Zenon Waszczyszyn, Dominika Ziaja, Marzena Kłos
Rzeszów University of Technology
Faculty of Civil and Environmental Engineering
Department of Structural Mechanics

APPLICATION OF THE METHOD OF GAUSSIAN PROCESSES TO IDENTIFICATION PROBLEMS OF STRUCTURES AND SOIL MECHANICS

Two problems are analyzed in the paper: 1) prediction and identification of critical loads and concrete strength in compressed R/C columns, 2) identification of compaction characteristics in granular soils. The main goal of the paper is to compare the numerical efficiency of the Method of Gaussian Processes with the results obtained by means of other methods (Extended Back Propagation Neural Network, Semi Bayesian Neural Networks and Bayesian methods).

Key words: Gaussian Processes, artificial neural networks, Bayesian methods, reinforced compressed columns, grain soils, critical loads, compaction characteristics.

Проаналізовано: 1) передбачення і визначення критичних навантажень і міцності бетону в стиснутих залізобетонних колонах, 2) визначення характеристик стисливості в сипучих грунтах. Основна мета статті: порівняти чисельну ефективність методу гауссових процесів з результатами, отриманими із застосуванням інших методів.

Ключаві слова: гаусові процеси, штучні нейронні мережі, байєсівські методи, залізобетонні стиснуті колони, сипучі грунти, критичні навантаження, характеристики стисливості.

1. Introduction

A research group, under the supervision of Prof. Waszczyszyn, has developed research on applications of Artificial Networks (ANNs) in the analysis of structural and civil engineering problems [1, 2]. It has been stated that ANNs are especially suitable in the inverse analysis, i.e. regression analysis in which either excitations or material parameters are identified for known structural responses or material features.

Quite recently, the authors have focused their attention on the Bayesian methods and their applications in the Semi Bayesian Neural Network (SBNN) and, especially, in the Method of Gaussian Processes (MGP). In case of SBNN, the vector of connection weights w is the model main attribute, similarly as in the deterministic Extended Back Propagation Neural Network (BPNNext). MGP is similar to radial basis function neural networks but without application of w. The main attribute of GPM is the correlation matrix of input data, see [3, 4]. On the base of analyzed examples, it has been proved that for small and medium sets of data MGP seems to be more numerically efficient than BPNNext and SGNN.

Two problems of structure and soil mechanics are discussed in the presented paper. The first problem concerns the prediction or identification of the critical load or the strength of concrete in reinforced compressed columns. The other problem deals with the identification of compaction parameters of granular soils.

In both discussed problems the target output sets were adopted, either from buckling tests on laboratory models or from in situ tests on granular soils. The main goal of the paper is to conclude that the application of MGP can give results of identification parameters of high accuracy and that the GP method is numerically more efficient than BPNNext and SBNN.

2. Some basics of Gaussian Processes

The Gaussian Processes are based on the covariant functions of the distance between pattern points in the input space:

$$c_{mn} = k(\mathbf{x}^m, \mathbf{x}^n) + \sigma_v{}^2 \, \delta_{ij} \ , \tag{8.75}$$

where: $k\,(\mathbf{x}^m - \mathbf{x}^n) = k\,\|\mathbf{x}^m - \mathbf{x}^n\|$ – kernel functions, $\sigma_v{}^2$ – variance of the target data distribution. The functions c_{mn} are components of the covariance matrix **C** with the regularization parameter $\sigma_v{}^2$ at the matrix diagonal, cf. [3].

The Gaussian Process is a model for predicting a new target value t^{N+1} starting from the known target set $\mathbf{t} = \{t_1, \cdots, t_N\}$ and assuming the Gaussian conditional distribution $p(t^{N+1} | \mathbf{t})$. Then, the covariance matrix can be partitioned:

$$\mathbf{C}_{N+1} = \begin{bmatrix} \mathbf{C}_N & \mathbf{k} \\ \mathbf{k}^{\mathrm{T}} & c \end{bmatrix}, \tag{2}$$

where: $k = \{k_1, \cdots, k_N\}$, $c = k(\mathbf{x}^{N+1}, \mathbf{x}^{N+1}) + \sigma_v{}^2 \delta_{mn}$. Thus, the conditional distribution $p(t^{N+1} | \mathbf{t})$ is a Gaussian distribution with mean and covariance given estimated by:

$$m(\mathbf{x}^{N+1}) = \mathbf{k}^T \mathbf{C}_N^{-1} \, \mathbf{t}, \quad \sigma^2(x_{N+1}) = c - \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{k} \ . \tag{3}$$

Computations in our paper were based on the squared exponential covariance function SE, see [4], of weighted distance between points m and n :

$$c(\mathbf{x}^m, \mathbf{x}^n) = v_0 \exp[-\frac{1}{2} \sum_{i=1}^{D} a_i \, (x_i^m - x_i^n)^2] + b \ , \tag{4}$$

where the vector of parameters **Θ** has D+3 parameters, where D is the dimensionality of the input space: space :

$$\mathbf{\Theta}_{(D+3)\times 1}^{\mathrm{Ex}} = \{v_0, b, a_1, \cdots, a_D, \sigma_N{}^2\} \ . \tag{5}$$

In the second identification problem, see below Fig. 4a, the other function, called Rational Quadratic function RQ, see [4], was also applied.

3. Prediction of critical loads and identification of concrete stress in R/C compressed columns

Computational data

This study case is based on laboratory tests, related to two databases: i) PEER (Pacific Earthquake Eng. Res. Center, Univ. of California), see [5], ii) K. Chudyba Ph.D. Thesis [6]. From theses databases P = 92 patterns were completed. cf. [7], and then this set was randomly split into the learning and testing sets, composed of L = 65 and T = 35 patterns, respectively.

Two problems were analyzed:

I) Prediction of the critical load P_{cr}:

$$x_{(7\times 1)} = \{\mathbf{Ca}, f_{\mathrm{c}}\}, \ \ y = P_{\mathrm{cr}} \ , \tag{6I}$$

II) Identification of the concrete strength f_{c}:

$$x_{(7\times 1)} = \{\mathbf{Ca}, P_{\mathrm{cr}}\}, \ \ y = f_{\mathrm{c}} \ , \tag{6II}$$

where the subvector $\mathbf{Ca}_{(6\times1)}=\{B, H, L, \rho, rb, f_y\}$ has the following components: B, H, L – dimensions of columns, ρ, rb, f_y – reinforcement percentage, number of longitudinal bars and steel yield stress.

BPNNext and neurocomputing

Back Propagation Neural Network (BPNNext) was applied for the validation of results obtained by MGP. The BPNNext was formulated according to the manual MATLAB Neural Networks Toolbox, see [8]. The error measure of BPNNext was formulated in the following form:

$$F(\mathbf{w}) = \frac{\gamma}{2}\frac{1}{S}\sum_{p=1}^{S}(y^p - t^p)^2 + \frac{1-\gamma}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w}. \tag{7}$$

In the measure (7) the first term corresponds to the mean square error and the second term is related to the neural network MLP (Multilayer Perceptron) for $\gamma = 0$. The regularization parameter $\gamma \neq 0$ is automatically optimized by means of procedures listed in [7].

The one hidden layer networks, see Fig.2, were designed by means of the classical cross-validation method, applying the training set T as the validation set. Bipolar sigmoid functions were adopted in the hidden layers and the linear output was assumed.

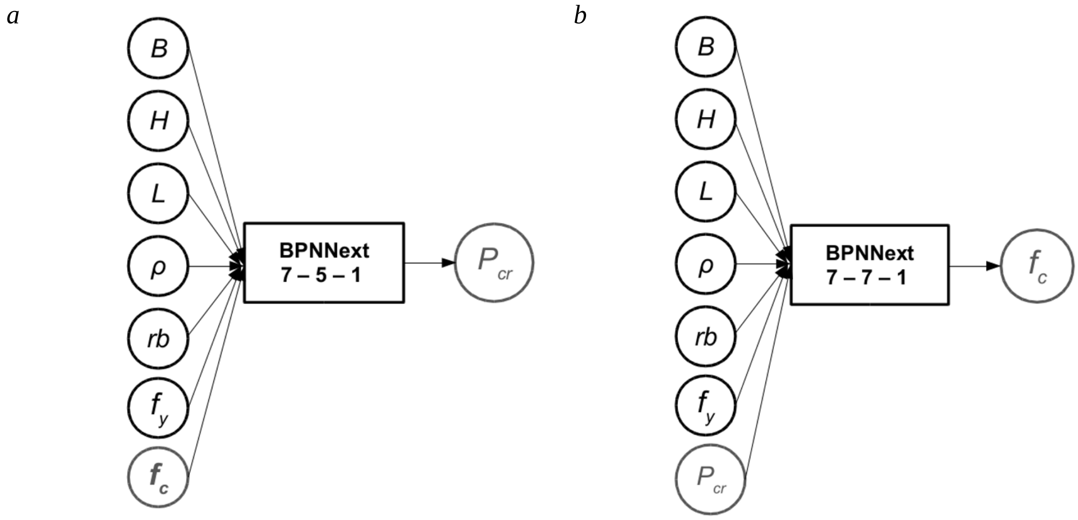

Fig. 1. BPNNext for:
a) Prediction of critical load P_{cr} , b) Identification of concrete strength f_c

The networks were trained applying the Levenberg-Marquardt learning method. After the training procedure was over, the distribution of training and testing point was obtained, as shown in Figures 2a, b. The distribution of points p on the plane (t^p, $y^{p\cdot}$) was explored for the formulation of cumulative curves.

Following the previous papers, see e.g. [9], the bounds $Re \leq \pm B\%$ are introduced in Figs 2 as lines corresponding to relative errors for points p, for which $|Rep| \leq B\%$, where:

$$Rep = (y^{p\cdot}/\,t^p - 1)\times 100\% \,. \tag{8}$$

Having defined Re, the cumulative function, called in [9] the Success Ratio function $SC(Re)$, can be formulated:

$$SR = (SRe/S)\times 100\% \,, \tag{9}$$

where: SRe – number of points (t^p, $y^{p\cdot}$) in the area $\pm Re$, S – total number of points corresponding to the learning and testing data sets, completed of either L or T pattern points, respectively. After the application of BPNNext the $SC(Re)$ curves are plotted in Figs 3a, b.

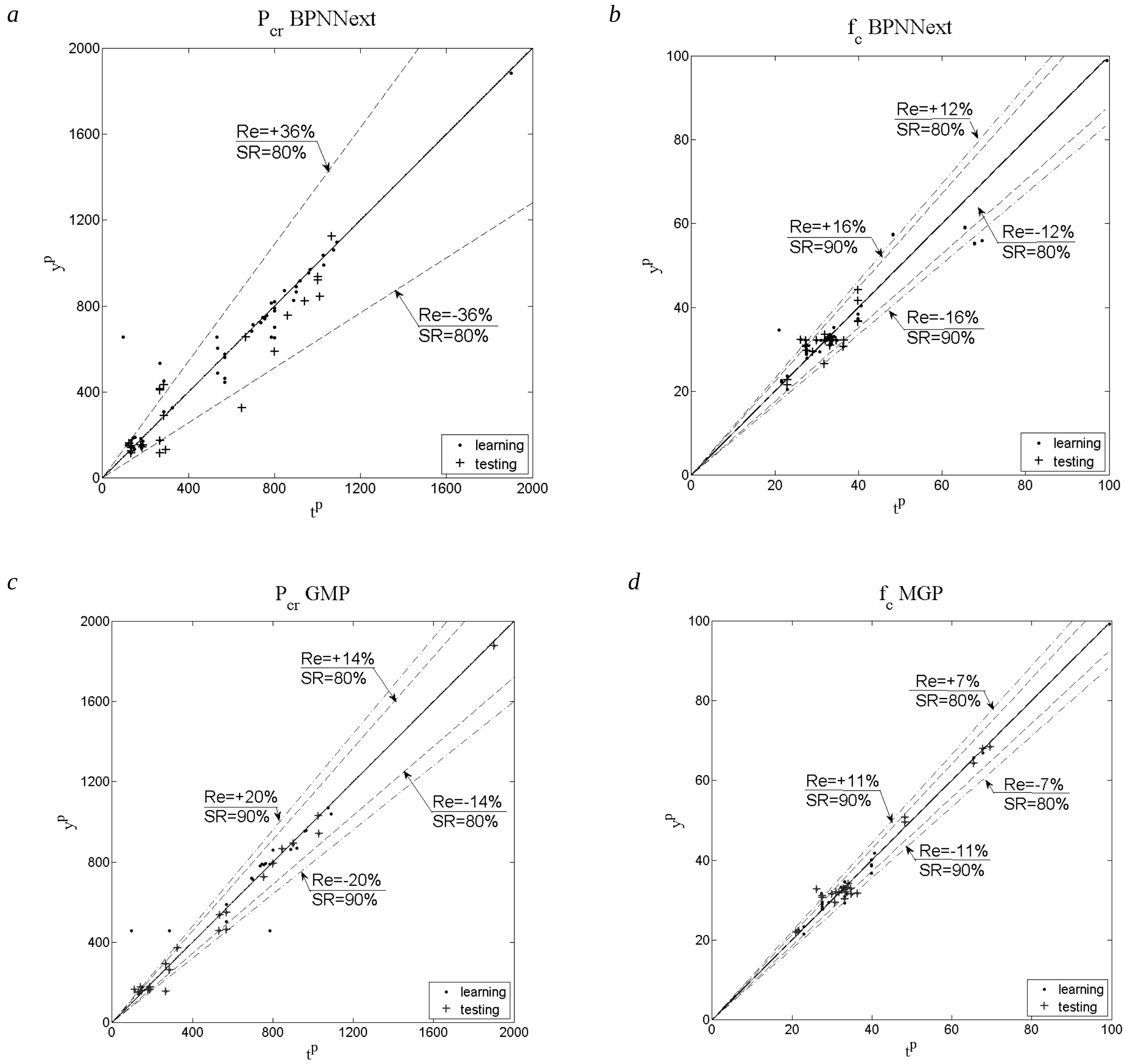

Fig. 2. Distribution of points (t^p, y^p) computed by:
a, b) BPNNext, c, d) MGP

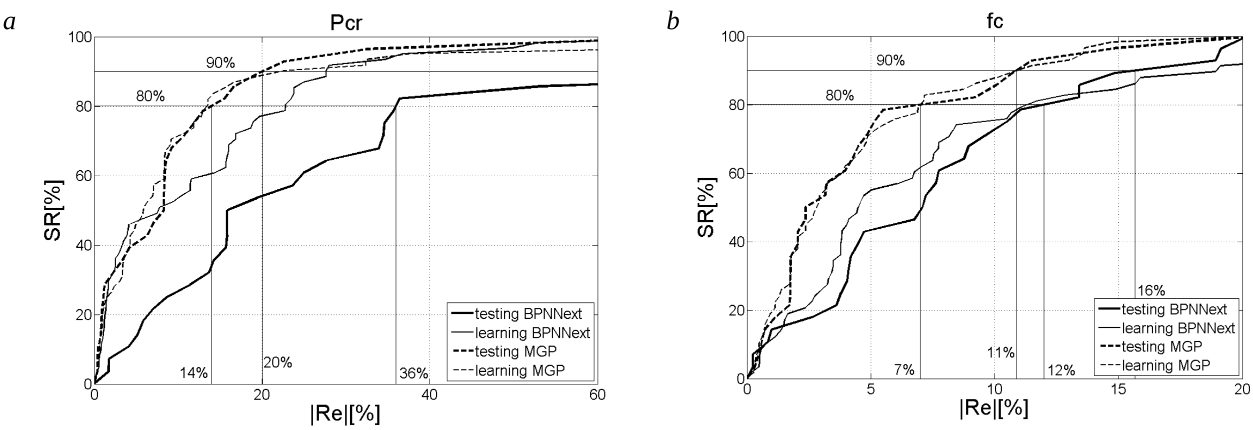

Fig. 3. Successes Ratio curves for:
a) Critical load prediction P_{cr} , b) Concrete strength f_c

MGP and comparison of results

The computation by MGP was carried out by means of procedures of the NETLAB system, briefly described in the book [4]. The same learning and testing patterns were used as those in the application of BPNNext. In Figs 2c and 2d, there are shown results obtained by the application of MGP for two considered cases (6I) and (6II). The same concerns Fig. 3a and Figs 3b, where the MGP cumulative curves $SR(Re)$ are compared with those curves corresponding to BPNNs application.

Starting from the cumulative curves drawn for the prediction of P_{cr} by MGP, it is visible that for SC = 80% and 90% the corresponding values of error bounds are $|Re|$ = 14% and 20% , respectively. This means that 80% or 90% of the total number of patterns were predicted correctly within the error bounds mentioned above. In the case of the BPNNext network only the corresponding value |Re| = 36% was found for SC = 80%. The values listed above of the error bounds are plotted in Figs 2a, 3a.

In the case of concrete strength f_c identification by BPNNext, the corresponding values for 80% and 90% patterns are correctly predicted with $|Rep| \le$ 12% and 16% . The application of MGP enables obtaining better evaluation of correctly identified values of the concrete strength f_c. The corresponding figures are $|Rep| \le$ 7% and 11% for SC = 80% or 90%, respectively.

4. Identification of compaction parameters in granular soils

Engineering structures involving earthwork often require compaction to improve soil conditions. In case of granular soils, the Optimum Water Content (OWC) and Maximum Dry Density (MDD) are essential characteristics for the design of compacted earthwork. These characteristics can be found experimentally by means of the grain size distribution measured 'in situ', cf. [10]. ANNs were recently applied for the identification of OWC and MDD. It was proved that the networks supported on Bayesian methods, e.g. the so called Semi Bayesian Neural Networks (SBNNs), are computationally efficient.

SBNN vs. GPM

The SBNN is based on the extended error measure, similarly as BPNNext:

$$F_i = \frac{\beta_i}{2} E_{D\,i}(\mathbf{w}) + \frac{\alpha_i}{2} E_{W\,i}(\mathbf{w}) = \frac{\beta_i}{2} \sum_{n=1}^{N} \{t_i^n - t_i^n\}^2 + \frac{\alpha_i}{2} \sum_{i=1}^{W} w_i^{\,2} \,, \tag{10}$$

where: i – number of output OWC, MDD; $\mathbf{w} \in R^W$ – vector of synaptic weights of the network. The Bayesian criterion MLL (Maximum Marginal Likelihood) was applied for computing the optimal number of neurons H_{opt} in the hidden layer using only the learning set of patterns. The same criterion was applied for finding optimal values of hyperparameters $\alpha_i \ \beta_i$, see [4].

The inclusion of MLL criterion into the Levenberg-Marquardt learning method makes such an approach more computationally not only efficient but also more 'costly' because of the increase of the number of operations. That is why the MGP was applied as a simpler and quicker solution than by SBNN .

Data and neurocomputing

In the considered study case a set of pattern pairs $\{\mathbf{x}^p, t^p\}_{p=1}^{P}$ was composed of P =121 tests corresponding to the in situ measurement of postglacial soils from the north of Poland, see [9]. This set of data was randomly split into $L = 0.7\times P = 85$ and $T = 0.3\times P$ patterns.

According to [8] the input data are related to grain size distribution. After the correlation analysis a sequence of nine grain diameters $\{D_x\}_{(9\times1)} = \{D_{10}, \dots, D_{90}\}$ was adopted, where x [%] is the percentage of grain diameters below which the soil mass is placed. Besides these parameters, the uniformity parameter $C_U = D_{10} / D_{10}$ was also included into the vector of input data. Thus, the following input/output data were adopted:

$$\mathbf{x}_{(10\times1)} = \{C_U, D_x \,|\, x = 10\%, \dots, 90\%\} \,, \qquad y_1 = OWC \text{ and } y_2 = MDD \,. \tag{10}$$

Following [9, 8] only one neural network output was adopted. When SBNN was applied, the Evidence procedure in the NATLAB system [4] was used to design the networks. The architecture 10–2–1 was found for computing OWC values and 10–4–1 for computing MDD.

The NATLAB system was also used for computing by the MGP. In Figures 4 the distribution of points p on the planes $(t^p, y^{p\cdot})$ is shown. In Figures 5 the SR curves are presented, obtained by MGP and SBNNs.

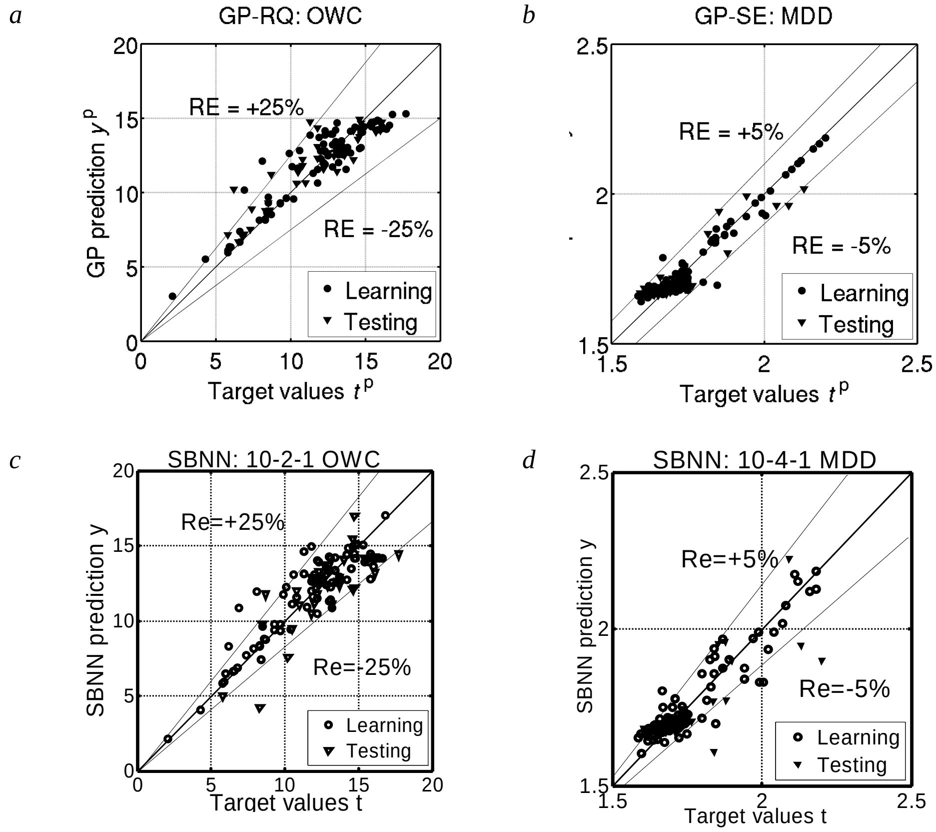

Fig. 4. Distribution of points $(t^p, y^{p\cdot})$ for the compaction characteristics, computed by SBNNs and MGP

Comparison of results

The approximate average bounds *Re* =± 25 % for OWC and *Re* =± 25 % MDD were drawn for the results obtained by GPM and SBNNs. As can be seen, the SC curves are close to each other, so we can deduce that the accuracy of computations by the GPM and SBNNs methods is comparable. This was proved in [9], where the network approximation errors were computed for different error measures.

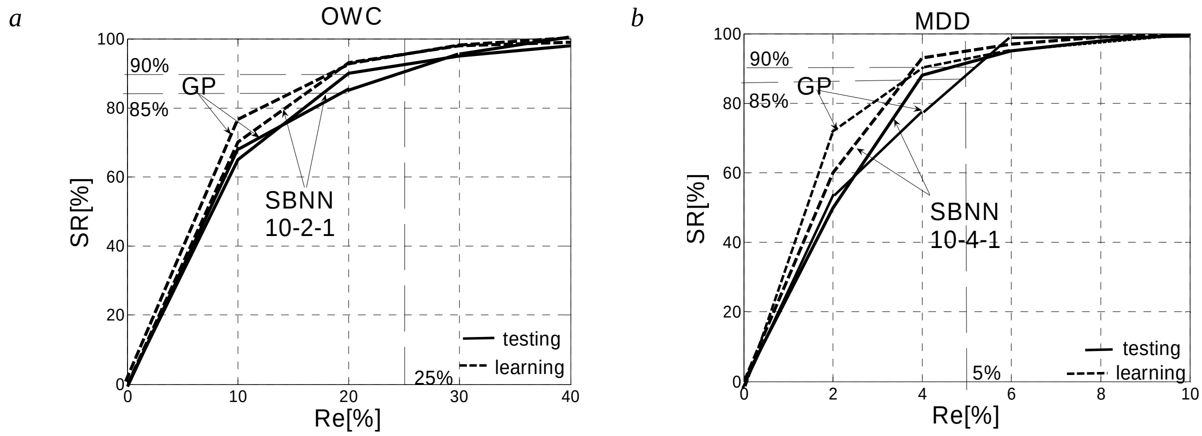

Fig. 5. Success Ratio curves SC (Re) for compaction parameters OWC and MDD computed by SBNNs and GPM

5. Final remarks and some conclusions

1. Two different problems were analyzed in the present paper. In the frame of regression analysis the internal reverse problems were analyzed. These problems correspond to the parametric identification of material characteristics (the strength of concrete f_c in columns and compaction parameters *OWC* and *MDD* in granular soils).

2. The accuracy of computed solutions was estimated by the Success Ratio curves *SC* (*Re*). These curves enable us to see the relation between desirable percent of correct placements of points (t^p, $y^{p\cdot}$), e.g. SC = 80% or 90%, in the areas $|Re| \leq B$ % .

3. Looking at Success Ratios it is visible that in the first problem the accuracy reached by MGP in the first problem is much better than by BPNNs. This concerns both the critical load prediction and concrete strength identification.

4. In case of the identification of compaction parameters in grain soils, the application of MGP gives results comparable with those by SBNN. It is, however, worth emphasizing that the application of MGP is numerically 'cheaper' because of a significantly smaller number of operations than those needed in the application of SBNN

1. Waszczyszyn Z.: Artificial neural networks in civil and structural engineering: ten years of research in Poland. Computer Assisted Mech. and Eng. Sc., 13: 489-512, 2006. 2. Waszczyszyn Z.: Artificial neural networks in civil engineering: another five years of research in Poland. Computer Assisted Mech. and Eng. Sci., 18: 131-146, 2013. 3. Bishop C.M.: Pattern Recognition and Machine Learning. Springer Science & Business Media, LLC, 2006. 4. Nabney I. T.: NETLAB – Algorithms for Pattern Recognition. Springer, London, 2004. 5. http://www.ce.washinton.edu/peera1/, 2007. 6. Chudyba K.: The influence of material and geometric characteristics on the compressive load capacity of reinforced concrete columns (in Polish). Ph.D. Thesis, Cracow Univ. of Technology, 1999. 7. Jakubek M.: Neural network prediction of load carrying capacity for eccentrically loaded reinforced concrete columns. Computer Assisted Methods in Eng. and Sci., 19: 339-549, 2012. 8. Neural Network Toolbox for Use with MATLAB. User's Guide, Version 3. The Math Works Inc, Nattik, MA,1998. 9. Kłos M, Sulewska M., J. and Waszczyszyn Z.: Neural identification of compaction characteristics for granular soils. Computer Assisted Mech. and Eng. Sci., 18: 265-274, 2013. 10.Sulewska M.: Artificial Neural Networks for Evaluation of Compaction Parameters of Cohesionless Soils (in Polish). Committee of Civil Eng. of the Polish Acad. Sci., Warsaw – Białystok, 2009.