COMPDYN 2015
5th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, V. Papadopoulos, V. Plevris (eds.)
Crete Island, Greece, 25–27 May 2015

A MESHLESS NEURAL NETWORK APPROACH FOR THE DYNAMIC ANALYSIS OF ELASTIC SYSTEMS

Luca Facchini[1] **and Michele Betti**[2]

[1] Department of Civil and Environmental Engineering,
University of Florence
via di S. Marta 3 – I-50139 Florence – Italy
e-mail: luca.facchini@unifi.it

[2] Department of Civil and Environmental Engineering,
University of Florence
via di S. Marta 3 – I-50139 Florence – Italy
e-mail: mbetti@dicea.unifi.it

Keywords: Static and dynamic problem in linear elasticity, Artificial Neural Networks, Meshless methods, Mixed formulation, Hu-Washizu principle, Energy-based training.

Abstract. *A meshless approach is presented for the computation of the approximated solution of static and dynamic problems in linear elasticity in terms of displacement fields. The displacement field is modeled by means of two different kinds of Artificial Neural Networks (ANN). This task is accomplished by means of a meshless approach coupled to a net training based on the weak formulation of the differential problem, related to Hu-Washizu principle. A common benchmark, namely the Timoshenko cantilever beam, is analyzed and discussed in detail; several researches have shown that severe difficulties are encountered with the Galerkin and the collocation approach since the neural networks never satisfy essential boundary conditions (EBC): in the proposed meshless approach, the trial functions can be modified in order to satisfy EBC. A possibility to overcome such difficulty in elasticity is to employ an energy-based training, that is, to employ an approach (such as the Hu-Washizu functional) which can take into account EBC in the error function to be minimized. An example is given in one dimension, analyzing the deflection of a horizontal beam subject to transverse loads. The presented examples clearly show the importance of the optimization of the non-linear pa-rameters of the network, which control the shape and location of the activation functions. It is shown that such parameters can be optimized for a static problem and subsequently employed for a dynamic problem. The paper in fact aims to extend the results investigating the bench-mark problem in the dynamic field.*

1 INTRODUCTION

Meshless approaches which are based on Artificial Neural Networks (ANN) have lately become widely used and investigated by the scientific community to find reliable approximations for the solutions of a great number of problems, described by one or more differential equations. Extensive summaries can be found f.i. in [1] and [2].

In the field of structural mechanics one of the most promising perspectives is the analysis of cracked structures and more generally the behavior of materials which cannot undergo tensile states. One of the most attractive features is that such approaches do not need the definition of a mesh, and many of them need a so-called background mesh of the structure for only numerical integration purposes. This leads, especially in the field of fracture mechanics, to model the cracked materials in a more reliable manner, as the crack pattern computed with a FEM-based algorithm often shows a strong sensitivity in the mesh definition in the surroundings of the crack itself.

Usually, meshless approaches employ a weighted residual method (see f.i. [3] and [4]) to evaluate the solution; the activation functions are defined by means of a set of parameters – called non-linear parameters in the following – which control the shape and location of the support of the functions. The present work emphasizes the importance of the optimization of such parameters, showing that the non-linear parameters can be optimized with respect to a static problem, and subsequently employed to determine the dynamic characteristics of the investigated structure.

2 THEORETICAL REMARKS

A great amount of different kinds of neural networks can be found in specialized literature, but many of them can be represented as a linear combination of nonlinear functions of their arguments in the form

$$\boldsymbol{f}(\boldsymbol{x}) = \sum_{h=1}^{N_\phi} \boldsymbol{w}^{(h)} \phi_h(\boldsymbol{x}) \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ is the independent variable and $\boldsymbol{f}: \mathbb{R}^n \to \mathbb{R}^m$ is a vector function. It follows from the definition (1) that $\boldsymbol{w}^{(h)} \in \mathbb{R}^m$ are vectors coefficients and $\phi_h(x)$ are scalar functions of $\boldsymbol{x}$.

Typically, functions $\phi_h(x)$ determine the kind of neural network; in the present work, examples are presented which employ either sigmoid functions (utilized in feed-forward back-propagation – FFBP – networks) or radial basis functions (RBF, giving rise to RBF networks).

2.1 RBF Networks

In the meshless approaches for structural mechanics, RBF are typically employed, such as multi-quadric, Gaussian bells, or compact support functions as the so-called Chebichev's hat or similar functions. RB functions can be expressed by means of the combination of the distance of the generic point $\boldsymbol{x}$ from the so-called center of the associated function, and a real-valued function of a scalar variable which decreases for increasing argument (see f.i. [5], [6] and [7]):

$$\begin{cases} \phi_h(\boldsymbol{x}) = \varphi\left(\| \boldsymbol{x} - \boldsymbol{c}^{(h)} \|\right) \\ \varphi(r) = \exp(-r^2/\sigma^2) \end{cases} \tag{2}$$

In expression (2), $\varphi(r)$ can be seen as an *underlying* function which enables to define the RBFs $\phi_h(\boldsymbol{x})$, and in this case a Gaussian bell is employed (Figure 1); usually, compact support functions are preferred to other kind of functions owing to the difficulty involved in the calculation of integrals, but such difficulty can be overcome by means of proper numerical techniques.

As it is clear from definitions, RB functions depend on a limited number of parameters which control their shape and location. In information theory, trial functions are in fact called activation functions, as they change their value in a limited interval of the real axis: this means that, if the independent variable $\boldsymbol{x}$ falls into a limited region $\Omega_h \in \mathbb{R}^n$, then the activation function attains non-negligible values.

The region Ω_h is typically a hyper-sphere in $\mathbb{R}^n$ in the case of RBFs, so that its characterizing parameters are the coordinates of its center and its radius. Nevertheless, it can also attain different shapes, as its expression is mainly a function of the distance of the generic point $\boldsymbol{x}$ from the so-called center of the associated function as expressed in (2).

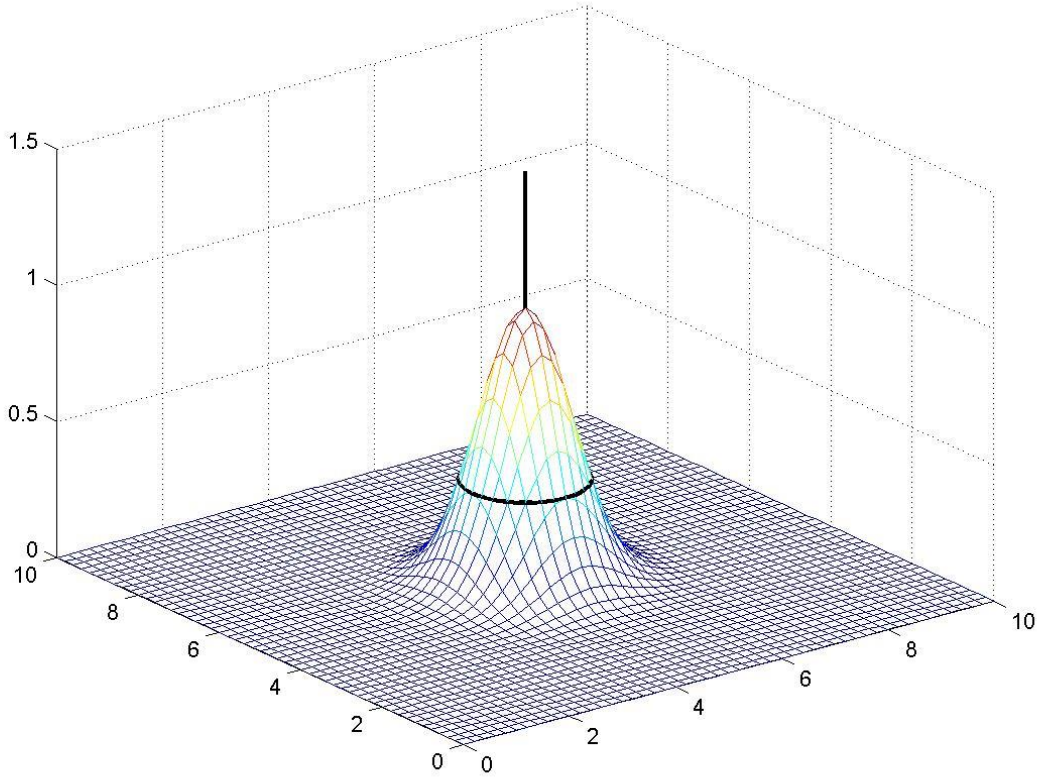

Figure 1: A typical RB function defined by means of a Gaussian bell; the center is located at coordinates (5,5) and is highlighted by the black vertical line, while the parameter σ determines the radius of the black circle.

An interesting example of an RB function is described in [5] and called the *up function*; such function is compactly supported and, if properly scaled and overlapped, possesses the following notable characteristics:

$$\sum_h u_h(x) = 1 \quad ; \quad \sum_h x_h u_h(x) = x \tag{3}$$

2.2 Feed-forward networks

An interesting approach is represented by FFBP networks which make use of sigmoid trial functions: these are monotonically increasing functions of the independent variable, and among the most used are:

$$\text{(sigmoid function)} \begin{cases} \phi_h(\boldsymbol{x}) = \varphi(\sigma_{hk} x_k + b_h) \\ \varphi(r) = \frac{1}{1+\exp(-r)} \end{cases} \tag{4}$$

$$\text{(hyperbolic tangent function)} \begin{cases} \phi_h(\boldsymbol{x}) = \psi(\sigma_{hk} x_k + b_h) \\ \psi(r) = \frac{\exp(r)-\exp(-r)}{\exp(r)+\exp(-r)} \end{cases} \tag{5}$$

The use of functions (4) or (5) as activation functions is equivalent: as a matter of fact, it can be shown that:

$$\psi(r) = \frac{1-\exp(-2r)}{1+\exp(-2r)} = 2\phi(2r) - 1 \tag{6}$$

2.3 Meshless approach for static problem

The described functions are herein utilized to obtain a satisfactory approximation of a linear elastostatic problem, cast in the following terms: given a linearly elastic continuum B subject to body forces $\boldsymbol{b}$ and to surface forces $\hat{\boldsymbol{f}}$ on the boundary ∂B_N (natural boundary conditions, NBC), and to restraints $\hat{\boldsymbol{s}}$ on the boundary ∂B_E, find a displacement field $\boldsymbol{s}(\boldsymbol{x})$ such that:

$$\text{(essential boundary conditions, EBC)} \quad \boldsymbol{s}(\boldsymbol{x}) = \hat{s} \quad \text{on } \partial B_E \tag{7}$$

$$\text{(strain-displacement relation)} \quad \boldsymbol{\epsilon} = Sym(\nabla \boldsymbol{s}) \tag{8}$$

$$\text{(equilibrium equations)} \begin{cases} \text{div}\boldsymbol{\sigma} + \boldsymbol{b} = \boldsymbol{0} & \text{on } B \\ \boldsymbol{\sigma n} = \hat{\boldsymbol{f}} & \text{on } \partial B_N \end{cases} \tag{9}$$

where $\boldsymbol{\sigma} = C[\boldsymbol{\epsilon}]$, C being the elastic tensor. Even though it is simpler than the dynamic case, the solution of the elastostatic problem can nevertheless give some insight on the main issues of the solution procedure: usually, a weighted residual procedure is adopted to determine the unknown parameters $\boldsymbol{w}^{(h)}$ in (1), while the parameters which control the shape and the collocation of the generic function ϕ_h are decided and fixed by the user, and the goodness of the obtained approximation mainly depend on the number N_ϕ of functions chosen.

Some indications can be found in specialized literature on the best possible way to choose the location and shape parameters of the activation functions. As an example, Peng *et al.* [8] observe that a larger support size and higher order of RB functions will yield better convergence results. This fact can be partially explained in the light of the stationarity of the total energy of the continuum, as explained in the following paragraph.

The total energy of the structure, given by the sum of the elastic deformation energy and the potential energy of the applied loads, is defined as follows:

$$E(\boldsymbol{s}) = \frac{1}{2}\int_B \boldsymbol{\sigma} \cdot \boldsymbol{\epsilon}\, dV - \int_B \boldsymbol{b} \cdot \boldsymbol{s}\, dV - \int_{\partial B_N} \hat{f} \cdot \boldsymbol{s}\, dS \tag{10}$$

Finding the vector field which minimizes (10) and satisfies the EBC (7) and the strain-displacement conditions (8) yields the solution of the elastostatic problem.

Functional (10) can usefully be slightly transformed, making use of the symmetries of the elastic tensor and of the deformation and tension components, yielding:

$$E(\boldsymbol{s}) = \frac{1}{2}\int_B \nabla \boldsymbol{s} \cdot C[\nabla \boldsymbol{s}] - \boldsymbol{b} \cdot \boldsymbol{s}\, dV - \int_{\partial B_N} \hat{\boldsymbol{f}} \cdot \boldsymbol{s}\, dS \tag{11}$$

If the displacement is expressed by means of a neural network in the form:

$$\boldsymbol{s}(\boldsymbol{x}) = \sum_{h=1}^{N_\phi} \boldsymbol{w}^{(h)} \phi_h(\boldsymbol{x}) \tag{12}$$

then its gradient is given by the expression:

$$\nabla \boldsymbol{s}(\boldsymbol{x}) = \sum_{h=1}^{N_\phi} \boldsymbol{w}^{(h)} [\nabla \phi_h(\boldsymbol{x})]^t \tag{13}$$

and, by substitution in relation (11), the stiffness matrix $\boldsymbol{K}$ and the load vector $\boldsymbol{q}$ can be computed such that the minimization of the total energy of the system yields:

$$\boldsymbol{K}\boldsymbol{d} = \boldsymbol{q} \tag{14}$$

where $\boldsymbol{d}$ is a vector which groups all the components of the vectors $\boldsymbol{w}^{(h)}$, which are therefore called *linear parameters*. Equation (14) represents the minimization of the structural energy with respect to the linear parameters, but not to the other characterizing parameters of the network, which will be discussed later.

As a matter of fact, the stiffness matrix $\boldsymbol{K}$ and the load vector $\boldsymbol{q}$ are given by the integration, over the continuum body and surface, of the gradient of the displacement field, and therefore depend on the location and shape parameters of the activation functions $\phi_h(\boldsymbol{x})$, which will be grouped in the vector $\boldsymbol{\theta}$ of *non-linear parameters*. In this way, it is possible to write that:

$$\boldsymbol{K} = \boldsymbol{K}(\boldsymbol{\theta}) \quad \boldsymbol{q} = \boldsymbol{q}(\boldsymbol{\theta}) \quad \Rightarrow \quad \boldsymbol{d} = \boldsymbol{K}^{-1}(\boldsymbol{\theta}) q(\boldsymbol{\theta}) = \boldsymbol{d}(\boldsymbol{\theta}) \tag{15}$$

and therefore the total energy of the structure will be approximated by the expression

$$E(\boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{d}^t(\boldsymbol{\theta})\boldsymbol{K}(\boldsymbol{\theta})\boldsymbol{d}(\boldsymbol{\theta}) - \boldsymbol{d}^t(\boldsymbol{\theta})\boldsymbol{q}(\boldsymbol{\theta}) = -\frac{1}{2}\boldsymbol{q}^t(\boldsymbol{\theta})\boldsymbol{K}^{-t}(\boldsymbol{\theta})\boldsymbol{q}(\boldsymbol{\theta}) \tag{16}$$

In other words, the minimum value of the total energy of the structure (with respect to the linear parameters) is a function of the non-linear parameters which define the activation functions, and it is therefore possible to optimize again the energy functional with respect to the non-linear parameters only, as shown in [9].

2.4 Meshless approach for dynamic problem

In the case of an elastodynamic problem, the displacements, deformations, tensions and loads depend, of course, also on time. Following standard approaches for the numerical solution of the dynamical problem, a consistent mass matrix and a damping matrix can be computed such that:

$$\boldsymbol{M}(\boldsymbol{\theta})\ddot{\boldsymbol{d}} + \boldsymbol{C}(\boldsymbol{\theta})\dot{\boldsymbol{d}} + \boldsymbol{K}(\boldsymbol{\theta})\boldsymbol{d} = \boldsymbol{q}(\boldsymbol{\theta}, t) \tag{17}$$

where the dependence of $\boldsymbol{d}$ on time and $\boldsymbol{\theta}$ have been omitted for sake of simplicity. In the following, non-linear parameters will be optimized for the elastostatic problem, and the non-linear parameters determined in this manner will be utilized to compute mass and damping matrices.

3 APPLICATIONS

In order to show the feasibility of the examined approach, two applications will be analyzed and discussed in the following, namely a 1D horizontal beam with a clamped and a hinged end, and a 2D horizontal cantilever, the solution of which can be found f.i. in [10].

3.1 1D application with a clamped-hinged horizontal beam

The considered structure is represented in Figure 2 and is simply a beam with a clamped end on the left and a hinged end on the right.

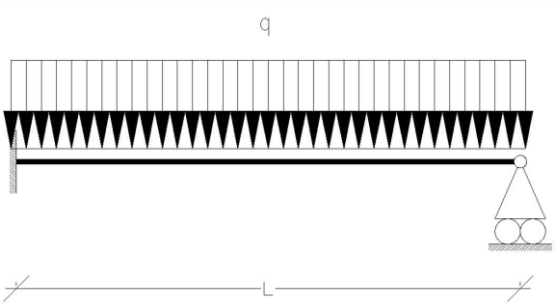

Figure 2: The examined structure.

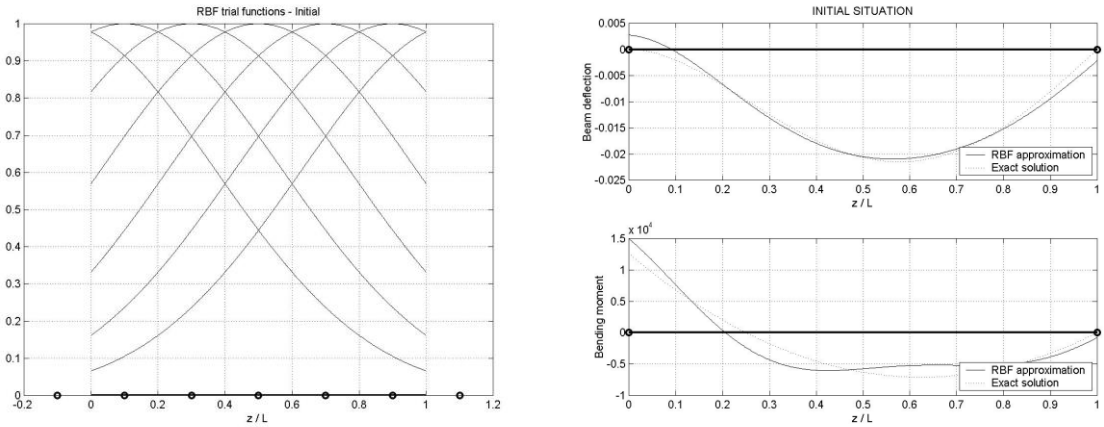

Figure 3: The initial arrangement of the RBFs (left) and the corresponding deflection (top right) and bending moment (bottom right).

At first, a set of 7 RBF were chosen, with centers (shown by the black circles in Figure 3) located at $z_h = [-1.0, 0.10, 0.30, 0.50, 0.70, 0.90, 1.10]L$ where L is the length of the beam.

If the non-linear parameters of the network are left as they are, which means no optimization is carried out on them, but just the linear parameters are evaluated, the obtained

solution is poor from the point of view of both deflections and bending moment, as it is clear from a look at Figure 3 (right).

On the other hand, if the optimization of the non-linear parameters is carried out, the results shown in Figure 4 are obtained, thus demonstrating the importance of the optimization of the non-linear parameters.

As a matter of fact, the optimization of the non-linear parameters is not strictly necessary, but the number of functions N_ϕ employed in equation (12) must be much higher.

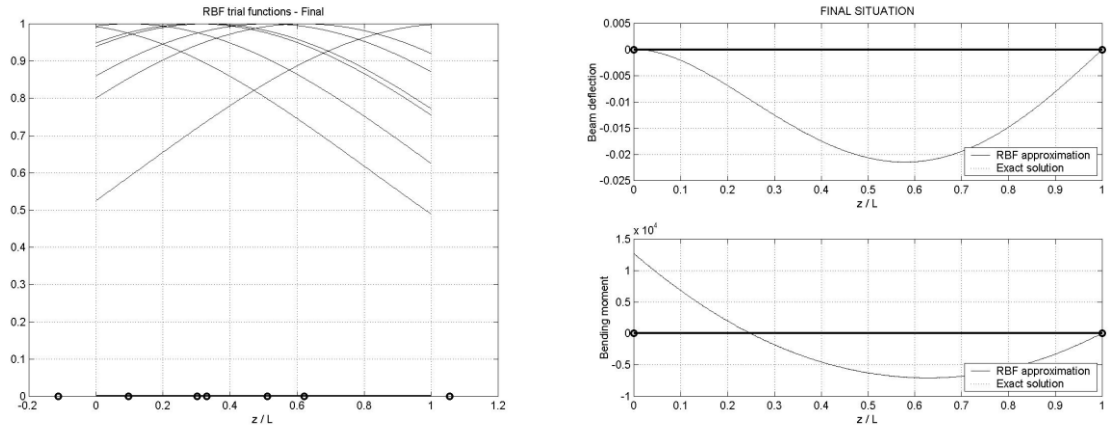

Figure 4: The final arrangement of the RBFs (left) and the obtained solution (top right: deflection, bottom right: bending moment).

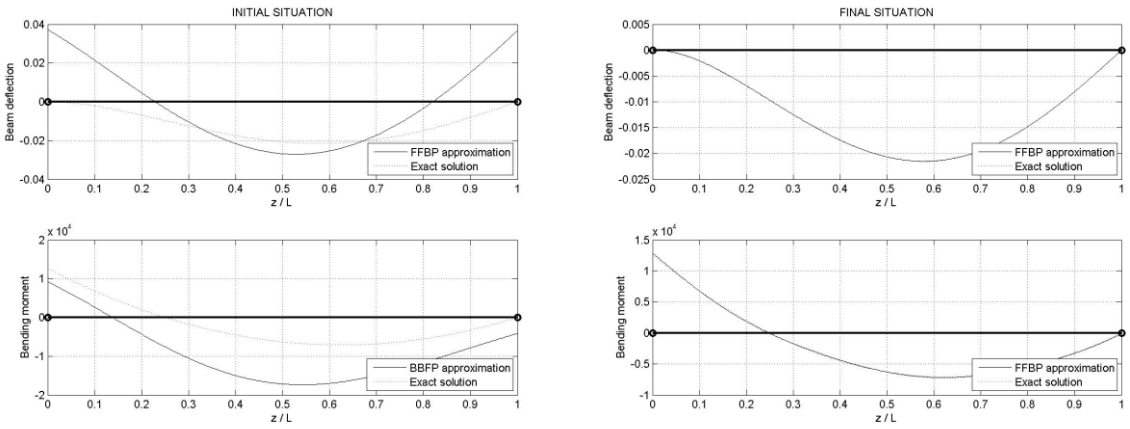

Figure 5: The solution (top: deflection, bottom: bending moment) with hyperbolic tangent activation functions; left: no optimization of non-linear parameters, right: optimized non-linear parameters.

A similar situation is observed with the use of a FFBP network in the form (5) with hyperbolic tangent activation functions, as shown in Figure 5.

3.2 2D application (Timoshenko cantilever)

In this example, a horizontal cantilever is considered, but the displacement is modelled as a bi-dimensional vector field which depends on two independent variables x and y, as shown in Figure 6. The dimensions of the cantilever are 5.00 m (length) by 1.00 m (height).

In the current example, the displacement is modeled by means of equation (12); in order to obtain functions which vanish at $x = 0$, the RBFs were slightly modified, and the following functions were used:

$$\begin{cases} \tilde{\phi}_h(\boldsymbol{p}) = \phi_h(x,y) - \phi_h(0,y) \\ \phi_h(\boldsymbol{p}) = \exp\left(-\| \boldsymbol{p} - \boldsymbol{c}^{(h)} \|^2/\sigma^2\right) \\ \boldsymbol{p} \equiv [x,y]^t \quad ; \quad \boldsymbol{c}^{(h)} \equiv [x_{ch}, y_{ch}]^t \end{cases} \tag{18}$$

Such modification is significant only when the RBF is centered near the clamped end, as is clear from a look at Figure 7.

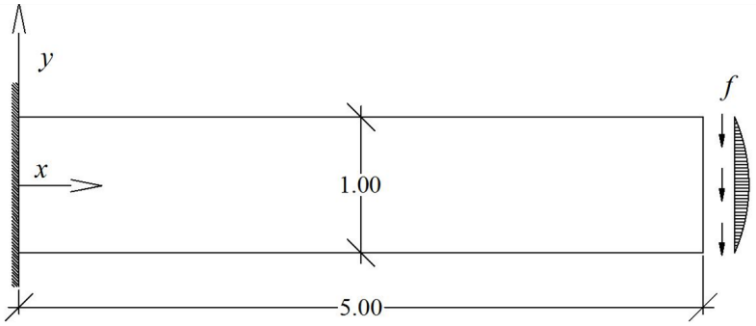

Figure 6: The considered cantilever: the left end is clamped, the right end is free and loaded by a parabolic vertical load *f*(*y*).

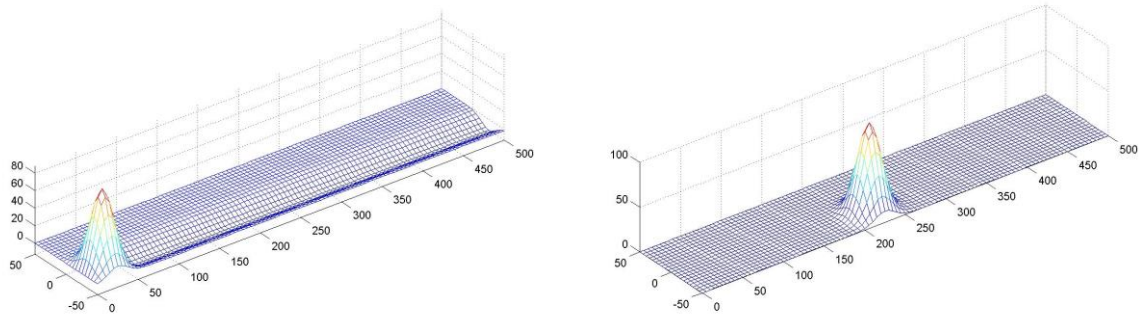

Figure 7: Two examples of the employed RBFs: on the left a function centered near the clamped edge, while on the right a function centered far from the clamped edge.

Element number and type	Degrees of freedom	Tip deflection
20 × CST	40	25,00%
10 × Q4	40	67,00%
10 × LST	48	99,00%
2 × Q8	20	93,00%
2 × Q9	24	95,00%

Table 1: Results reported in [11] for a cantilever beam. The tip deflections are the percentages of the solution obtained for a 1D horizontal slender beam.

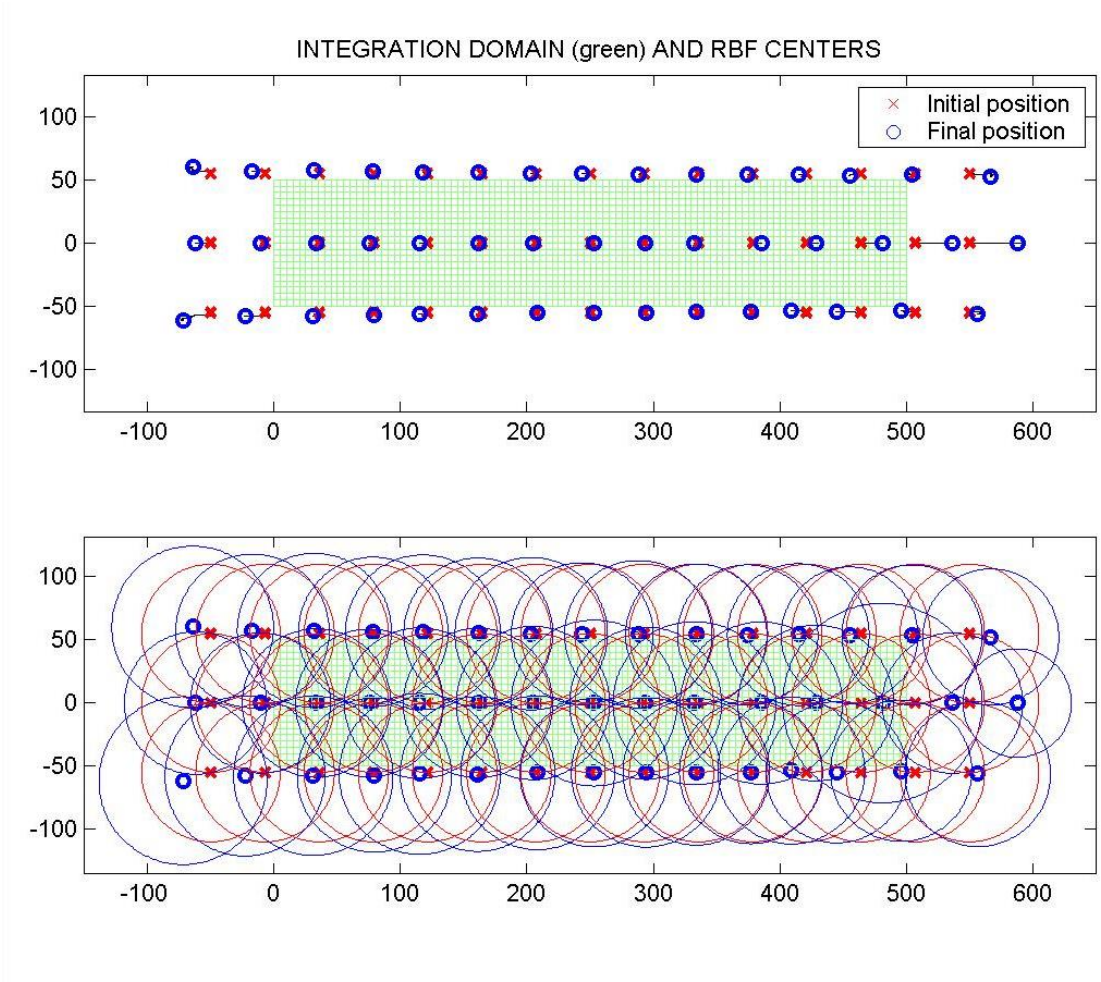

Figure 8: The training of the RBF network in the case of the 2D cantilever: above, the positions of the centers (red crosses denote initial position and blue circles denote final positions), while below the radii of the circles indicate the scaling parameters (red: initial values; blue: final values). The green area indicates the position of the cantilever.

In this case, the optimization of the RBF centers $\boldsymbol{c}^{(h)}$ and scaling parameters σ_h shows that convergence is observed towards a pattern where the centers are placed at more or less equal distances from each other, and the scaling parameters allow a wide overlapping between the functions.

Following such indications, several arrangements of RB functions were tested: a noticeable result is that, after a proper training of the non-linear parameters, a limited number of functions is sufficient to obtain a satisfactory approximation of the solution (see Figure 9 and Figure 10 for details).

For sake of comparison, Table 1 reports results from [11] for a cantilever beam with aspect ratio $L/H = 10$, along with the element types employed in the analysis and the number of degrees of freedom. CST denotes triangular, 3-nodes plane stress elements with linear shape functions. LST denotes triangular, 6-nodes plane stress elements with parabolic shape functions. Q4 denotes quadrilateral, 4-nodes plane stress elements with bi-linear shape functions. Q8 and Q9 denote quadrilateral, 8- and 9-nodes plane stress elements with cubic and quartic shape functions, respectively.

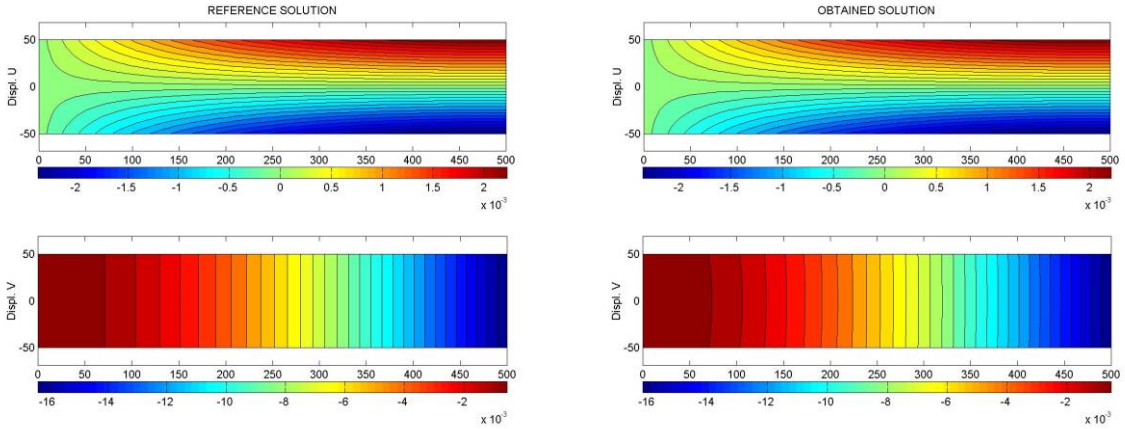

Figure 9: The solution for the horizontal (above) and vertical (below) displacement given in [10] (left) and obtained (right) with a network with only 13 RBFs.

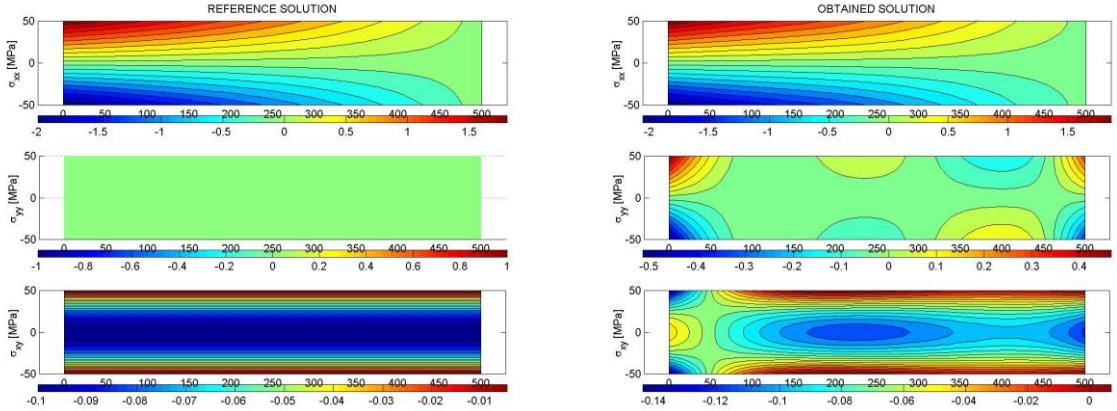

Figure 10: The solution for σ_{xx} (above), σ_{yy} (middle) and τ_{xy} (below) given in [10] (left) and obtained (right) with a network with only 13 RBFs (which corresponds to 26 linear and 39 non-linear parameters).

3.3 2D dynamic application (modal shapes for the 2D cantilever)

The non-linear parameters, optimized for a static load such that of the previous example, can be usefully employed to determine the consistent mass matrix of the structure; in the case of a time-dependent displacement field, equation (12) is changed in

$$\boldsymbol{s}(\boldsymbol{x},t) = \sum_{h=1}^{N_\phi} \boldsymbol{w}^{(h)}(t)\phi_h(\boldsymbol{x}) \tag{19}$$

and, upon differentiation, the velocity field is given by:

$$\boldsymbol{v}(\boldsymbol{x},t) = \dot{\boldsymbol{s}}(\boldsymbol{x},t) = \sum_{h=1}^{N_\phi} \dot{\boldsymbol{w}}^{(h)}(t)\phi_h(\boldsymbol{x}) \tag{20}$$

where the dot indicates differentiation with respect to time.

The kinetic energy of the continuum is therefore given by:

$$E_K(t) = \frac{1}{2}\int_B \boldsymbol{\rho}(\boldsymbol{x}) \parallel \boldsymbol{v}(\boldsymbol{x},t) \parallel^2 dV \tag{21}$$

thus leading to the usual evaluation of the consistent mass matrix $\boldsymbol{M}$; it is to be noted that such matrix can be regarded as a function of the non-linear parameters again: $\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{\theta})$.

It is to be noted that a careful optimization of the non-linear parameters carried out in the case of a static load works well also for the determination of the dynamic characteristics of the structure, as can be seen by Figure 11.

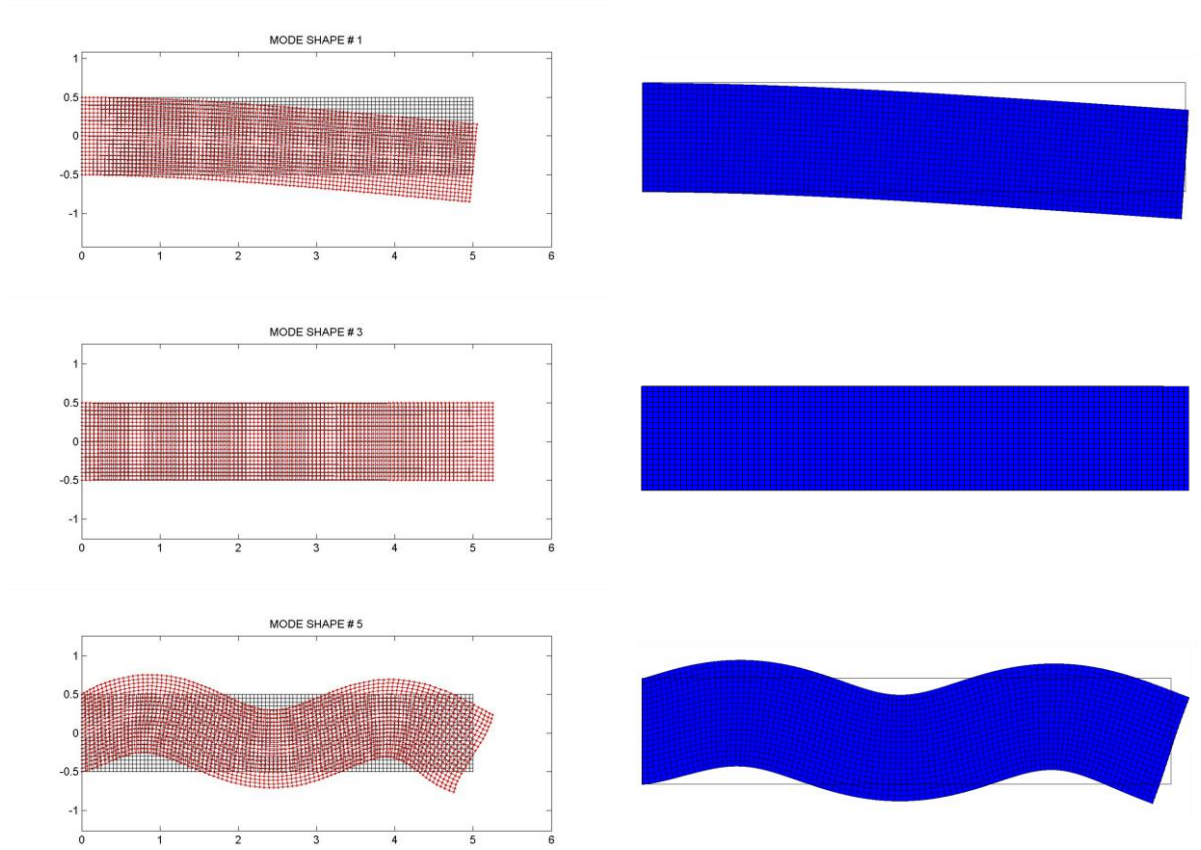

Figure 11: Some examples of the obtained modal shapes of the cantilever: first transverse (top), first axial (middle), fourth transverse (bottom). On the left side the shapes obtained with the meshless approach are shown, while on the right side the results from ANSYS FEM code are shown for comparison.

CONCLUDING REMARKS

The paper presents some results obtained by a RBF meshless approximation of the elastostatic and elastodynamic problems. Following the philosophy of neural networks, a set of non-linear parameters and a set of linear parameters have been defined.

It is clear that the optimization of the non-linear parameters of the network can bring notable benefits on the quality of the obtained approximation, and enable to greatly reduce the numbers of activation functions employed in the network.

The optimization performed for a static load can be extremely effective also in the computation of the dynamic characteristics of the examined structure.

REFERENCES

[1] S.D. Daxini, J.M. Prajapati, A review on recent contribution of meshfree methods to structure and fracture mechanics applications. *The Scientific World Journal*, Article ID 247172, 13 pages, 2014, http://dx.doi.org/10.1155/2014/247172.

[2] T. Belytschko, Y. Krongauz, D. Organ, M. Fleming, P. Krysl, Meshless methods: an overview and recent developments. *Computer Methods in Applied Mechanics and Engineering*, **139**, 3-47, 1996.

[3] G.R. Liu, Y.T. Gu, *An introduction to meshfree methods and their programming*. Springer, Dordrecht 2010.

[4] G. R. Liu, *Meshfree methods: moving beyond the Finite Element Method*. CRC Press, Boca Raton 2010.

[5] B. Gotovac, V. Kozulic, On a selection of basis functions in numerical analyses of engineering problems. *International Journal for Engineering Modelling*, **12(**1-4), 24-41, 1999.

[6] D.S. Broomhead, D. Lowe, Multi-variable functional interpolation and adaptive networks, *Complex Systems*, **2**, 269-303, 1988.

[7] T. Poggio, F. Girosi, Networks for approximation and learning, *Proceedings of the IEEE*, **78**(9), 1481-1497, 1990.

[8] L.X. Peng, S. Kitipornchai, K.M. Liew, Analysis of rectangular stiffened plates under uniform lateral load based on FSDT and element-free Galerkin method, *International Journal of Mechanical Sciences*, **47**(2), 251-276, 2005.

[9] L. Facchini, P. Biagini, Neural networks for meshless and PUFEM approaches, *22nd International Congress of Theoretical and Applied Mechanics (ICTAM 2008)*, Adelaide, Australia, August 24-29, 2008.

[10] S.P. Timoshenko, J.N. Goodier, *Theory of elasticity, 3rd edition*, Mc Graw-Hill International editions, Singapore 1970.

[11] R.D. Cook, D.S. Malkus, M.E. Plesha, R.J. Witt, *Concepts and applications of finite element analysis, 4th edition*, John Wiley and Sons inc., 2002.