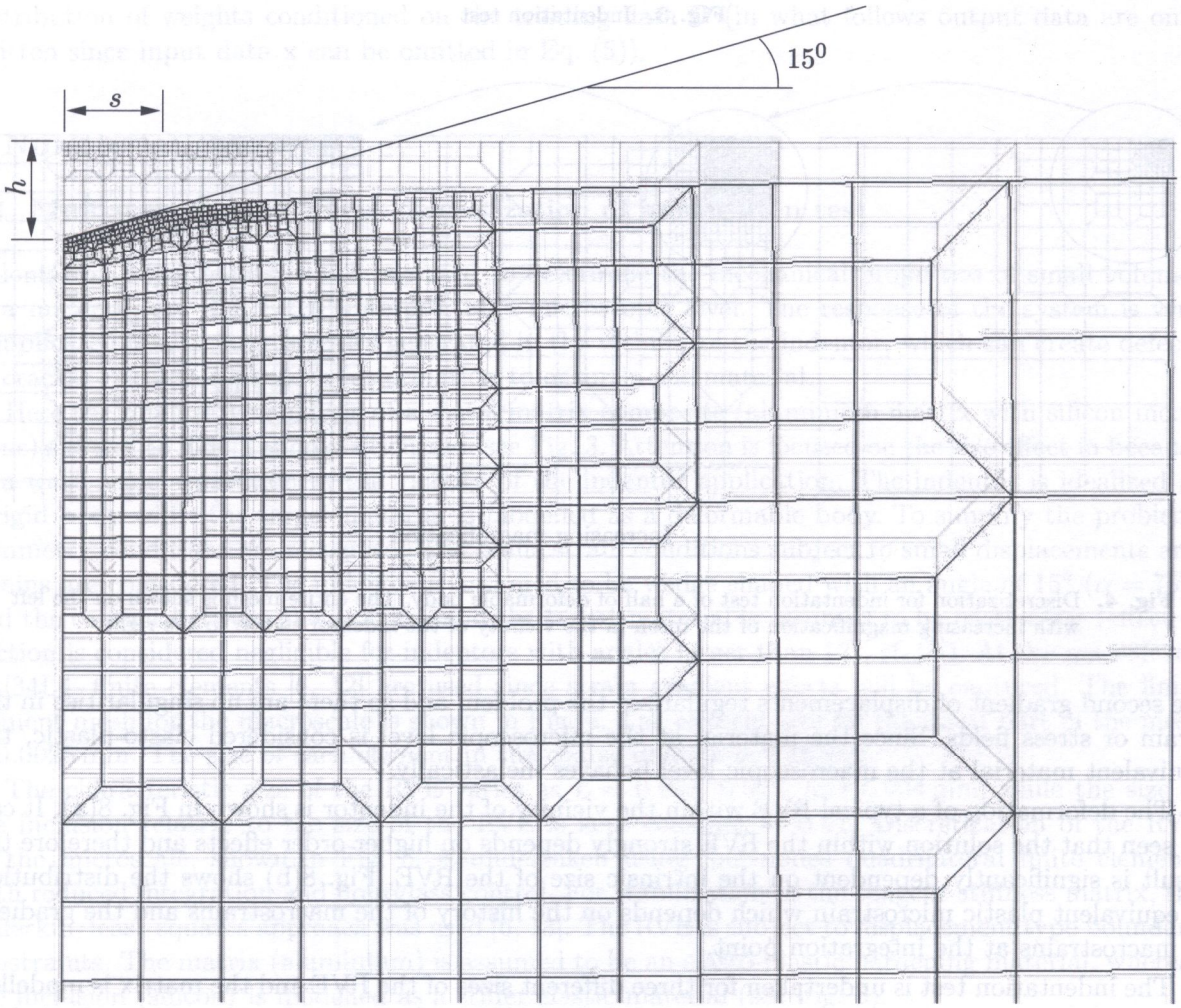


Fig. 6. Mesh deformation near the indentor

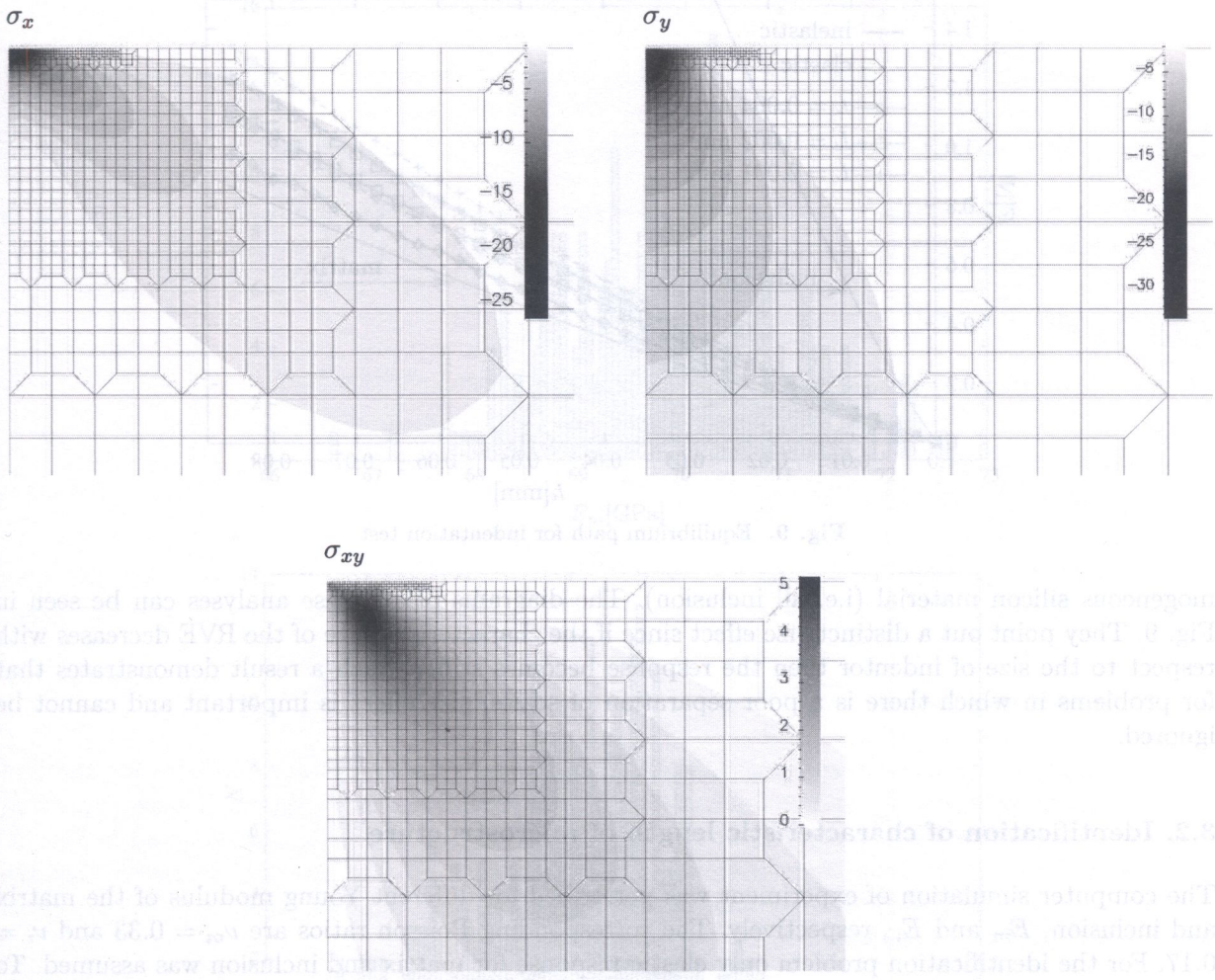


Fig. 7. Macroscopic stress components in the vicinity of indentor (inelastic matrix)

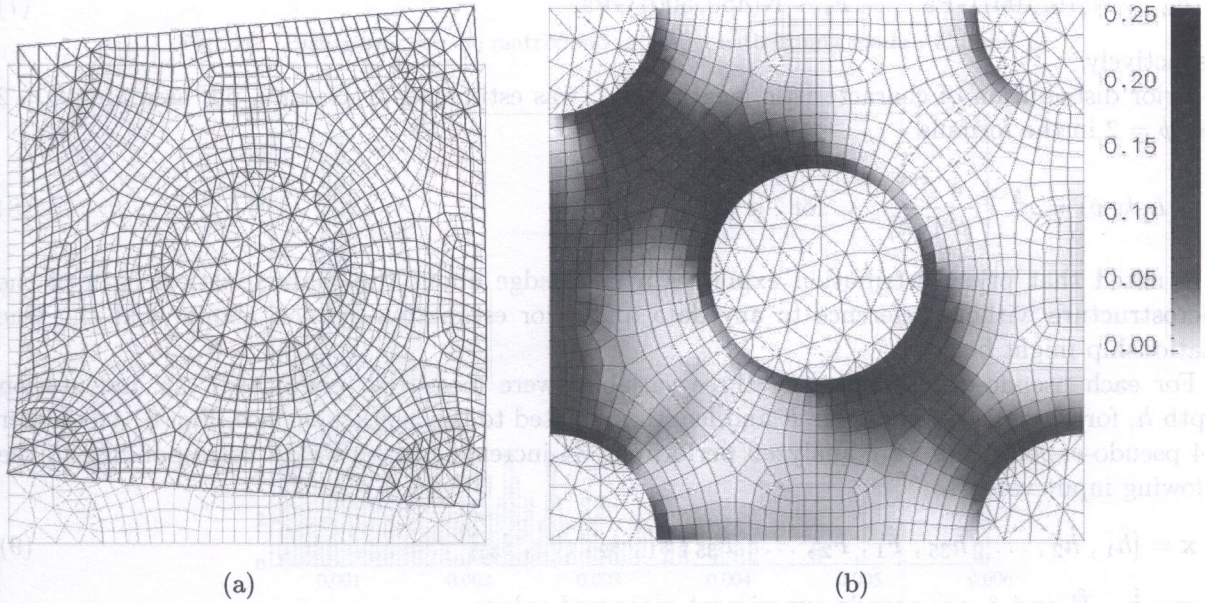


Fig. 8. (a) Deformation and (b) distribution of equivalent plastic microstrain in a typical RVE in the vicinity of indentor.

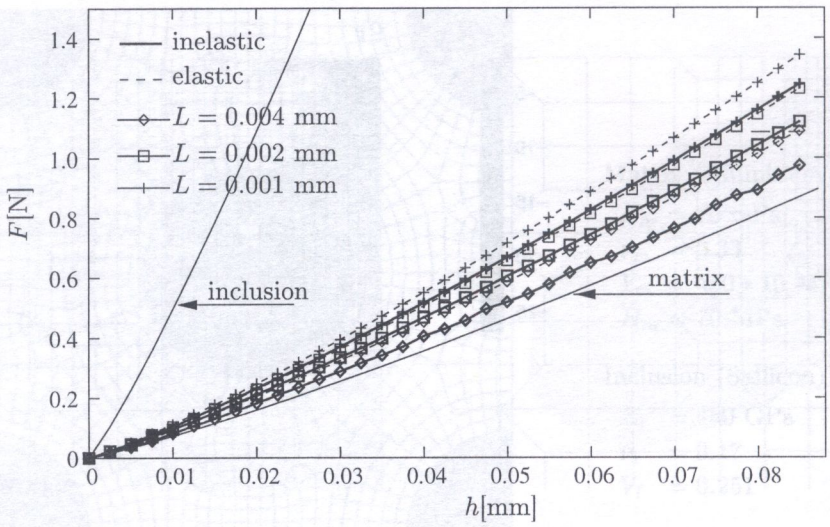


Fig. 9. Equilibrium path for indentation test

mogeneous silicon material (i.e. all inclusion). The diagrams of all these analyses can be seen in Fig. 9. They point out a distinct size effect since if the characteristic size of the RVE decreases with respect to the size of indentor then the response becomes stiffer. Such a result demonstrates that for problems in which there is a poor separation of scales, size effect is important and cannot be ignored.

3.2. Identification of characteristic length of microstructure

The computer simulation of experiment was performed for different Young modulus of the matrix and inclusion, E_m and E_i, respectively. The corresponding Poisson ratios are $\nu_m = 0.33$ and $\nu_i = 0.17$. For the identification problem only elastic response for matrix and inclusion was assumed. To simulate the pseudo-experiment, random parameters were adopted. The Gaussian normal pdf of Young moduli for matrix and inclusion were assumed:

$$E_m \sim \mathrm{N}(70, 100)\,\mathrm{GPa}, \qquad E_i \sim \mathrm{N}(450, 500)\,\mathrm{GPa}, \tag{7}$$

respectively.

Prior distribution of characteristic length size L was estimated by Gamma pdf assuming $k = 2$ and $\phi = 2$ in the formula

$$p(L, k, \phi) = x^{k-1} \frac{e^{-x/\phi}}{\phi^k \Gamma(k)} \qquad \text{for} \quad x \geq 0. \tag{8}$$

It is noted that prior distribution expresses a knowledge about the characteristic length of the microstructure without reference to any data and prior estimation opinion about what the true relationship might be.

For each pseudo-experiment only three variables were measured (computed), i.e. indentation depth h, force applied to indentor F and length s, related to the contact surface, Fig. 3. Altogether 344 pseudo-experiments were analyzed performing 34 incremental steps each corresponding to the following input vector,

$$\mathbf{x} = [\hat{h}_1\,, \hat{h}_2\,, \ldots\,, \hat{h}_{35}\,, \hat{F}_1\,, \hat{F}_2\,, \ldots\,, \hat{F}_{35}\,, \hat{s}_1\,, \hat{s}_2\,, \ldots\,, \hat{s}_{35}]^{\mathrm{T}}, \tag{9}$$

where: $\hat{h}_i$, $\hat{F}_i$ and $\hat{s}_i$ are pseudo-experiment measured values.

In Fig. 10 the histograms of Young moduli E_m and E_i for the matrix and inclusion are presented. The histogram of characteristic length L is shown in Fig. 11.

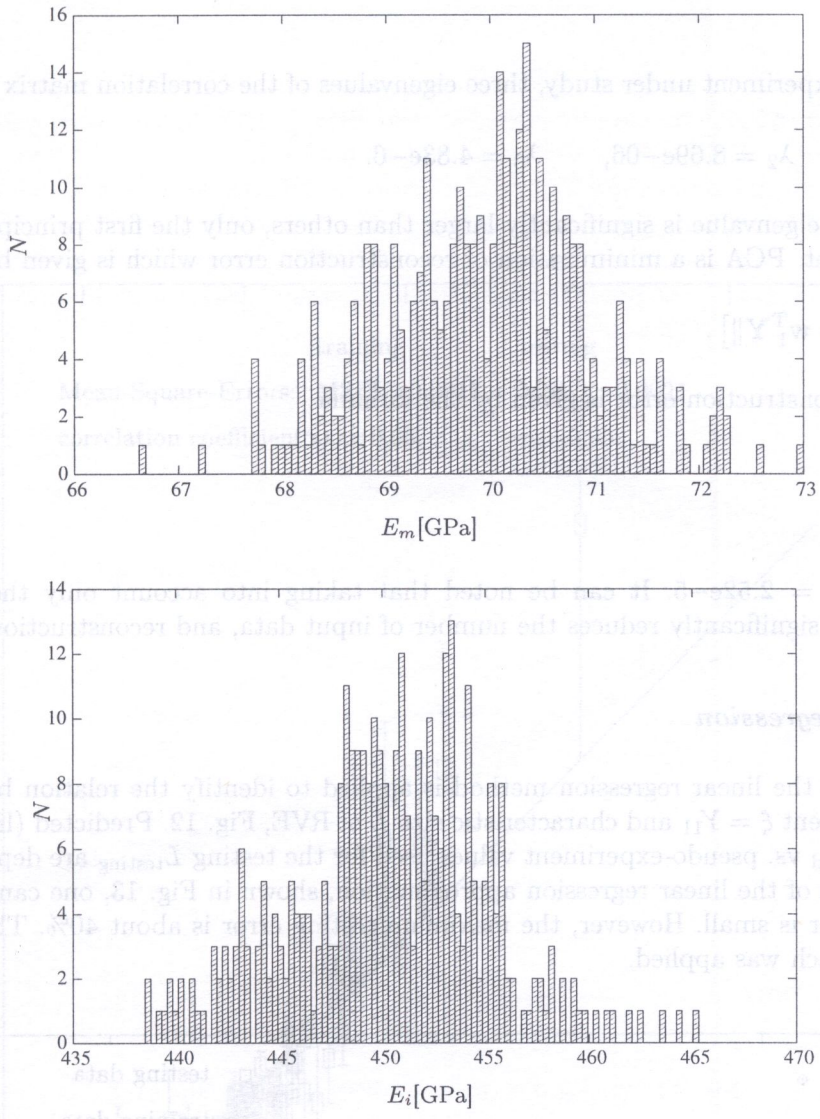


Fig. 10. Histograms for the matrix and inclusion with Young moduli E_m and E_i

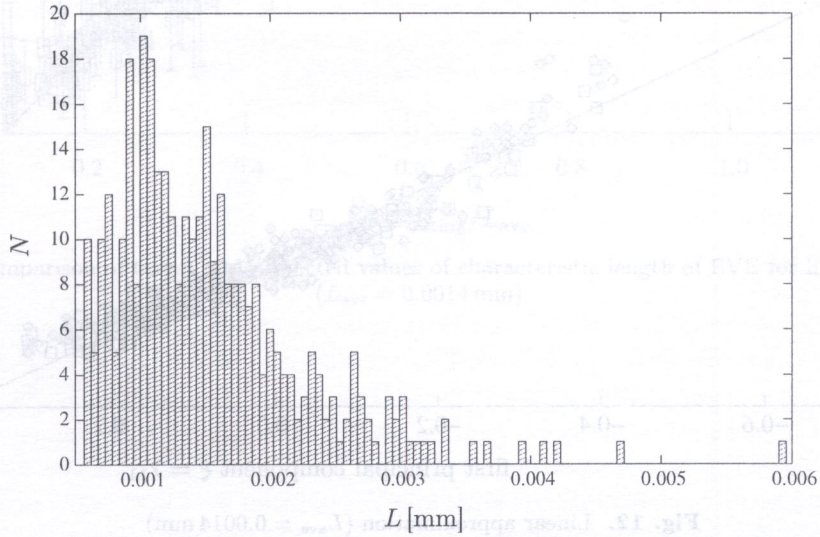


Fig. 11. Histogram of characteristic length L

3.2.1. PCA

For the pseudo-experiment under study, three eigenvalues of the correlation matrix were computed,

$$\lambda_1 = 0.055, \qquad \lambda_2 = 8.69\text{e}{-06}, \qquad \lambda_3 = 4.83\text{e}{-6}.$$

Because the first eigenvalue is significantly larger than others, only the first principal component is taken into account. PCA is a minimization of reconstruction error which is given by

$$E_r = E\left[\|\mathbf{X} - \mathbf{w}_1^{\mathrm{T}}\mathbf{Y}\|\right]. \tag{10}$$

After [5], the reconstruction error is given by the formula

$$E_r = \sum_{i=2}^{p} \lambda_i \tag{11}$$

and is equal $E_r = 2.52\text{e}{-5}$. It can be noted that taking into account only the first principal component PCA significantly reduces the number of input data, and reconstruction error is small.

3.2.2. Linear regression

After using PCA the linear regression method is applied to identify the relation between the first principal component $\xi = Y_{11}$ and characteristic size L of RVE, Fig. 12. Predicted (linear regression) values of $L_{\text{predicted}}$ vs. pseudo-experiment values used for the testing L_{testing} are depicted in Fig. 13. Looking at errors of the linear regression approximation, shown in Fig. 13, one can notice that the mean square error is small. However, the maximum testing error is about 40%. This is the reason why BNN approach was applied.

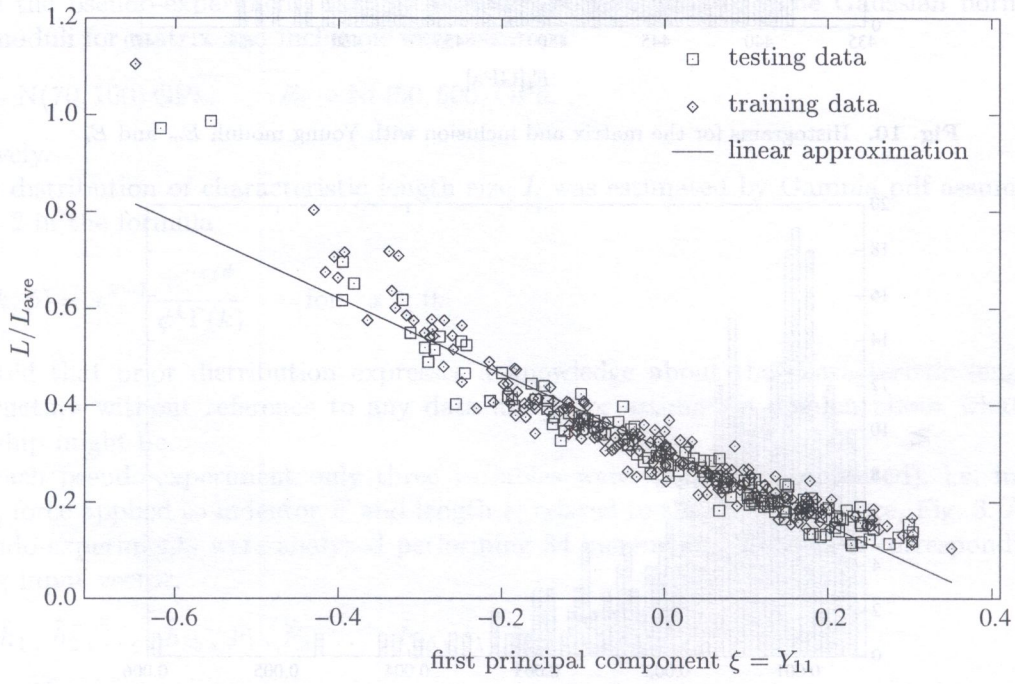


Fig. 12. Linear approximation ($L_{\text{ave}} = 0.0014\,\text{mm}$)

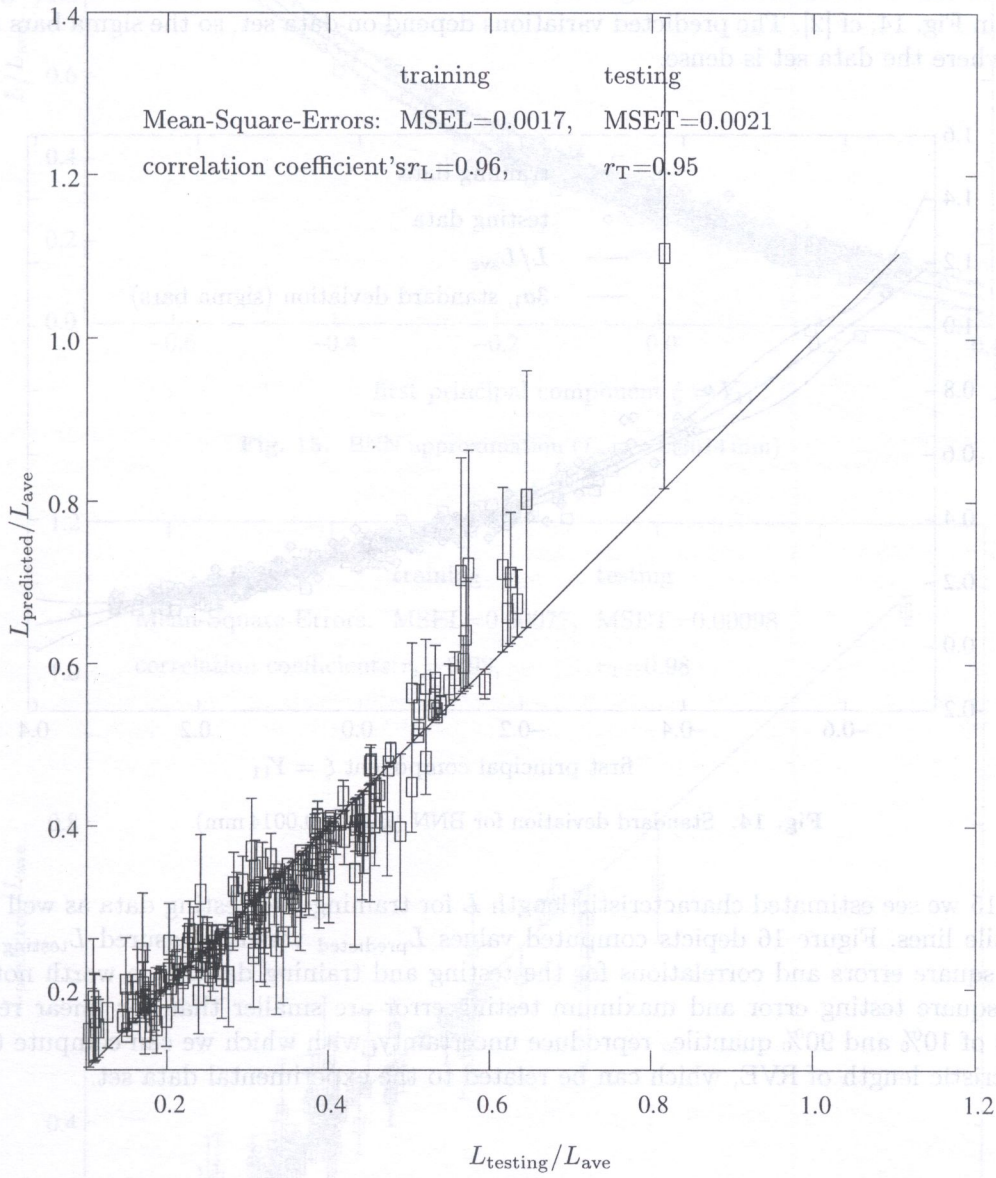


Fig. 13. Comparison of target and predicted values of characteristic length of RVE for linear regression ($L_{ave} = 0.0014\,mm$)

3.2.3. BNN

From among 344 pseudo-experiments 200 data were randomly selected and completed the training set. The remaining 144 experiments were used for the testing. Then the software FBM [11] was used for training a family of BNNs. After the cross-validation procedure was applied, cf. [5], the network BNN: 1-16-1 with 16 sigmoidal hidden units and linear output was accepted. The input corresponds to the first principal component ξ and the output is the characteristic length size L.

Predicted values of characteristic length size of RVE and its 3σ standard deviations (sigma bars) are shown in Fig. 14, cf [2]. The predicted variations depend on data set, so the sigma bars are small at points where the data set is dense.

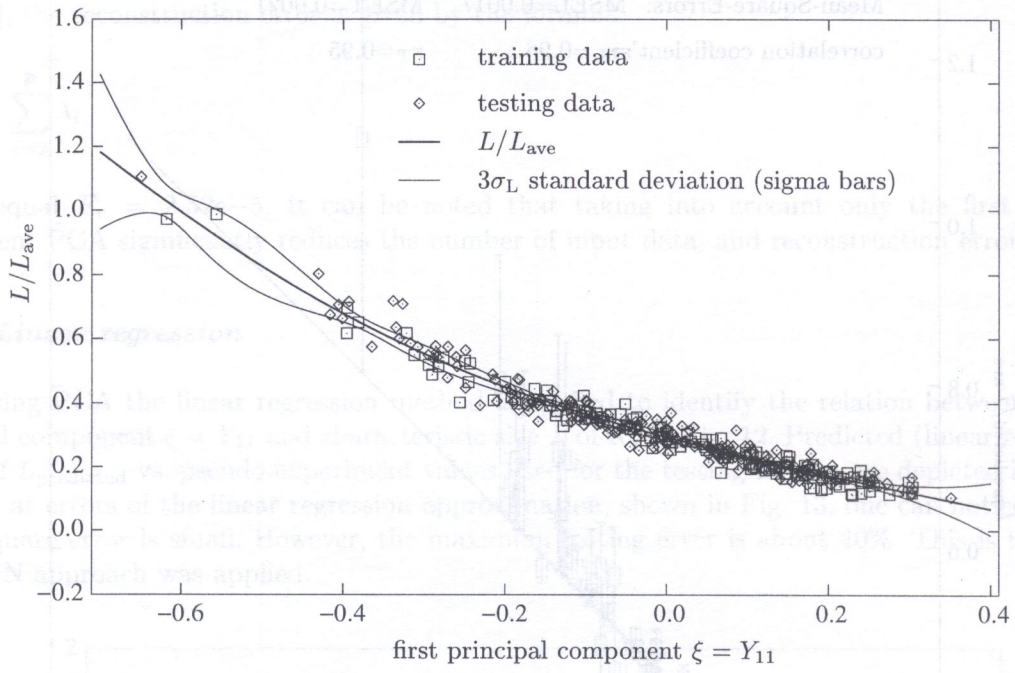


Fig. 14. Standard deviation for BNN ($L_{ave} = 0.0014\,\text{mm}$)

In Fig. 15 we see estimated characteristic length L for training and testing data as well 10% and 90% quantile lines. Figure 16 depicts computed values $L_{\text{predicted}}$ versus measured L_{testing}, median and mean square errors and correlations for the testing and training data. It is worth noting that the mean square testing error and maximum testing error are smaller than for linear regression. The values of 10% and 90% quantiles reproduce uncertainty, with which we can compute the value of characteristic length of RVE, which can be related to the experimental data set.

4. Final remarks

A computational homogenization concept for heterogeneous materials undergoing small deformations has been explored. We have applied the Bayesian approach for neural network training and demonstrated the advantages of this approach to analyze an inverse problem.

The problem of identification of characteristic length scale has been discussed. The presented approach demonstrates the advantages of the applied neural network and the data compression algorithm. PCA enables us to reduce significantly the size of input vector to only one scalar input x_1. The network BNN:1-16-1 gives not only mean values of the characteristic length of RVE $\hat{L}$ but also its normal pdf for computing the standard deviation σ_L or sigma bars.

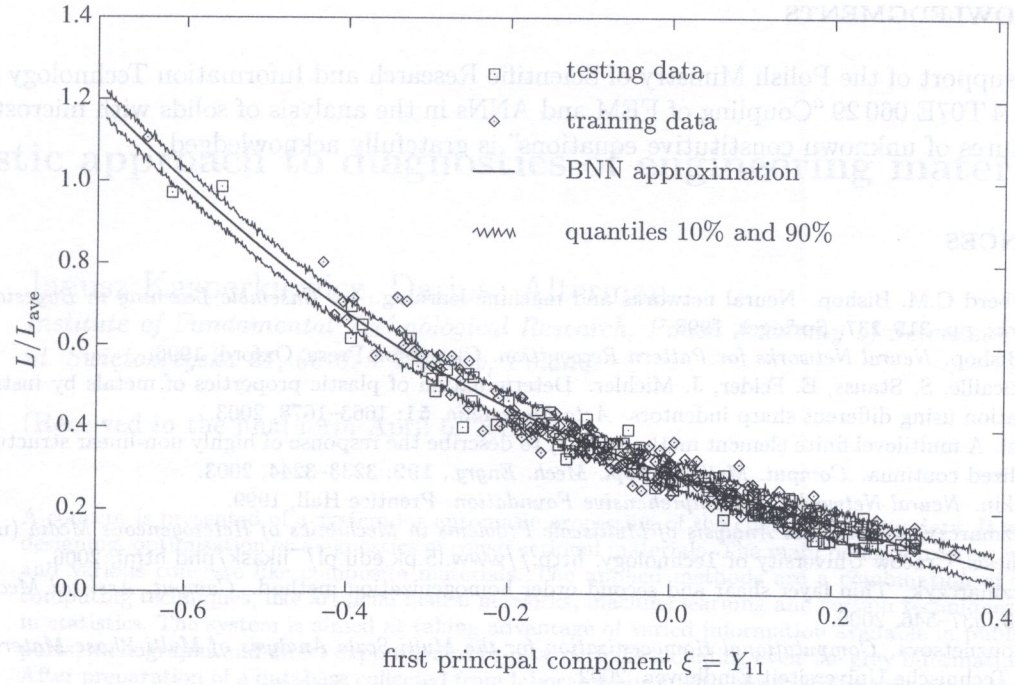


Fig. 15. BNN approximation ($L_{\text{ave}} = 0.0014\,\text{mm}$)

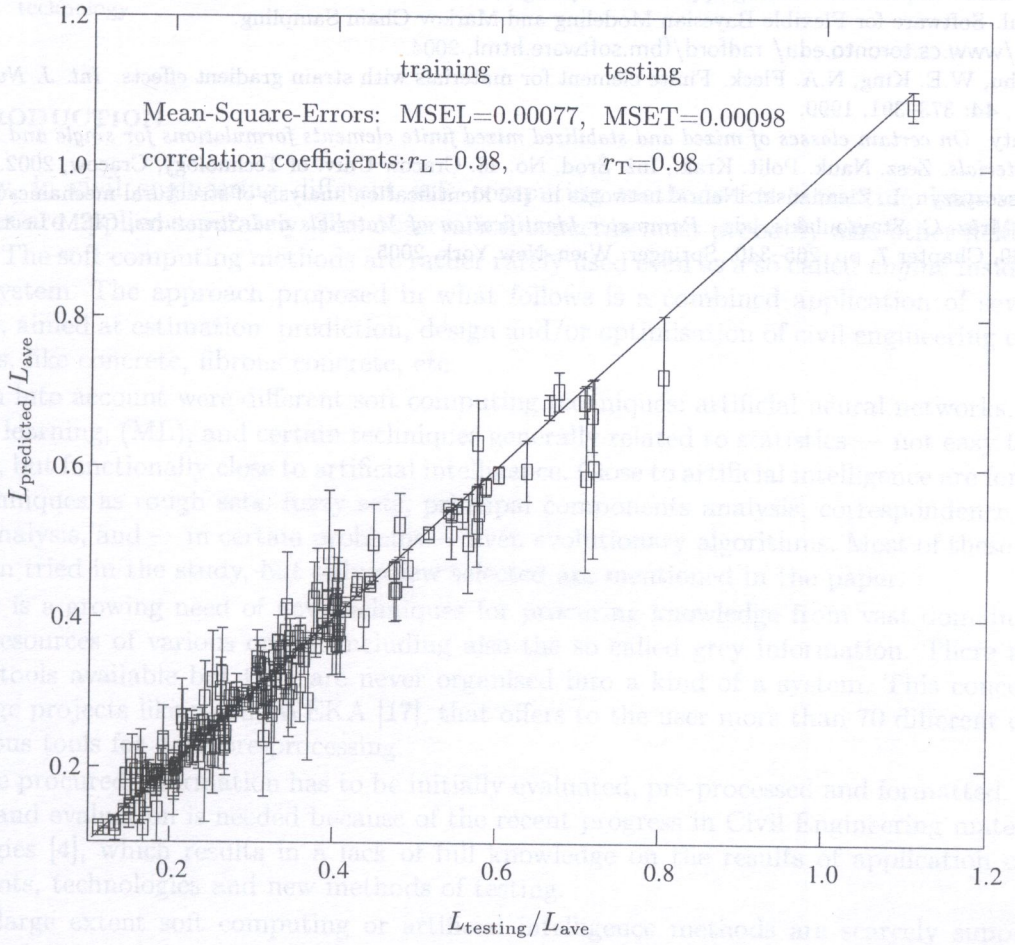


Fig. 16. Comparison of target and predicted values of characteristic length of RVE for BNN ($L_{\text{ave}} = 0.0014\,\text{mm}$)

5. Acknowledgments

Financial support of the Polish Ministry of Scientific Research and Information Technology Society, grant: No. 4 T07E 060 29 "Coupling of FEM and ANNs in the analysis of solids with microstructure and structures of unknown constitutive equations" is gratefully acknowledged.

References

[1] D. Barberd C.M. Bishop. Neural networks and machine learning. In: *Ensemble Learning in Bayesian Neural Networks*, pp. 215–237. Springer, 1998.

[2] M.C. Bishop. *Neural Networks for Pattern Recognition.* Clarendon Press, Oxford, 1996.

[3] J.L. Bucaille, S. Stauss, E. Felder, J. Michler. Determination of plastic properties of metals by instrumented indentation using different sharp indentors. *Acta Materialia*, **51**: 1663–1678, 2003.

[4] F. Feyel. A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua. *Comput. Methods Appl. Mech. Engrg.*, **192**: 3233–3244, 2003.

[5] S. Haykin. *Neural Networks. A Comprehensive Foundation.* Prentice Hall, 1999.

[6] Ł. Kaczmarczyk. *Numerical Analysis of Multiscale Problems in Mechanics of Heterogeneous Media* (in Polish). PhD thesis, Cracow University of Technology, http://www.l5.pk.edu.pl/~likask/phd.html, 2006.

[7] Ł. Kaczmarczyk. Thin layer shear and second order homogenization method. *Comput. Assisted Mech. Engrg. Sci.*, **13**: 537–546, 2006.

[8] V.G Kouznetsova. *Computational Homogenization for the Multi-Scale Analysis of Multi-Phase Materials.* PhD thesis, Technische Universiteit Eindhoven, 2002.

[9] J. Lampinen, A. Vehtari. Bayesian approach for neural networks. *Neural Networks*, **14**(3): 7–24, 2001.

[10] M.E. Tipping. *Advanced Lectures on Machine Learning*, chapter 'Bayesian inference: an introduction to principles and practice in machine learning', pp. 41–62. Springer, 2004.

[11] R. Neal. Software for Flexible Bayesian Modeling and Markov Chain Sampling. http://www.cs.toronto.edu/ radford/fbm.software.html, 2004.

[12] J.Y. Shu, W.E. King, N.A. Fleck. Finite element for materials with strain gradient effects. *Int. J. Num. Meth. Engrg.*, **44**: 373–391, 1999.

[13] A. Truty. *On certain classes of mixed and stabilized mixed finite elements formulations for single and two-phase geomaterials.* Zesz. Nauk. Polit. Krak., Inż. Środ. No. 48. Cracow Univ. of Technology, Cracow, 2002.

[14] Z. Waszczyszyn, L. Ziemiański. Neural networks in the identification analysis of structural mechanics problems. In: Z. Mróz, G. Stavroulakis, eds., *Parameter Identification of Materials and Structures*, CISM Lecture Notes No. 469, Chapter 7, pp. 265–340. Springer, Wien–New York, 2005.