Efficient Neural Network Models for Structural Reliability Analysis and Identification Problems

CONFERENCE PAPER · JANUARY 2005

DOI: 10.4203/ccp.82.41

Yiannis Tsompanakis
Technical University of Crete

Nikos Lagaros
National Technical University of Athens

Efficient neural network models for structural reliability analysis and identification problems

Yiannis Tsompanakis[1], *Nikos D. Lagaros*[2] *and Georgios E. Stavroulakis*[3]

[1] *Division of Mechanics, Department of Applied Sciences, Technical University of Crete, Greece*
e-mail: jt@mechanics.tuc.gr

[2] *School of Civil Engineering, National Technical University of Athens, Greece*
e-mail: nlagaros@central.ntua.gr

[3] *Institute of Mechanics, University of Ioannina, Greece and Institute of Applied Mechanics, Technical University of Braunschweig, Germany,*
email: gestavr@cc.uoi.gr

Abstract

The objective of this paper is to investigate the efficiency of soft computing methods, in particular methodologies based on neural networks, when incorporated into the solution of computationally intensive engineering problems. Two types of applications have been investigated, namely flaw identification and structural reliability analysis. Artificial neural networks (ANNs) based metamodels are used in order to replace the time-consuming repeated structural analyses. The back propagation algorithm is employed for training the ANN, using data derived from selected analyses. The trained ANN is then used to predict the values of the necessary data. The numerical tests demonstrate the computational advantages of the proposed methodologies.

Keywords: artificial neural networks, simulation, inverse problems, structural reliability analysis, structural identification.

1 Introduction

The advances in computational hardware and software resources since the early 90's resulted in the development of new, non-conventional data processing and simulation methods. Among these methods soft computing has to be mentioned as one of the most eminent approaches to the so-called intelligent methods of information processing. Artificial neural networks (ANNs), expert and fuzzy systems, evolutionary methods are the most popular soft computing techniques. Especially ANNs have been widely used in many fields of science and technology, as well as, in an increasing number of problems in structural engineering. From among general problems that can be analyzed by means of ANNs the simulation and identification problems can be classified as follows [1]:

- *simulation* is related to direct methods of structural analysis, i.e., for known inputs (e.g., excitations of mechanical systems (MSs)) and characteristics of structures or materials outputs (responses of MSs) are searched;
- *inverse simulation* (partial identification, for example, of an unknown excitation) takes place if inputs correspond to known responses of MSs and excitations are searched as outputs of ANNs; and
- *identification* is associated with the inverse analysis of structures and materials, i.e., excitations and responses are known and MS characteristics are searched.

Over the last ten years artificial intelligence techniques like ANNs have emerged as a powerful tool that could be used to replace time consuming procedures in many engineering applications. Some of the fields where ANNs have been successfully applied are: (i) pattern recognition, (ii) regression (function approximation/fitting) and (iii) optimization. Additionally, ANNs are presented in [2] as an alternative approach to nonlinear system modeling. In the past the first field of application of ANNs was mostly used for predicting the behavior of a structural system in the context of structural optimal design [3-7], structural damage assessment [8,9] or structural reliability analysis [10,11]. Function approximation involves approximating the underlying relationship from a given finite input-output data set. Feedforward ANNs, such as multi-layer perceptrons (MLP) and radial basis function networks have been widely used as an alternative approach to function approximation since they provide a generic functional representation and have been shown to be capable of approximating any continuous function with acceptable accuracy [12].

In this work the application of ANNs is focused on the simulation, i.e. structural reliability analysis, and identification, i.e. flaw detection, problems. Many sources of uncertainty (material, geometry, loads etc) are inherent in structural design and functioning. Reliability analysis leads to safety measures that a design engineer has to take into account due to the aforementioned uncertainties. Reliability analysis problems, especially when earthquake loadings are considered, are highly computationally intensive tasks since in order to calculate the structural behavior under seismic loads a large number of dynamic analyses (such as multi-modal response spectrum analyses) are required [13]. Soft computing techniques are used in order to reduce the aforementioned computational cost. An ANN is trained utilizing available information generated from selected multi-modal response spectrum analyses of a typical multi-storey building. The trained ANN is then used to predict the maximum inter-storey drift due to different sets of random variables. After the maximum inter-storey drift is predicted, the probability of failure is calculated by means of Monte Carlo Simulation (MCS). The results of the proposed methodology in the test examples show its efficiency and its potential for treating large-scale practical problems.

Another very promising field of soft computing applications in computational mechanics is flaw or damage detection, which basically can be considered as an inverse problem. For example, material or parameter identification problems, which can be formulated as output-error optimization problems, can be solved very efficiently with the proposed technique. For the classical formulation and solution with classical or less classical (e.g. filter-driven or genetic) optimization algorithms and neural networks details can be found in references [14-16] and the review article on inverse analysis [1]. In this work the methodology is extended by using a neural network technique for the replacement of the mechanical analysis modelling and, consequently, the genetic optimization is applied for the solution of the inverse crack or defect optimization problem. The effectiveness is compared with the results of the previously used single-method techniques in characteristic test examples.

2 Multi-layer perceptrons

ANNs metamodels have the ability of learning and accumulating expertise and have found their way into applications in many scientific areas. There is an increasing number of publications that cover a wide range of computational structures technology applications, where most of them are heavily dependent on extensive computer resources that have been investigated or are under development. This trend demonstrates the great potential of ANNs.

2.1 The Back Propagation learning algorithm

A multi-layer perceptron (MLP) is a feed-forward ANN, consisting of a number of units (neurons) linked together and attempts to create a desired relation in an input/output set of learning patterns. A neural network consists of an input layer, one or more hidden layers and an output layer. Each layer has its corresponding neurons or nodes and weight connections. A single training pattern is an I/O vector of pairs of input-output values in the entire matrix of I/O training set.

The inputs x_i, i=1, 2,...,n which are received by the input layer are analogous to the electrochemical signals received by neurons in human brain. In the simplest model these input signals are multiplied by connection weights $w_{p,ij}$ and the effective input $net_{p,j}$ to neurons is the weighted sum of the inputs

$$\mathrm{net}_{p,j} = \sum_{i=1}^{n} \mathrm{w}_{p,ij}\, \mathrm{net}_{q,i} \tag{1}$$

where $w_{p,ij}$ is the connecting weight of the layer p from the i neuron in the q (source) layer to the j neuron in the p (target) layer, $net_{q,i}$ is the output produced at the i neuron of the layer q and $net_{p,j}$ is the output produced at the j neuron in the layer p. Inputs x_i correspond to $net_{q,i}$ for the input layer.

In the biological system, a typical neuron may only produce an output signal if the incoming signal builds up to a certain level. This output is expressed in ANNs by

$$\mathrm{out}_{p,j} = \mathrm{F}(\mathrm{net}_{p,j}) \tag{2}$$

where F is an activation function which produce the output at the j neuron in the p layer. The type of activation function that has been used in the present study is the sigmoid function.

At the output layer the computed output(s), otherwise known as the observed output(s), are subtracted from the desired or target output(s) to give the error signal

$$\mathcal{E}(\mathrm{w}) = \frac{1}{2\mathrm{m}}\left\|\mathrm{E}(\mathrm{w})\right\|^2 \tag{3}$$

$$\mathrm{E}_{\mathrm{i}}(\mathrm{w}) = \sum_{\mathrm{j}=1}^{\ell}[\mathrm{out}_{\mathrm{k,j}} - \mathrm{tar}_{\mathrm{k,j}}] \tag{4}$$

where m is the number of training pairs, $tar_{k,i}$ and $out_{k,i}$ are the target and the observed output(s) for the node i in the output layer k, respectively. This is called supervised learning.

A learning algorithm tries to determine the weights, in order to achieve the right response for each input vector applied to the network. The numerical minimization algorithms used for the training generate a sequence of weight matrices through an iterative procedure. To apply an algorithmic operator $\mathcal{A}$, a starting value of the weight matrix $w^{(0)}$ is needed, while the iteration formula can be written as follows

$$\mathrm{w}^{(t+1)} = \mathcal{A}(\mathrm{w}^{(t)}) = \mathrm{w}^{(t)} + \Delta\mathrm{w}^{(t)} \tag{5}$$

All numerical methods applied in ANNs are based on the above formula. The changing part of the algorithm $\Delta w^{(t)}$ is further decomposed into two parts as

$$\Delta\mathrm{w}^{(t)} = \mathrm{a}_t \mathrm{d}^{(t)} \tag{6}$$

where $d^{(t)}$ is a desired search direction of the move and a_t the step size in that direction.

The training methods can be divided into two categories. Algorithms that use global knowledge of the state of the entire network, such as the direction of the overall weight update vector, which are referred to as global techniques. In contrast local adaptation strategies are based on weight specific information only such as the temporal behavior of the partial derivative of this weight. The local approach is more closely related to the ANN concept of distributed processing in which computations can be made independent to each other. Furthermore, it appears that for many applications local strategies achieve faster and reliable prediction than global techniques despite the fact that they use less information [17].

2.2 Global Adaptive Techniques

The algorithms most frequently used in the ANN training are the steepest descent, the conjugate gradient and the Newton's methods with the following direction vectors:

- *Steepest descent method*: $\mathrm{d}^{(t)} = -\nabla\mathcal{E}(\mathrm{w}^{(t)})$
- *Conjugate gradient method*: $\mathrm{d}^{(t)} = -\nabla\mathcal{E}(\mathrm{w}^{(t)}) + \beta_{t-1}\mathrm{d}^{(t-1)}$ where β_t is defined as

  $\beta_{t-1} = \nabla \mathrm{E}_t \cdot \nabla \mathrm{E}_t / \nabla \mathrm{E}_{t-1} \cdot \nabla \mathrm{E}_{t-1}$ (Fletcher-Reeves)
- *Newton's method*: $\mathrm{d}^{(t)} = -\left[\mathrm{H}(\mathrm{w}^{(t)})\right]^{-1} \nabla\mathcal{E}(\mathrm{w}^{(t)})$

The convergence properties of optimization algorithms for differentiable functions depend on properties of the first and/or second derivatives of the function to be optimized. When optimization algorithms converge slowly for NN problems, this suggests that the corresponding derivative matrices are numerically ill-conditioned. It has been shown that these algorithms converge slowly when rank-deficiencies appear in the Jacobian matrix of a NN, making the problem numerically ill-conditioned [18]. Furthermore, like in all local minimization algorithms, training of ANN will be interrupted without success at local minima of the error function if the latter, due to the complexity of the problem at hand, happens to be non-convex. The introduction of momentum terms in the local step may help the algorithm to avoid premature stop at local minima.

2.3 Local Adaptive Techniques

In order to improve the performance of weight updating, two completely different approaches have been proposed, namely Quickprop [19] and Rprop [20].

The Quickprop method

This method is based on a heuristic learning algorithm for a multi-layer perceptron, developed by Fahlman [19], which is partially based on the Newton's method. Quickprop is one of most frequently used adaptive learning paradigms. The weight updates are based on estimates of the position of the minimum for each weight, obtained by solving the following equation for the two following partial derivatives

$$\frac{\partial \mathcal{E}_{t-1}}{\partial w_{ij}} \text{ and } \frac{\partial \mathcal{E}_{t}}{\partial w_{ij}} \tag{7}$$

and the weight update is implemented as follows

$$\Delta w_{ij}^{(t)} = \frac{\dfrac{\partial \mathcal{E}_{t}}{\partial w_{ij}}}{\dfrac{\partial \mathcal{E}_{t-1}}{\partial w_{ij}} - \dfrac{\partial \mathcal{E}_{t}}{\partial w_{ij}}} \Delta w_{ij}^{(t-1)} \tag{8}$$

The learning time can be remarkably improved compared to the global adaptive techniques.

The Rprop method

Another heuristic learning algorithm with locally adaptive learning rates based on an adaptive version of the Manhattan-learning rule and developed by Riedmiller and Braun [20] is the **R**esilient back**prop**agation abbreviated as Rprop. The weight updates can be written

$$\Delta w_{ij}^{(t)} = -\eta_{ij}^{(t)} \operatorname{sgn}\left(\frac{\partial \mathcal{E}_{t}}{\partial w_{ij}} \right) \tag{9}$$

where

$$\eta_{ij}^{(t)} = \begin{cases} \min(\alpha \cdot \eta_{ij}^{(t-1)}, \eta_{max}), \text{ if } \frac{\partial \mathcal{E}_t}{\partial w_{ij}} \cdot \frac{\partial \mathcal{E}_{t-1}}{\partial w_{ij}} > 0 \\ \max(b \cdot \eta_{ij}^{(t-1)}, \eta_{min}), \text{ if } \frac{\partial \mathcal{E}_t}{\partial w_{ij}} \cdot \frac{\partial \mathcal{E}_{t-1}}{\partial w_{ij}} < 0 \\ \eta_{ij}^{(t-1)}, \text{ otherwise} \end{cases} \tag{10}$$

where α=1.2, b= 0.5, η_{max}=50 and η_{min}=0.1 [21]. The learning rates are bounded by upper and lower limits in order to avoid oscillations and arithmetic underflow. It is interesting to note that, in contrast to other algorithms, Rprop employs information about the sign and not the magnitude of the gradient components.

3 Structural identification

Inverse analysis concerns the determination of material or other data (in general, parameter identification) of a structure from measurements of its mechanical response caused by a given loading. Dynamic loadings give the most promising results and are used in many non-destructive evaluation procedures in industry (ultrasonic evaluation, ambient vibration or modal testing). Another 'ambitious' goal of identifying when the status of a structure has been changed beyond a critical limit or not, is structural health monitoring which has great practical importance.

The inverse crack or damage identification problem has the following general formulation as an output error minimization problem: (a) the unknown quantities like number and type of defects, their position and other geometric parameters are expressed with the help of certain variables, (b) a number of mechanical tests are considered, (c) for each value of the unknown defect parameters the corresponding responses of the structure are considered and compared with the target (measured) responses. Let z be the vector of unknown parameters involved in the mechanical problem, u_0 be the measured response of the mechanical system and $u(z)$ be the response of the mechanical system for a fixed value of z. The minimization problem reads

$$\min\ z : \{\frac{1}{2}(u(z) - u_0)^T \cdot M \cdot (u(z) - u_0)\} \tag{11}$$

where M is an appropriate symmetric and positive semi-definite weight matrix. The implicit nonlinear dependence of the structural response u on the unknown parameters z, i.e. u(z), makes the above-mentioned least square minimization problem complicated. The error function may become non-convex, with local minima, in the case of difficult and complicated inverse problems where no sufficient estimate of the solution is available.

Several methods have been presented for the effective numerical solution of this problem (among others, numerical optimization, genetic algorithms, soft computing). A comparison on crack identification problems in elasticity has been discussed in [15]. Summarizing the results one could state that a method based on genetic optimization, a method of global optimization, which is able to avoid local minima, will solve the problem in almost all cases [15]. The usage of more than one loading cases may be necessary and certainly makes the solution easier (compare the three-dimensional defect identification using static data presented in [22]). On the other hand, the computer resources needed for the use of genetic algorithms are almost prohibitive for a regular use in an industrial environment. A cheaper approach, computationally, is based on the artificial neural network methodology or on filter-based optimization techniques (for recent results see [23,24]).

ANNs (usually feedforward multilayer networks trained with various backpropagation techniques) are used for the direct approximation of the mapping between measurements and unknown parameters (the so-called inverse mapping). They have been used with great efficiency for the post-processing of impact-echo data in two-dimensional elastic structures [15]. Other practical applications of this method include the crack-depth determination of a vertical crack emanating from the hidden surface of a plate from ultrasonic back-scattering data [25] and the depth determination of surface-breaking cracks [26].

In the case that the aforementioned optimization problem has local minima the ANN approach alone may be ineffective, while the genetic optimization algorithm may still be effective. From the discussion of training algorithms in the previous sections, this result could be expected; since an optimization problem is hidden within the training phase of the ANN and this (distributed and parallel optimization) technique is able to overcome only moderate nonconvexity and may stop at local minima. In this paper a hybrid strategy made of two components is proposed. First a neural network is used for the replacement of the direct mechanical model. This part has already been used in the community of structural optimization, it is known as response surface method and it is especially beneficial for demanding modelling tasks like dynamical or computational fluid mechanics simulations. The approximation of the response mapping is, in general, easier than the approximation of the inverse mapping. The resulting model of the mechanical system is used in connection with the genetic algorithm for the solution of the inverse problem. At this point the efficiency of the genetic algorithm is enhanced, since the neural network approximator of the structural response is usually far less demanding in terms of computational cost than evaluating the actual mechanical model.

4 Seismic reliability analysis

In the design of structural systems, limiting uncertainties and increasing safety is an important issue to be considered. Structural reliability, which is defined as the probability that the system meets some specified demands for a specified time period under specified environmental conditions, is used as a probabilistic measure to evaluate the reliability of structural systems. The performance function of a structural system must be determined to describe the system's behavior and to identify the relationship between the basic parameters in the system. It should be noted that in the earthquake loading environment the uncertainties related to seismic demand and structure’s capacity are strongly coupled.

The probability of failure p_f can be determined using a time invariant reliability analysis procedure with the following expression

$$p_f = p[R < S] = \int_{-\infty}^{\infty} F_R(t)\, f_S(t)dt = 1 - \int_{-\infty}^{\infty} F_s(t)\, f_R(t)dt \tag{12}$$

where R denotes the structure's bearing capacity and S the external loads. The randomness of R and S can be described by known probability density functions $f_R(t)$ and $f_S(t)$, with $F_R(t) = p[R<t]$, $F_S(t) = p[S<t]$ being the cumulative probability density functions of R and S, respectively.

Most often a limit state function is defined as $G(R,S) = S\text{-}R$ and the probability of structural failure is given by

$$p_f = p[G(R,S) \geq 0] = \int_{G \geq 0} f_R(R)\, f_S(S)dRdS \tag{13}$$

It is practically impossible to evaluate p_f analytically for complex and/or large-scale structures, especially in the case of structural problems under seismic loads that are considered in the present study. In such cases the integral of Eq. (13) can be calculated only approximately using either simulation methods, such as the Monte Carlo Simulation (MCS), or approximation methods like the first order reliability method (FORM) and the second order reliability method (SORM), or response surface methods (RSM). Despite its high computational cost, MCS is considered as an efficient method and is commonly used for the evaluation of the probability of failure in computational mechanics, either for comparison with other methods or as a standalone reliability analysis tool.

4.1 Monte Carlo simulation

In reliability analysis the MCS method is often employed when the analytical solution is not attainable and the failure domain can not be expressed or approximated by an analytical form. This is mainly the case in problems of complex nature with a large number of basic variables where all other reliability analysis methods are not applicable. Expressing the limit state function as $G(x)<0$, where $x=(x_1,x_2,...,x_M)$ is the vector of the random variables, Eq. (13) can be written as

$$p_f = \int_{G(x) \geq 0} f_x(x)dx \tag{14}$$

where $f_x(x)$ denotes the joint probability of failure for all random variables. Since MCS is based on the theory of large numbers (N_∞) an unbiased estimator of the probability of failure is given by

$$p_f = \frac{1}{N_\infty} \sum_{j=1}^{N_\infty} I(x_j) \tag{15}$$

in which $I(x_j)$ is a binary indicator for failure and successful simulations defined as

$$I(x_j) = \begin{cases} 1 & \text{if} \quad G(x_j) \geq 0 \\ 0 & \text{if} \quad G(x_j) < 0 \end{cases} \tag{16}$$

It is important in structural reliability using simulation methods, to efficiently and accurately evaluate the probability of failure for a given performance function. In order to estimate p_f an adequate number of N_{sim} independent random samples is produced using a specific, usually uniform, probability density function of the vector x. The value of the failure function is computed for each random sample x_j and the Monte Carlo estimation of p_f is given in terms of sample mean by

$$p_f \cong \frac{N_H}{N_{sim}} \tag{17}$$

where N_H is the number of failure simulations. In order to reduce the number of simulations and the computational cost of the standard MCS many efficient sampling reduction techniques have been used [11,27]. In the present study there is no need of implementing such techniques, since ANNs replace the computationally expensive structural response evaluations and the sample size has trivial importance.

4.2 Design under seismic loading

The equations of equilibrium for a finite element system in motion can be written in the usual form

$$M\ddot{u}(t) + C\dot{u}(t) + Ku(t) = R(t) \tag{18}$$

where M, C, and K are the mass, damping and stiffness matrices; R_t is the external load vector, while $u_t, \dot{u}_t$ and $\ddot{u}_t$ are the displacement, velocity, and acceleration vectors of the finite element assemblage, respectively. Design approach based on the multi-modal Response Spectrum (mMRS) analysis, which is based on the mode superposition approach, has been used in the present study [27].

The mMRS method is based on a simplification of the mode superposition approach with the aim to avoid time history analyses which are required by both, the direct integration and mode superposition approaches. In the case of the multi-modal response spectrum analysis Eq. (18) is modified according to the modal superposition approach to a system of independent equations

$$\bar{M}_i \ddot{y}_i(t) + \bar{C}_i \dot{y}_i(t) + \bar{K}_i y_i(t) = \bar{R}_i(t) \tag{19}$$

where

$$\bar{M}_i = \Phi_i^T M \Phi_i, \ \bar{C}_i = \Phi_i^T C \Phi_i, \ \bar{K}_i = \Phi_i^T K \Phi_i \text{ and } \bar{R}(t) = \Phi_i^T R(t) \tag{20}$$

are the generalized values of the corresponding matrices and the loading vector, while Φ_i is the i-th eigenmode shape matrix. According to the modal superposition approach the system of N differential equations, which are coupled with the off-diagonal terms in the mass, damping and stiffness matrices, is transformed to a set of N independent normal-coordinate equations. The dynamic response can therefore be obtained by solving separately for the response of each normal (modal) coordinate and by superposing the response in the original coordinates.

In the mMRS analysis a number of different formulas have been proposed to obtain reasonable estimates of the maximum response based on the spectral values without performing time history analyses for a considerable number of transformed dynamic equations. The simplest and most popular one is the Square Root of Sum of Squares (SRSS) of the modal responses. According to this estimate the maximum total displacement is approximated by:

$$u_{max} = \left(\sum_{i=1}^{N} u_{i,max}^2 \right)^{1/2} \qquad (21)$$
$$u_{i,max} = \Phi_i y_{i,max}$$

where $u_{i,max}$ corresponds to the maximum displacement vector corresponding to the i-th eigenmode.

4.3 The ANN training for seismic reliability analysis

In the present implementation the main objective is to investigate the ability of the ANNs to predict the structural performance in terms of maximum inter-storey drift. These objectives comprise the following tasks: (i) Select the proper training set. (ii) Find suitable network architecture. (iii) Determine the appropriate values of characteristic parameters of ANN. For the back propagation algorithm to provide good results the training set must include data over the entire range of the output space. The appropriate selection of I/O training data is one of the important factors in ANN training. Although the number of training patterns may not be the only concern, the distribution of samples is of greater importance. The selection of the I/O training pairs is based on the requirement that the full range of possible results should be represented in the training procedure [11]. In the present study the sample space for each random variable is divided into equally spaced distances for the application of the ANN simulation and for the selection of the suitable training pairs. After the selection of the suitable ANN architecture and the performance of the training procedure, the network is then used to produce predictions of structural failure corresponding to different values of the input random variables. The results are then processed by means of MCS to calculate the probability of failure p_f using Eq. (15).

The modulus of elasticity, the dimensions b and h of the I-shape cross-section and the earthquake loading have been considered as random variables. In this study three test cases have been taken into account depending on the random variables considered. In the first test case the modulus of elasticity and the earthquake loading are considered to be random variables. In the second test case the dimensions b and h of the I-shape cross section and earthquake loading are taken as random variables, while in the third test case all three groups of random variables are considered. For the implementation of the ANN-based metamodel in all three test cases a two level approximation is employed using two different ANNs. The first ANN predicts the values of the significant eigen-periods, the inputs of the ANN are the random variables while the outputs are the values of the significant eigen-periods. The second ANN is used to predict the maximum inter-storey drift, which is used to define the failure or not of each simulation required by MCS (Eq. (15)), the spectral acceleration values both in X and Y directions are the inputs of the ANN while the maximum inter-storey drift is the output.

5 Numerical examples

5.1 Defect identification in a composite beam using static measurements

This test example consists of a clamped beam subjected to a distributed loading. The beam is equipped with piezoelectric actuators and sensors, which can be used for defect identification (see schematic configuration in Figure 1). Measurements of the deflections from the piezoelectric sensors are used for the identification of the defect

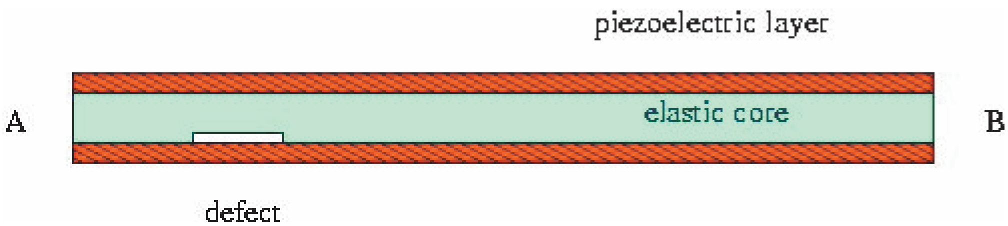

Figure 1: Configuration of a composite plate with a defect

The finite element model has been developed and tested for static and dynamic optimal control problems in [28] and [29], respectively, where technical details can be found. The beam has been discretized with 30 finite elements. A defect is approximately modelled by reduction of the stiffness in the corresponding finite element with a smeared-crack-like approach. The first assumption is that the complete deformation of the beam can be measured by using the distributed piezoelectric sensors or other suitable measurements. The second assumption is that only one defect will arise, therefore the two unknown variables of the defect identification problem are: (a) the position (number of damaged element) and (b) the extent of the defect. For the position it is assumed that this continuous variable takes values in the interval [1,30], which indicates the damaged element (after appropriate trimming to give the discrete value corresponding to the element number). The extent of damage is approximated by a continuous variable, taking values in the interval [0,10], which is linearly dependent on the extent of damage (between 0 and 50 percent of the nominal stiffness, respectively).

The direct approximation of the inverse mapping (i.e. the relation between measurements and damage variables) with neural networks does not lead to useful results. The genetic algorithm solves the inverse problem. Although this problem is rather small, the mechanical problem can be replaced by a response function approach, which is based either on interpolation functions or on neural networks. The accuracy of the inverse problem is not influenced by either of the two approaches while the required computational time is reduced. For a constant population size equal to 15 the solution of the inverse problem is documented in the following Table 1. The fitness variable transforms the error minimization problem to a maximization problem and includes a logarithmic scaling. The fitness maximization for this example is demonstrated in Figure 2 (where the best value within each population is depicted with red, and the mean value with yellow, respectively).

Example No	Exact position of defect	Calculated position of defect	Exact extent of damage (%)	Calculated extent of damage	Fitness value
1	5	4.9860	10	9.9915	42.9395
2	15	15.1133	10	10.078	41.0051
3	25	25.2182	10	10.463	42.9711

Table 1: Identification of the position and size of one defect in the composite beam using static measurements

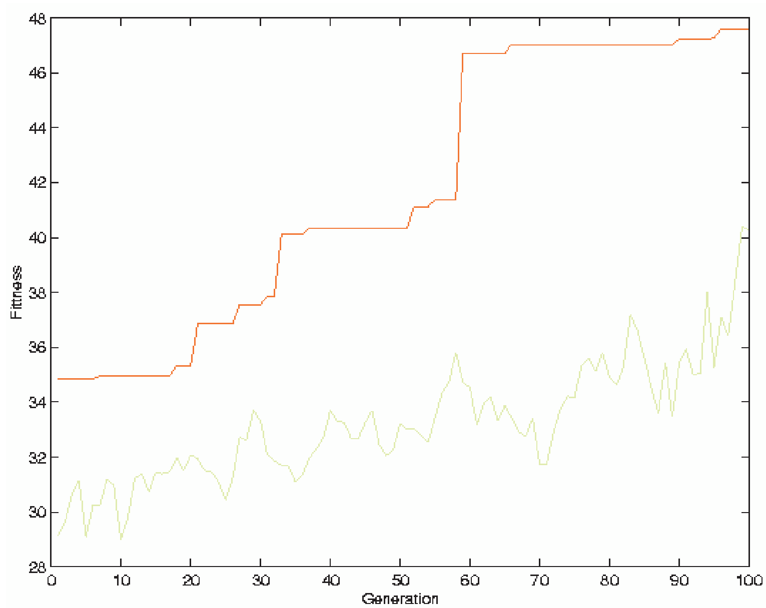

Figure 2: Documentation of one genetic optimization solution with best individual and mean value for each generation

In this example all vertical displacements of the beam, i.e. 30 measurements, have been taken into account for the solution of the inverse problem. By reducing the number of measurements (again, equally distributed along the length of the cantilever) the solution of the inverse problem is not affected, as it can be shown in Table 2.

Example No.	Number of measurements	Calculated position of defect	Calculated extent of defect	Fitness value
1	30	15.0009	25.0035	45.4718
2	15	14.9783	24.9285	41.8466
3	7	15.0058	25.0155	46.3440

Table 2: Damage identification of a defect at element no. 15 with extent of damage equal to 25 per cent loss of stiffness: effect of reduced number of measurements

5.2 Crack identification problem

In this example an orthogonal plane stress domain rigidly supported at the lower boundary and loaded at the upper boundary is considered. From measurements at the free boundaries of the plate one can identity cracks and holes in the interior of the plate [23]. A schematic description of the problem is shown in Figure 3.

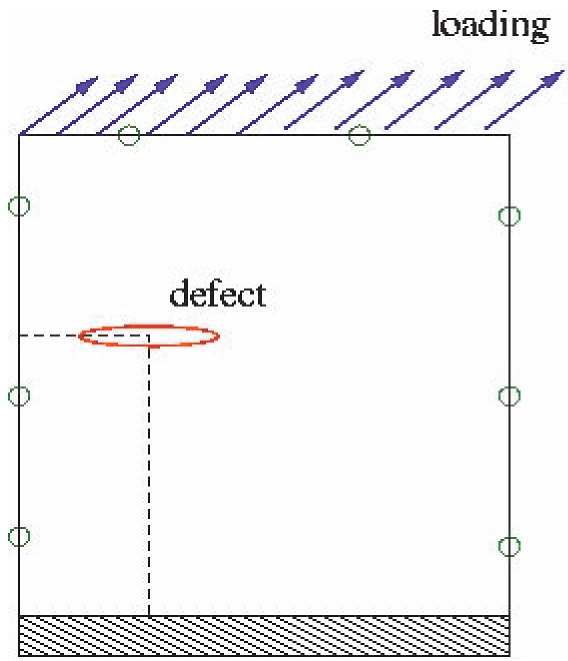

Figure 3: Configuration of a plate with a defect. Measurements at the external boundary can be used for defect identification

At first one example of the neural network based solution of the inverse problem is presented. Using a set of 206 measurements of displacements at the boundary of the plate in several representative time instances a suitably trained 206-7-2 neural network can be trained to reproduce the position of a given crack (horizontal, of given length) with acceptable accuracy. Here a boundary element model is used for the production of the training and test examples. A set of 81 training and 64 test examples has been used. In the following figures (Fig. 4 to Fig. 10) the training procedure, the accuracy of predicting the training data and of the test data are demonstrated.

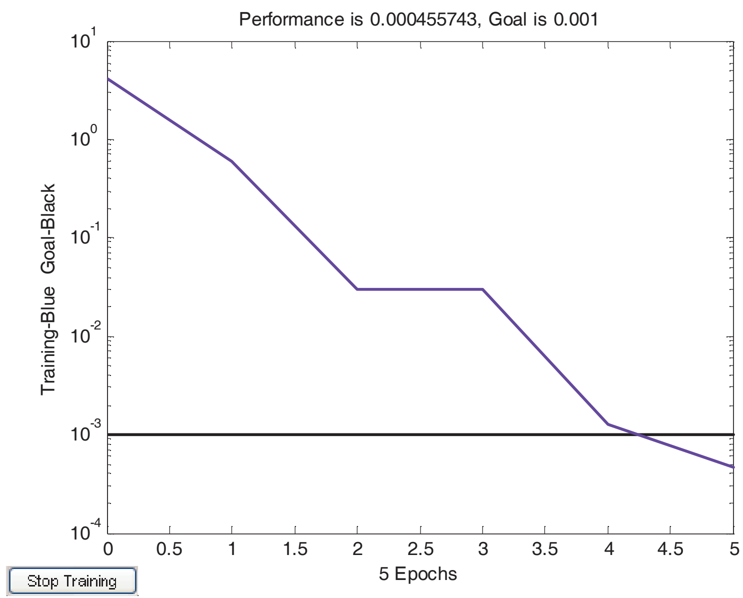

Figure 4: Training of the neural network for the defect identification problem

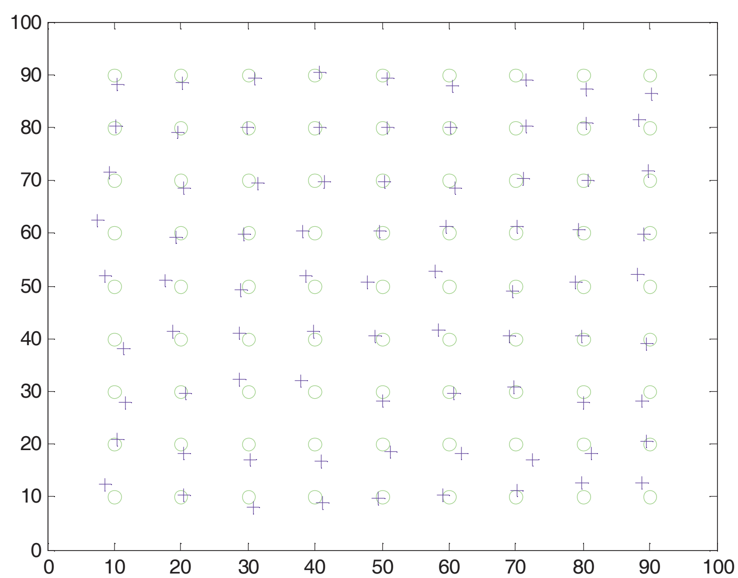

Figure 5: Training data: Predicted (+) and real (o) position of the defect for different places of the defect

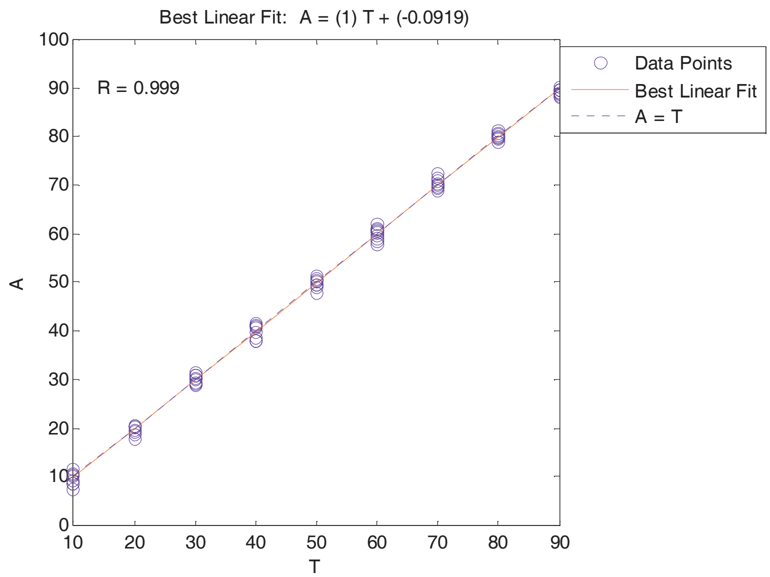

Figure 6: Training data: Accuracy of x-coordinate prediction

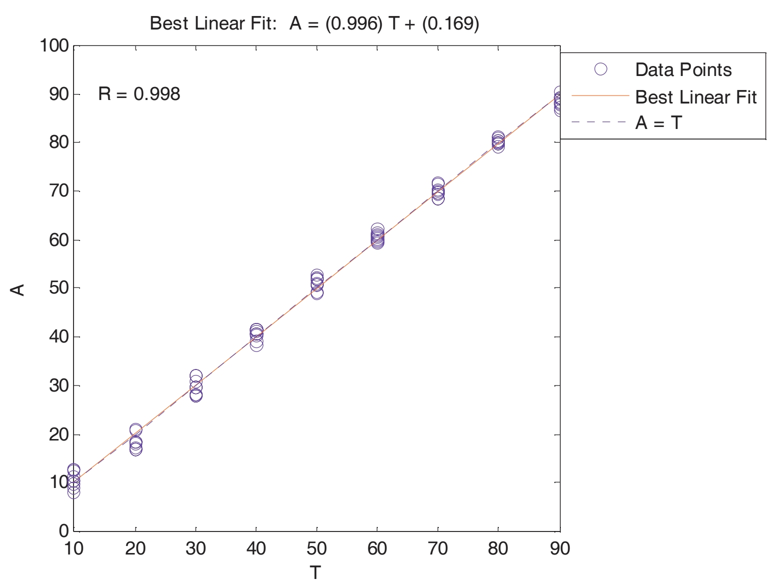

Figure 7: Training data: Accuracy of y-coordinate prediction

In order to enhance the accuracy of the results the same problem is solved with the genetic algorithm optimization procedure. In fact, in more complicated forms of the plate, only the genetic optimization can solve the inverse problem (see [22] for relevant investigations). For the technical realization of this concept the genetic optimization tool can be combined with the mechanical modeling software, if this is possible (for example if the analysis code is open, it is written in a compatible computer language, etc). The requirements on computer time and storage may become very large. Alternatively one may use the proposed hybrid strategy. First, a number of small neural networks approximate the response of each boundary point of interest as a function of the crack position (for all other parameters of the mechanical problem, like loading or boundary conditions, fixed). A similar performance can be obtained by using classical interpolation functions. The genetic optimization uses then the approximators to construct the fitness function and proceeds with the solution of the output error minimization problem and, eventually, of the inverse problem as in the previous example.

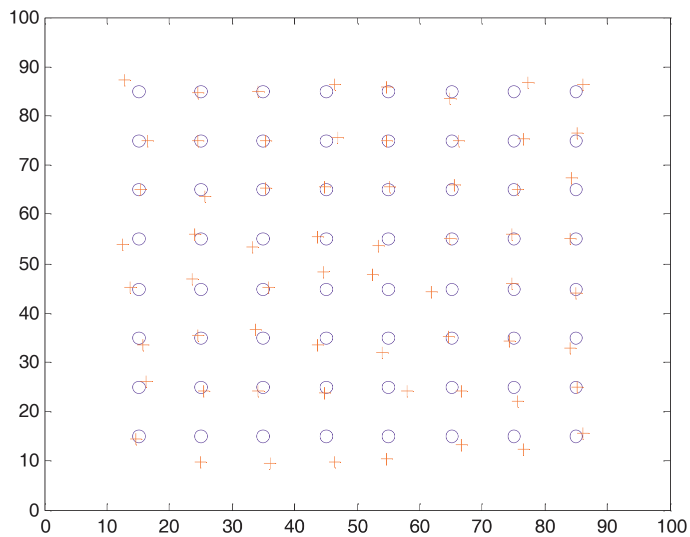

Figure 8: Test data: Predicted (+) and real (o) position of the defect for different places of the defect

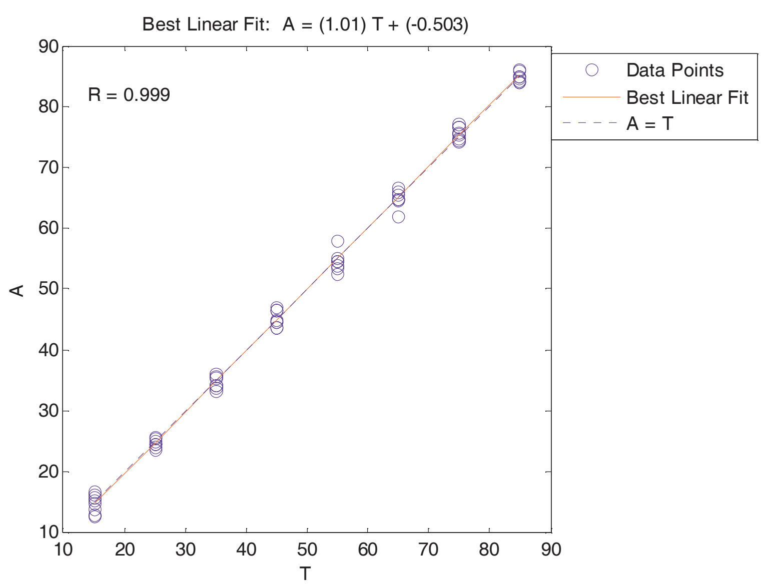

Figure 9: Test data: Accuracy of x-coordinate prediction

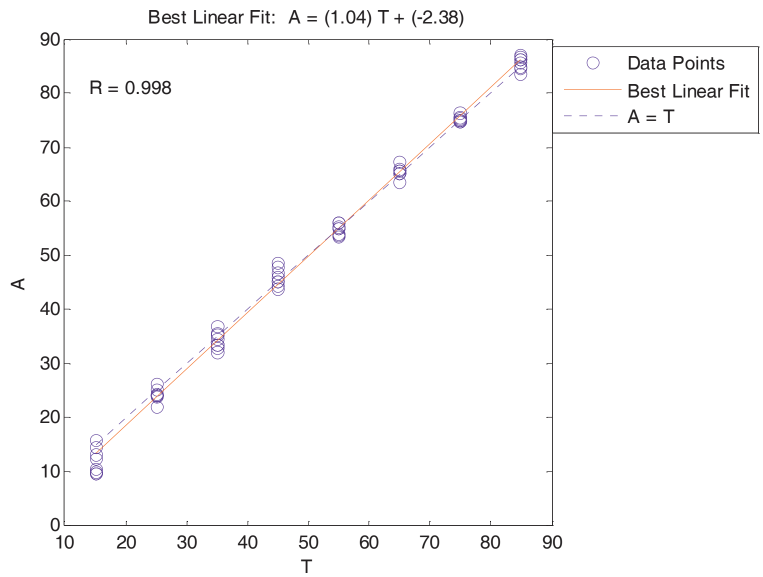

Figure 10: Test data: Accuracy of y-coordinate prediction

5.3 Structural reliability test example

The six storey space frame, shown in Figure 11**,** has been considered for the purpose of the current study in order to assess the proposed metamodel assisted structural reliability analysis methodology. The space frame consists of 63 structural elements which are divided into five groups having the following cross sections: 1) IPB 650, 2) IPB 650, 3) IPE 450, 4) IPE 400 and 5) IPB 450. The structure is loaded with a permanent action of G = 3 kPa and a variable action of Q = 5 kPa. In addition, in order to take into account the non-linear behavior of the structure when the mMRS method is used the seismic loads are reduced by the behavior factor q = 4.0 as Eurocede 8 (EC8) suggests for steel frames [30]. The three test cases corresponding to different combinations of random variables are considered here.

The most common approach for the definition of the seismic input is the use of design code response spectrum. This is a general approach, which is easy to implement. However, if higher precision and more realistic simulation of the structural seismic response is required, the use of spectra derived from natural earthquake records is more appropriate. In order to avoid significant dispersion on the structural response, due to the use of different natural records, these spectra must be scaled to the same desired earthquake intensity. The most commonly applied scaling procedure is based on the peak ground acceleration (PGA).

In this study a set of nineteen natural accelerograms, shown in Table 3, is used. Each record corresponds to different earthquake magnitudes and soil properties. These time histories are from different earthquakes. Two are from the 1992 Cape Mendocino earthquake, two are from the 1978 Tabas, Iran earthquake and fifteen are from the 1999 Chi-chi, Taiwan earthquake. The records are scaled, to the same peak ground acceleration of 0.32g in order to ensure compatibility between the records. The response spectra for each scaled record, in x and y directions, are shown in

Figures 12 and 13, respectively. The type of probability density functions, mean values and standard deviations for all variables are presented in Table 4.

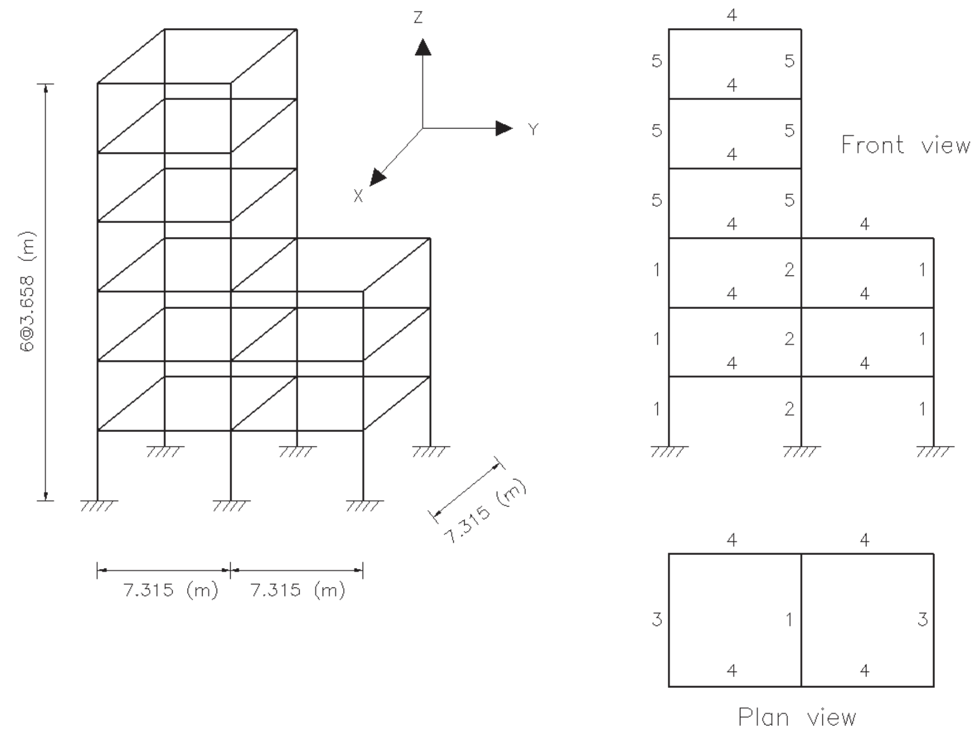

Figure 11: The six-storey space frame

It has been observed that the response spectra follow the lognormal distribution [31]. Therefore the median spectrum $\hat{x}$, also shown in Figures 12 and 13, and the standard deviation δ are calculated from the above set of spectra using the following expressions

$$\hat{x} = \exp\left[\frac{\sum_{i=1}^{n} \ln(R_{d,i}(T))}{n}\right] \tag{22}$$

$$\delta = \left[\frac{\sum_{i=1}^{n} \left(\ln(R_{d,i}(T)) - \ln(\hat{x})\right)^2}{n-1}\right]^{1/2} \tag{23}$$

where $R_{d,i}(T)$ is the response spectrum value for period equal to T of the ith record (i=1,…,n, where n=19 in this study). It has been observed that for the six storey test example considered, the number of significant eigen-periods is eight. For a given period value, the acceleration R_d is obtained as a random variable following the log-normal distribution whose mean value is equal to $\hat{x}$ and standard deviation is equal to δ.

Earthquake	Station	Distance	Site
Tabas 16 Sept. 1978	Dayhook	14	rock
	Tabas	1.1	rock
Cape Mendocino 25 April 1992	Cape Mendocino	6.9	rock
	Petrolia	8.1	soil
Chi-Chi 20 Sept. 1999	TCU052	1.4	soil
	TCU065	5.0	soil
	TCU067	2.4	soil
	TCU068	0.2	soil
	TCU071	2.9	soil
	TCU072	5.9	soil
	TCU074	12.2	soil
	TCU075	5.6	soil
	TCU076	5.1	soil
	TCU078	6.9	soil
	TCU079	9.3	soil
	TCU089	7.0	rock
	TCU101	4.9	soil
	TCU102	3.8	soil
	TCU129	3.9	soil

Table 3: List of the natural records

Random variable	**Probability density function**	**Mean value**	**Standard deviation**
E	N	210	10 (%)
b	N	b*	2 (%)
h	N	h*	2 (%)
Seismic load	Log-N	$\hat{x}$ (Eq. 22)	δ (Eq. 23)

* dimensions from the IPE and HEB databases

Table 4: Characteristics of the random variables

In the first test case, where the modulus of elasticity and the earthquake loading are considered as random variables, one hundred values of the modulus of elasticity are selected in order to train the first ANN and one hundred combinations of the spectral values are selected for the training of the second ANN. For each ANN ten of these combinations are selected for testing the generalization capabilities of the trained ANNs. In the second test case the training-testing set was composed by one hundred fifty pairs, while in the case of the third test case the set was composed by two hundred pairs. The neural network configurations used are the following: (i) NN1: 1-10-8, NN2: 16-20-1, (ii) NN1: 10-20-8, NN2: 16-20-1 and (iii) NN1: 11-10-8, NN2: 16-20-1, for the three test cases respectively.

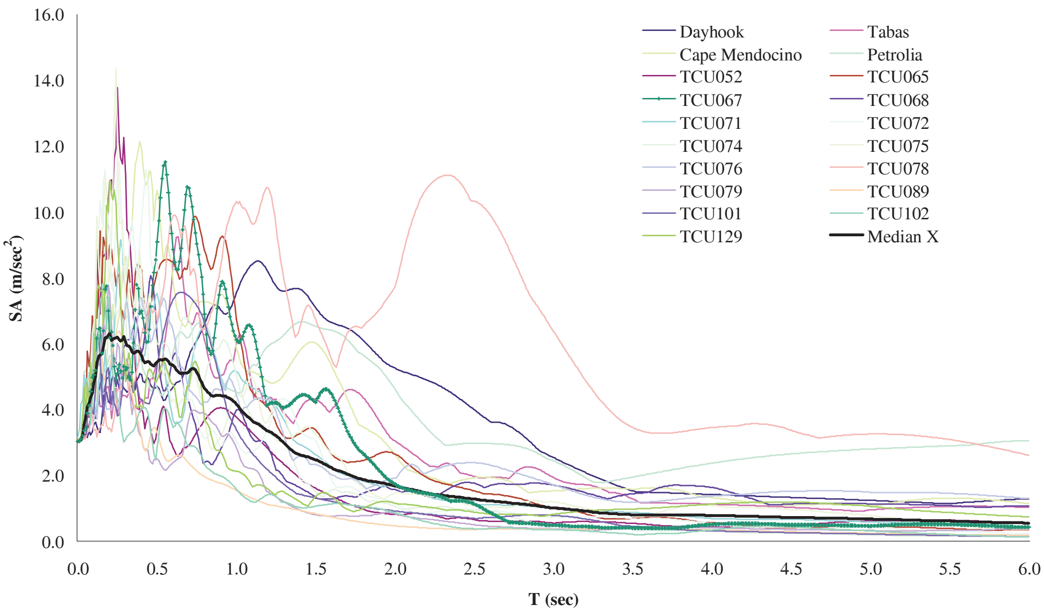

Figure 12: Natural record response spectra and their median (x axis)

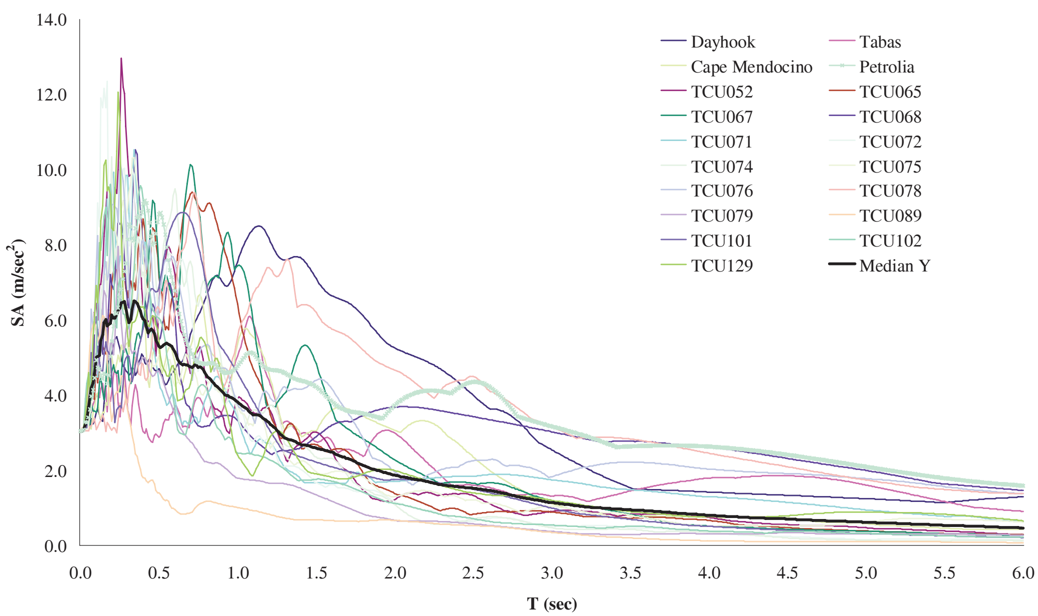

Figure 13: Natural record response spectra and their median (y axis)

The influence of the three groups of random variables with respect to the number of simulations is show in Figure 14. It can be seen that 20 to 50 thousand simulations are required in order to calculate with an adequate accuracy the target probability of failure. It is also observed that considering both modulus of elasticity and the cross section dimensions as random variables leads to an increase of 7% of the value of the probability of failure (1.37% for the third test case instead of 1.28% of the first and 1.24% of the second test case).

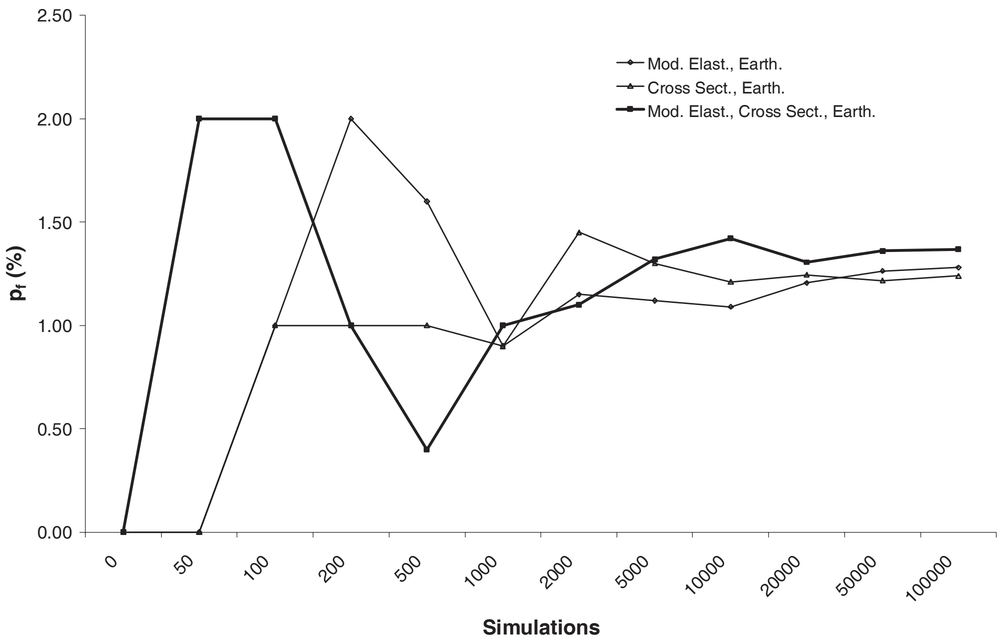

Figure 14: Influence of the number of MC simulations on the value of p_f for the three test cases

Once an acceptably trained ANN for predicting the maximum drift is obtained, the probability of failure for each test case is estimated by means of ANN based MCS. The results for various numbers of simulations are depicted in Table 5 for the three test cases examined. From these results it can be observed that, in the case of basic MCS, the error of the predicted probability of failure with respect to the "exact" one is rather marginal. On the other hand, the computational cost is drastically decreased, approximately 30 times, for all test cases.

6 Conclusions

The paper presents applications of hybrid techniques, involving ANNs, in computationally demanding tasks in mechanics. For inverse and parameter identification problems there exist, in principle, the possibility to use directly ANNs for the approximation of the inverse structural mapping. Nevertheless, this mapping becomes complicated for real-life applications and the training of the ANNs may be difficult or even impossible. On the other hand, the direct mapping relating unknown parameters with structural response is much easier to be approximated. This can be done very efficiently with ANNs. The resulting trained ANN can be combined with a powerful numerical optimization algorithm, like genetic algorithms, for the solution of the inverse problem. This approach has been demonstrated in this paper with both static and dynamic applications.

	Test case 1		Test case 2		Test case 3	
Number of simulations	**"exact" p_f (%)**	**NN p_f (%)**	**"exact" p_f (%)**	**NN p_f (%)**	**"exact" p_f (%)**	**NN p_f (%)**
50	0.00	0.00	0.00	0.00	2.00	0.00
100	1.00	2.00	1.00	2.00	2.00	1.00
200	2.00	1.60	1.00	1.00	1.00	3.00
500	1.60	1.26	1.00	1.70	0.40	0.70
1,000	0.90	0.81	0.90	0.67	1.00	0.86
2,000	1.15	1.06	1.45	1.41	1.10	0.97
5,000	1.12	1.04	1.30	1.24	1.32	1.18
10,000	1.09	1.03	1.21	1.14	1.42	1.29
20,000	1.21	1.14	1.25	1.31	1.31	1.19
50,000	1.26	1.16	1.22	1.31	1.36	1.21
100,000	1.28	1.16	1.24	1.31	1.37	1.21
CPU time (sec)						
Pattern selection	-	2	-	3	-	5
Training	-	7	-	9	-	12
Propagation	-	25	-	25	-	25
Total	1,154	34	1,154	37	1,154	42

Table 5: "Exact" and predicted values of p_f and the required CPU time

On the other hand the computational effort involved in the conventional MCS becomes excessive in large-scale problems, especially when earthquake loading is considered, because of the enormous sample size and the computing time required for each Monte Carlo run. The use of ANNs can practically eliminate any limitation on the scale of the problem and the sample size used for MCS.

Acknowledgements

Partial support from the Greek-Italian Bilateral Scientific Cooperation Project is gratefully acknowledged.

References

[1] G.E. Stavroulakis, G. Bolzon, Z. Waszczyszyn, L. Ziemianski, ''Inverse Analysis'', in ''Comprehensive Structural Integrity'', Eds in Chief: B. Karihaloo, R.O. Ritchie, I. Milne, ''Volume 3: Numerical and Computational Methods'', Volume Eds: R. de Borst, H.A. Mang, Elsevier Publishers, Chapter 13, 685-718, 2003.

[2] K.R. Gurley, M. Tognarelli, A. Kareem, ‘‘Analysis and simulation tools for wind engineering’’, Prob. Engrg. Mech., 12, 9-31, 1997.
[3] H. Adeli, H.S. Park, ‘‘Optimization of space structures by neural networks’’, J. of Neural Networks, 8, 769-782, 1995.
[4] D.S. McCorkle, K.M. Bryden, C.G. Carmichael, ‘‘A new methodology for evolutionary optimization of energy systems’’, Comp. Meth. Appl. Mech. Engrg., 192, 5021-5036, 2003.
[5] M. Papadrakakis, N.D. Lagaros, ‘‘Reliability-based structural optimization using neural networks and Monte Carlo simulation’’, Comp. Meth. Appl. Mech. Engrg., 191, 3491-3507, 2002.
[6] S. Sakata, F. Ashida, M. Zako, ‘‘Structural optimization using Kriging approximation’’, Comp. Meth. Appl. Mech. Engrg., 192, 923-939, 2003.
[7] L. Zhang, G. Subbarayan, ‘‘An evaluation of back-propagation neural networks for the optimal design of structural systems: Part I. Training procedures’’, Comp. Meth. Appl. Mech. Engrg., 191, 2873-2886, 2002.
[8] G.R. Liu, Y.G. Xu, Z.P. Wu, ‘‘Total solution for structural mechanics problems’’, Comp. Meth. Appl. Mech. Engrg., 191, 989-1012, 2001.
[9] J. Zacharias, C. Hartmann, A. Delgado, ‘‘Damage detection on crates of beverages by artificial neural networks trained with finite-element data’’, Comp. Meth. Appl. Mech. Engrg., 193, 561-574, 2004.
[10] J.E. Hurtado, D.A. Alvarez, ‘‘Neural-network-based reliability analysis: a comparative study’’, Comp. Meth. Appl. Mech. Engrg., 191, 113-132, 2002.
[11] M. Papadrakakis, V. Papadopoulos, N.D. Lagaros, ‘‘Structural reliability analysis of elastic-plastic structures using neural networks and Monte Carlo simulation’’, Comp. Meth. Appl. Mech. Engrg., 136, 145-163, 1996.
[12] J.N. Hwang, ‘‘Nonparametric Multivariate Density Estimation: A Comparative Study’’, IEEE Trans. Signal Processing, 42, 2795-2810, 1994.
[13] Y. Tsompanakis, N.D. Lagaros, M. Papadrakakis, ‘‘Reliability analysis of structures under seismic loading’’, in H.A. Mang, F.G. Rammerstorfer, J. Eberhardsteiner (Eds.), Proc. 5th World Congress on Computational Mechanics WCCM-V, Vienna, Austria, July 7-12, 2002.
[14] G.E. Stavroulakis, H. Antes, ‘‘Flaw identification in elastomechanics: BEM simulation with local and genetic optimization’’, Str. Opt., 16(2/3), 162-175, 1998.
[15] G.E. Stavroulakis, ‘‘Inverse and crack identification problems in engineering mechanics’’, Habilitation Thesis, Technical University Braunschweig. Kluwer Academic Publishers - Springer Verlag, Dordrecht, Boston, London, 2000.
[16] M. Engelhardt, A. Likas, G.E. Stavroulakis, ‘‘Neural crack identification’’, in H.A. Mang, F.G. Rammerstorfer, J. Eberhardsteiner (Eds.), Proc. 5th World Congress on Computational Mechanics WCCM-V, Vienna, Austria, July 7-12, 2002.
[17] W. Schiffmann, M. Joost, R. Werner, ‘‘Optimization of the back-propagation algorithm for training multi-layer perceptrons’’, Technical Report, Institute of Physics, University of Koblenz, 1993.

[18] N.D. Lagaros, M. Papadrakakis, ‘‘Learning improvement of neural networks used in structural optimization’’, Adv. Engrg. Soft., 35, 9-25, 2004.
[19] S. Fahlman, ‘‘An empirical study of learning speed in back-propagation networks’’, Technical Report, Carnegie Mellon: CMU-CS-88-162, 1988.
[20] M. Riedmiller, H. Braun, ‘‘A direct adaptive method for faster back-propagation learning: The RPROP algorithm’’, in H. Ruspini (Ed.), Proc. of the IEEE International Conference on Neural Networks (ICNN), San Francisco, 586-591, 1993.
[21] M. Riedmiller, ‘‘Advanced Supervised Learning in Multi-layer Perceptrons: From Back-propogation to Adaptive Learning Algorithms’’, Technical Report, University of Karlsruhe: W-76128 Karlsruhe, 1994.
[22] M. Engelhardt, ‘‘Numerische Verfahren zur Identifizierung von Fehlstellen aus Randdaten’’, PhD Thesis, Carolo Wilhelmina Technical University, Braunschweig, Germany, Braunschweig Series on Mechanics No. 56, 2004.
[23] G.E. Stavroulakis, M. Engelhardt, A. Likas, R. Gallego, H. Antes, ‘‘Neural network assisted crack and flaw identification in transient dynamics’’, J. Theor. Appl. Mech., Polish Academy of Sciences, 42(3), 629-649, 2004.
[24] M. Engelhardt, G.E. Stavroulakis, H. Antes, ‘‘Crack and flaw identification in elastodynamics using Kalman filter techniques’’, Comp. Mech., 2005 (in press).
[25] A. Oishi, K. Yamada, A. Yoshimura, G., Yagawa, ‘‘Quantitative nondestructive evaluation with ultrasonic method using neural networks and computational mechanics’’, Comp. Mech., 15, 521-533, 1995.
[26] M. Kitahara, J.D. Achenbach, Q.C. Guo, M.L. Peterson, T. Ogi, M. Notake, ‘‘Depth determination of surface-breaking cracks by a neural network’’, Review Prog. Qual. Nondestr. Eval., 10, 689-696, 1991.
[27] M. Papadrakakis, Y. Tsompanakis, N.D. Lagaros, M. Fragiadakis, ‘‘Reliability based optimization of steel frames under seismic loading conditions using evolutionary computation’’, J. Theor. Appl. Mech., Polish Academy of Sciences, 42(3), 585-608, 2004.
[28] E.P. Hadjigeorgiou, G. Foutsitzi, G.E. Stavroulakis, ‘‘Shape control of beams with piezoelectric actuators’’, HERMIS Int. J., 5, 65-78, 2004.
[29] G.E. Stavroulakis, G. Foutsitzi, E. Hadjigeorgiou, D. Marinova, C.C. Baniotopoulos, ‘‘Design and robust optimal control of smart beams with application on vibration suppression’’, Adv. Engrg. Soft., 2005 (in press).
[30] ‘‘Eurocode 8: Design provisions for earthquake resistant structures’’, CEN, ENV, 1998-1-1/2/3, 1994.
[31] C. Chintanapakdee, A.K. Chopra, ‘‘Evaluation of modal pushover analysis using generic frames’’, Earth. Engrg. & Str. Dyn., 32, 417-442, 2003.