Some aspects of application of artificial neural network for numerical modeling in civil engineering

M. LEFIK*
Department of Geotechnics and Engineering Structures, Technical University of Łódź, 6 Politechniki Ave., 90-924 Łódź, Poland

Abstract. In order to obtain reliable results of computations in civil engineering, the numerical procedures that are used at the stage of design should be calibrated by comparison of the theoretical results with an observed behavior of previously modeled and then executed structures. The hybrid Finite Element code with an Artificial Neural Network inserted as a representation of a constitutive law offers a possibility to adjust not only parameters of the constitutive relationships but also its qualitative form. Because of this, the representation of constitutive law by the ANN is presented in this paper. The constitutive data should be calibrated to fit well the observable values, measured in experiments. If the constitutive law is expressed by ANN, the inverse problem can be reduce to a training of the ANN inserted into the Finite Element code. An example of a solution of the inverse problem in calibration of constitutive law is presented. An identification of parameters of flow of pollutant in soils is described as another example of application of ANN in engineering.

Key words: hybrid code FEM-ANN, inverse solution with ANN.

1. Introduction

We present some applications of artificial neural networks in practice of civil engineering in this paper. We focus our attention on phenomenological representation of constitutive relationships. Since inverse analysis is of special importance in this issue, some solutions of an inverse problem with artificial neural networks is also addressed here.

Artificial neural networks (ANN) can be considered as non-linear operators that transform a set of suitably interpreted variables at its input into another set of numerical data at its output. The values of input and output should be known from observations. The ANN is “trained” with the known data to perform such the transformation and to become the “numerical, phenomenological model” of the considered phenomenon. The internal structure of the neural operator ANN remains the same for very different problems. For applications - it is enough to suitably interpret the input and the output of the ANN. This facilitates miscellaneous, qualitatively very different use of the ANN in various domains of engineering, including civil engineering.

ANNs are very efficient for analysis of large data systems, like for example elaboration of image in informatics or large databases in civil engineering. This technique is suitable for problems described with many, mutually dependent variables. The problem is always defined by examples of input – output pairs. The required definition of the problem can be not-exact (soft), superposed with noises, it can contain redundant or repeated information. The ANN works very well with such the difficult data sets.

If the ANN-based phenomenological model of the problem is constructed correctly, we can be sure that the essential rules or laws governing the problem are inscribed into the structure of weights of connections of the nodes (neurons) of the ANN. In this sense the ANN “discovers” the law governing the problem, acting still as a black box that behaves like the physical system that it models. It is believed that the approximation by ANN handles the real, deep dependence between two sets of data much better than any theoretical approach. It simulates very well a complicated, complex behavior in which influences of various physical factors and features are superposed.

These properties (and many others) are common to a set of techniques called “soft computing”. In this family of computational strategies, the ANNs appear among Genetic Algorithms, Fuzzy Logic, Believe Networks, Cellular Automata and many others.

From the point of view of the quoted above properties, it seems to be natural that the domains of civil engineering such like mechanics of soils and geotechnics are mostly explored using the ANN technique. This is primarily because the soil is much more difficult to be theoretically described than other engineering materials. Also experimental or in situ data are usually neither sufficiently reliable nor complete in soil engineering. The second domain, the ANN is widely applied in modeling of composite materials used in constructions like sandwich plates, rope wires, porous materials.

It is impossible to review all applications of the ANN in civil engineering. It was a true explosion of papers and algorithms since eighties of the twentieth century. We mention here only the most important and the most promising for numerical modeling in civil engineering developments of the ANNs and we illustrate the paper with some results of our own experiences in this subject. We are going to describe also some applications of very special ANN that performs its computing not on numbers but on words, acting thus as an

*e-mail: marek.lefik@p.lodz.pl
expert program that operates over a data base containing a description of a given problem. Our interest is also focused on hybrid computational systems, in which the soft computing performed by ANN complements only the classical hard computations (carried out by Finite Element code for example).

1.1. Structure of artificial neural network. ANN is composed of a collection of simple but nonlinear processing units (called nodes or artificial neurons) that are mutually interconnected. This is illustrated in Fig. 1. The connections modify the signals exchanged between neurons multiplying it by variable weights. This system of units is organized to transform an input signal into an output signal. Both input and output signals are suitably defined according to their physical interpretation. In this way we construct a functional dependence between the data at the input and at the output: the values at the output nodes can be seen as functions of the independent variables at the input nodes. This is obtained by a correct choice of coefficients (synaptic weights) that scale the signal transmitted between each pair of nodes belonging to the ANN. The weights of interconnections are modified by an iterative procedure to force the desired output signal – the response on a given input pattern. This process is called learning or training and it is continued until the error between the response of the neural network and the desired output (called target) is minimized for a whole set of pairs: given input – known output. Once the ANN is trained, it is used in recall mode to obtain the output of the problem at hand.

![Fig. 1. General scheme of the neural-like operator (ANN). Neurons are split into its own input and output, possibly different activation functions \( f, g \) that act on the neural input are shown. Arrows are used to mark the signal transmission. Weights (real numbers) are attributed to each arrow. The value of each output from neuron, multiplied by weight, is put on the input of another one as it is indicated by the arrows.](image)

Many algorithms that allow finding the weights \( w \) and biases \( b \) transforming the given input into the given output are created. For a fundamental introduction to the ANN technique the reader is referred to general textbooks and codes, for example [1–3]. Here we describe this method very briefly. We focus our attention on the widely used ANN with nodes organized in layers. Each node of one layer is connected (and exchanges its weighted activity) with each node of the next and previous layers. The strength of these connections (weights) can be modified. Within one layer the direct communication between neurons is forbidden. Input and output layers are accessible to the user. The values obtained at the nodes of the output layer (result of the network activity) as well as those given by the user to the input nodes, are interpreted in terms of the modeled problem.

The nature of nonlinear processing units – *neurones*, the architecture of their mutual interconnections and the algorithms of learning imply some very interesting properties of ANN.

The most important one is an ability to simulate numerically nonlinear behaviour, even very complex, basing on phenomenological knowledge about it. This knowledge is defined by a sufficiently large set of examples. Each example is composed of values of independent variables describing the problem and another set of values defining the output, considered as dependent variables. Formally, the ANN can be understood as a universal approximate of any functional dependence between the two set of data, if such a relation exists. It is also proved that any operator that acts from a set of variables into another set of data can be approximated by ANN as well (for details, see for example [3]). Unfortunately, the theorems that ensure that the ANNs can approximate any function of many variables are non-constructive. It means that the correct number of the internal degrees of freedom of this approximate (i.e. the number of hidden layers, the total number of nodes, their repartition between the hidden layers that determines the number of weights and biases) has to be decided by the user. These characteristics are problem dependent (the acceptable error of the approximation is given by the tolerance of the problem) and a general prescription cannot be proposed here.

1.2. Some most important application of ANN in civil engineering. The pioneer application of ANN in mechanics is attributed to Ghaboussi and al. They have proposed to use the ANN in civil engineering for direct representation of constitutive behavior of concrete. In his early works, he trained the ANN with experimental data collected from biaxial stress test of concrete cube. At the output of the ANN, increments of stress tensor were computed while strains increments have been presented at the input of the ANN. In this way, the ANN memorized, approximated and “generalized” the experimential data. This phenomenological description was neither constrained nor deformed by any “a priori” speculative, theoretical reasoning.

Many revolutionary ideas have been proposed by those authors who introduced the technique of soft computing in the field of numerical modeling of mechanics of structures. We mention for example the so called self-learning finite element procedure introduced first by Ghaboussi in [4–6] and developed then by Shin and Pande in [7, 8] as a very promising technique of data acquisition for the ANN training process. Presentation of this method that was applied immediately to problems of diaphragma wall dimensioning would exceed the
frame of this paper. We refer the reader to the papers [9, 10], containing the preliminary experience of the author in the application of this procedure. Also in [11] this method is used.

The use of ANNs for the direct identification of macroscopic mechanical properties of a composite system with a very complex internal structure – namely force-displacement relationship for a bundle of superconducting strands – can be found in [12] for a one-dimensional case and in [13] for a two-dimensional one. In [14–17] the non-linear constitutive behavior of the system has been represented by ANN and directly incorporated in a Finite Element (FE) code, while in [18] ANNs are used for the identification of the parameters of a constitutive law. In [15, 19–27] interesting reviews of possible applications of ANNs in nonlinear mechanics can be found.

ANN can serve to create some formulae, basing on training with experimental data, very useful in design process [28–30]. The most important advantage of these formulae is that an updating is always possible, when new experimental data appear, by retraining of the ANN. Another important application of ANN is creation of data bases containing physical description of soils and their measured properties (National Geotechnical Databases). Similar data bases are created for geotechnical technologies. The role of ANN in construction of the data bases is very important. ANN can directly represent or post-process the data. An “intelligent shell” created over the data base using ANNs of special construction assures a search for various, hidden correlation among the data.

We quote some papers that account for this very important application of ANNs: [19, 31].

Inverse problem is especially designated to be solved by ANN, since network approximates very well a solution of minimization problem. A broad literature of this problem is known, we suggest only few papers, the further references can be found there: [24, 32–36]. In the solution of inverse problem, not only numerical robustness is important but also the fact that the suitably trained ANN can be used as an inverse relation between observable data and parameters of the model. The last section of the present paper is devoted to this application of ANN.

Among more recent papers some examples of very important application of ANN should be recalled (also these examples concern the field of civil engineering). Two companion papers [37, 38] present a numerical approach designed for material parameter identification for the coupled hydro-mechanical boundary value problem (BVP) of the piezocone test (Cone Penetration Test – CPT) in normally and slightly over-consolidated clayey soils. The papers explore a possibility of using neural networks to solve the complex inverse problem of the penetration test, including partially drained conditions. It has been demonstrated that the development of ANN based inverse models can be based on training data sets which consist of pseudo-experimental measurements derived from numerical simulations of the piezocone test. The authors proved the effectiveness of correlating constitutive parameters with synthetic piezocone measurements, by means of the ANN modeling. Important for geotechnics, general conclusions related to the ANN training with the aid of numerical database have been drawn. The straightforward continuation of the method developed in the mentioned above papers was identification of constitutive model parameters using self-boring pressuremeter tests [34]. This idea of identification of soil properties, basing on geotechnical laboratory tests (as well as on the above described “in situ” tests) is very important and should become the leading method in the field of numerical modeling, in author’s opinion. Also some example of practical use of ANN in civil engineering can be easily fond. For instance – data measured with well known Falling Weight Deflectometer, used to check a quality of layered structure of pavements and roads, are interpreted with ANN. This is described in [33] and in many text books concerning this method of in situ testing. Another example – correlations between the data measured on the lateral, cylindrical and conical surfaces of the tip of the probe in various CPT (pressure and friction) and the properties that allow to classify the soil and to predict elements of its classical mechanical description [34].

The most promising method of data elaboration in the case of CPT is the prediction of the dimension of the pail for given bearing capacity, inserted in the soil the reaction on the CPT is known. In this case ANN technique is mostly adequate.

Development of application of ANN in civil engineering in Poland was triggered by a keynote lecture given by Z. Waszczyszyn at the 41st Polish Engineering Conference in Krynica in 1995, which brought attention of many engineers and researchers to this field. Since this moment, many important contribution concerning theory of ANN and its application, mainly in structural mechanics, appeared. More than 15 years of the activity in the field of neural computing was summarized in 2011, again during the Polish Engineering Conference in Krynica. Review of main achievements of the scientific groups from Technical Universities of Kraków, Rzeszów, Białystok ([30, 35, 39–41]) and others, chaired by Z. Waszczyszyn can be found in [42]. In parallel, soft computing (first of all genetic algorithms, recently also immunological algorithms, fuzzy computing and other generalized methods, inspired by biological observations) in a context of optimization, topological optimization, faults detection and others disciplines closely related to field of mechanics, mechanical engineering and civil engineering was being developed at the Silesian University by a group of T. Burczyński. The conference AIMETH, organized in Gliwice is an important forum for annual presentation of the current state of art in soft computing in Poland.

2. Representation of a constitutive relationships using ANN and finite element method

The civil engineers agree that the numerical methods, we dispose in design process, especially the most common – Finite Element Method (FEM), are not always reliable. To have trustworthy results, the numerical procedures should be calibrated by comparison of the theoretical results with an observed behavior of previously modeled and then executed structures.
The calibration of FEM codes becomes a very important problem. EUROCODE 7, that defines the rules of geotechnical design, explicitly imposes a necessity of monitoring of the realized structures. This is in context of reliability of the design procedures, especially numerical computations.

In the FEM code the only adjustable parameters are constitutive data. Of course, in frame of the classical description of the constitutive relationships, calibration of the code is well possible, but the ANN offers the possibility to adjust not only the values of the parameters of the constitutive relationships but also its qualitative form. Because of this, the representation of the constitutive relationships by ANN is, in the author’s opinion, the central issue for the calibration of the numerical procedure in civil engineering. This is especially important in soil mechanics, where the constitutive models are still very difficult to define. Also parameters of these relations are very difficult to obtain. Experimental test, both “in situ” and in laboratory are hardly interpreted, the constitutive data result rather from correlations with physical description of the soil and its state than from direct measurements.

Representation of constitutive law by the ANN is the subject of this section of the paper.

A hybrid ANN-FEM code is also very suitable in solving an inverse problem. Any laboratory or “in situ” experiment can be treated as a source of the inverse problem and the constitutive data should be calibrated to fit well the observable value, measured in experiment. If the constitutive law is expressed by ANN, the inverse problem can be reduced to the training of the ANN inserted into the FE code. This calibration is thus inseparable from FE model.

2.1. ANN as a numerical representation of constitutive relationships. The interest of the application of ANN in the case when the model is built directly from some available experimental data is obvious. In such a case an unknown conventional analytical constitutive description can be directly replaced with a suitably trained ANN.

In the most of applications of the ANN in civil engineering and in particular in mechanics of composites proposed in former papers of the author of the present article, for example in [18, 19, 20], the following form of constitutive model is privileged:

\[
(\sigma^{t+1}, F^{t+1}, \rho^{t+1}) = \text{ANN}(\sigma^t, F^t, \rho^t, \Delta F^t),
\]

\[
\Delta F^t = \Delta F^t \Delta t, \tag{1}
\]

It means that the ANN in the non-symbolic model is trained in such a manner that it associates with an increment \( \Delta F \) of a deformation gradient \( F \) at the current moment \( t \) and in a given state of stress \( \sigma \) and at the given current value of \( F - \) a new state of stress (in an incremental manner) and a new value of gradient of deformation. Density \( \rho \) is treated as a state variable. This form is derived from a hypo-plastic approach, and is very convenient since the only approximation of a tensorial function is needed to describe a non-linear material behavior in an admissible region of the stress space (any geometrical form in the space of stress, like for example – yield criterion, is not necessary to be defined).

In Fig. 2 a typical scheme of the ANN for representation of a constitutive relationship is shown. It is seen that the output of the network is a function of stress state, strains and increment of strains along a chosen path in the strains space.

In Fig. 2 it is marked that the deformation gradient can be replaced by an infinitesimal strains tensor.

Fig. 2. A typical ANN interpreted as a numerical approximation of a constitutive function. It is assumed that the numbers of neurons in the input layer and in the output layer are compatible with number of components of the respective tensors for the chosen dimensionality of the model. The \( \rho \) is a set of parameters. It can be, for example, density or porosity

2.2. Implementation of constitutive law represented by ANN inside the FE code. Technique of implementation of constitutive law represented by ANN depends strongly on the role the ANN plays inside the hybrid FEM-ANN procedure: training of the ANN inside the FE code or the action of the ANN inside the FE in the recall mode. Training of the ANN inside the FE code is governed usually by algorithms that are proposed by the above quoted authors: [6, 16]. These algorithms, fundamental for applications in civil engineering, are too complex to be analyzed here. Instead, the reader can be referred to papers recently published by E. Pabišek [17] for the most competent in the literature analysis of the so called “in line” and “off line” ANN training process. Some methods of practical integration of the two components are illustrated there with examples from the field of civil engineering: identification of constitutive behavior of the elements of plane truss and identification of an equivalent model for cohesion-less soils. The various techniques of fusion of the two component of the hybrid, FEM-ANN code are described; estimations concerning the numerical effectiveness are given. The situation, when the ANN is used inside FEM in the recall mode is much simpler for numerical realization. Various, important for civil engineering application of the ANN trained “off line”, as for example – definition of constitutive relationships for composites or functionally graded materials – are collected in [11, 43, 44]. The problem of iterative calibration of the constitutive relationships is not crucial for these applications. The only issue is how to incorporate the ANN into the formalism governed by a matrix notation of the FEM. This is described with some details in [10]. Here, the only draft of this solution is given.
Let us suppose that $\epsilon (du)$ is a strain measure – a function of the displacements increment computed during an iterative process of solving nonlinear BVP by FEM. This infinitesimal strain tensor is conjugate to material stress $\tau$ that is usually computed in points of integration at the level of a single element by the product (2). In formula (2) the constitutive tensor $D$ is assumed for the material attributed to the element while in the hybrid code it should be computed at the output of a suitably trained ANN. The first term under the integral (3) can be computed using a usual constitutive assumption:

$$d\tau = D : d\epsilon. \quad (2)$$

Using the assumed representation of constitutive law by ANN we have instead of (2):

$$d\tau = N_{d,\sigma} \odot d\epsilon. \quad (3)$$

Index $d$ denotes that the network quality is the best for some given value of increment $d$, $\sigma$ means that the value of stress increment is computed at the current value of $\tau = \sigma$. It is clear that following the scheme of the constitutive neural network, both increment of the stress tensor and its initial value at the current step of iterations are available. It is possible to include these values of stress $\tau$ with some manipulation at the level of the FEM code but the easiest way is to replace the neural operator in (3) by the matrix $D$ similar to the one in (2) but constructed directly using the given, neural representation of the constitutive law. This is done by trial incrementing of $\epsilon$.

Let us suppose that both tensors $d\tau$ and $d\epsilon$ are represented by column vectors:

$$[d\sigma] = [d\tau_1 \quad d\tau_2 \quad d\tau_3]$$

$$[d\epsilon^T] = [d\epsilon_1 \quad d\epsilon_2 \quad d\epsilon_3] \quad D = [d\sigma] [d\epsilon^T]^{-1}. \quad (5)$$

In the algorithm proposed in [10] the matrix of trial vectors $d\epsilon^T$ is always proportional to the strains at the last equilibrated point during the FE iterative solution (at the preceding step). Trial vectors cannot be arbitrary because $N_{d,\sigma} \odot d\epsilon \neq -N_{d,\sigma} \odot (-d\epsilon)$ (ANN is nonlinear). In fact, two different tangent stiffness matrices can be defined in any point: one for loading and the other for unloading it is supposed thus that the loading (unloading) is continued during the current increment in the Newton iterations. Computation of increments of the stress tensor can be also achieved by of computing the derivatives of the neural network with respect to input values (The method proposed in [12, 45]).

The stress (3) is computed using neural network in the recall mode for given, constant step $d\epsilon$, until the strain $\epsilon$ at the trial solution at the current step is reached. The ANN acts here in the autonomous activity mode as it was defined before. This process corresponds to the classical integration of incremental constitutive equation for updating $\sigma$. It starts always at the last equilibrated point and the increment $d\epsilon$ is proportional to the one, defined for this step (loading or unloading).

### 2.3. Example of application of ANN – virtual testing method in mechanics of composites

As it is stated in the previous sections, the available experimental or numerical data are sufficient sources of information to describe the constitutive dependence between the scalar, vector or tensor fields of interest. In the publications [14, 17, 43, 44] the only numerical tests made on a sufficiently large portion of the composite were used to identify the mechanical behavior of the homogenized material. For a non-linear composite or for a complex hierarchical heterogeneity, very frequent in materials used in civil engineering, an adequate description of the effective behavior is usually difficult to obtain on a purely theoretical way. The classical, symbolic constitutive law is usually identified from known properties of a representative volume element (RVE), based on a suitable version of some homogenization theory. An alternative to the theoretical development is given by the numerical testing of the composite. This approach is well known: numerical experiments can be carried out on a RVE of the composite using e.g. a FE code. Usually deformations are kinematically imposed and the material properties are obtained from the relation between average strain - average stress measures, computed from the FE solution. This method is also known as virtual testing and can be used to compose the input-output data set to train the network. ANN is used both as a numerical representation of the effective constitutive law and as a numerical tool for the analysis of the constitutive relations between averaged quantities. ANN approximation replaces the usual symbolic description of the effective constitutive law.

By performing a sufficient number of numerical tests the network can be trained. ANN representation of the constitutive law cannot be used out of a numerical environment since its parameter (weights and biases) have no physical meaning. After the training process, inside the FE code ANN is used in recall mode to identify the stress value or the stress increment for a given strain, strain increment and possibly other variables (e.g. temperature, porosity etc. depending upon the assumed model). The results of the virtual testing, can always be rearranged to form the system of equations (5), where $D$ contains the 21 independent elastic constants (the elements of the effective stiffness matrix).

The following overall algorithm is proposed to define effective properties of any hierarchical composite material and to solve the resulting, homogenized boundary value problem:

1. **Preparation of the learning data**: for casual values of the materials data and for each geometry of the cell of periodicity the effective material characteristics are computed by a FEM solution of BVP with periodic conditions and are suitably post-processed.
2. **Training of the network with the pairs of sets**: given random input and computed (as is said above), corresponding output. Interpretations of input and output data are defined in Subsection 1.1.

Having the well trained ANN, starting at the micro-level, for each structural level and for each kind of cell of periodicity at the current level:
iii Run the neural network in the recall mode with input data that characterize the current level of the structure.

iv Complete the sets of input data for each cell of higher structural level from ANN outputs obtained at the previous level.

v Run the same neural network in the recall mode with suitably completed input data plus information characterizing the geometry of the cell of periodicity of the higher structural level.

vi At the macro level algorithm stops.

2.4. Example of determination of mechanical properties of soils using ANN. While in [17] the constitutive law has been obtained by virtual pseudo-experiment, in this section we describe the idea of introduction into FEM of the “true”, phenomenological constitutive law for soils via ANN. Determination of mechanical properties of soils is a well known procedure regulated by geotechnical praxis and codes. Three methods are possible here: method A, when the soil properties are defined basing on laboratory or “in situ” tests; method B basing on the experiment indirectly, through a knowledge about correlation between some well defined geotechnical properties and state of soils (called “leading properties”) and the mechanical properties needed for estimation of deformation and strength on the other hand; finally – method C according to which the accumulated experience and some similarity to the existing examples of “good engineering practice” is sufficient. The three possibilities classify the geotechnical experience of a very difficult and costly experimental process that is at the basis of geotechnical engineering.

Application of ANN opens in this domain a lot of new possibilities. Constitutive modeling of geomaterials is often basing on non uniform or non classical material tests, thus unusual tools of interpretation are necessary. Non uniform or non classical material tests – means the tests, in which the state of stress and/or strain is neither homogeneous nor uniform (for example: various cone penetration tests – CPT [37], pull out-test [46]). Many papers have been published (for example: [29]), in which ANN is used to discover and describe correlations among measured, observable quantities and the properties that are classically used to model the soil within the frame of classical method like FEM or others. Clearly speaking: elastic moduli and failure parameters are defined as it has been imposed directly by practice of laboratory investigations of steel, concrete or other traditional material. The mechanical models of soils require the same classical description by Young moduli and parameters describing failure of the material structure while the geotechnical tests quite never result directly with these descriptors! We have cited above in the Sec. 1 some papers that present correlations between measured parameters and the ones needed for analysis established with various use of different ANNs. Below we describe shortly our own preliminary results concerning this subject.

According to the presented method of incorporation of the ANN that carries constitutive information into the FEM code, we exemplify the calibration of such a finite element using oedometric and triaxial test. The algorithmic scheme of the calibration process including the ANN training is the following:

i Information on material properties is furnished in the form of graphs of the observable data, observed during a sequence of experiments. This is a proposed “standard” element of the input data for the FEM code with ANN inside (such a code does not exists yet at the market, the research version is still under construction).

ii At the first step, calibration procedure is running: FEM dissertation of the laboratory tests is defined, boundary condition is applied; the observed data in the form of usual graphs are supplied (in our case - the oedometric test and the triaxial test). This is illustrated in Fig. 3. These boundary value problems are then solved and weights of the ANN are modified iteratively to assure the best accordance of the computed graphs and the experimental ones, given at the input.

iii If the process of calibration is successful within a given tolerance, we are sure that the ANN at the elementary level is well calibrated. We note that the ANN represents constitutive information concerning the material under consideration in the context of the assumed discretization and the kind of the element chosen for the model! The same types of element should be thus used in the numerical FE-ANN modeling of the engineering problem.

In the example, we have used for calibration, the most promising (in the authors opinion) method called by the authors of [8, 9] – “a self learning Finite Element code”. Detailed description of this method can be found also in [15, 16]. It is too specific to be explained here. We mention only that, according to this method, at each successive step of the iteration the ANN is re-trained with constitutive data corrected by results of Finite Element computations (stress, strains and/or corresponding increments) obtained for the minus error committed in preceding step of approximation.

![Fig. 3. Scheme of the oedometric test and its simplified numerical model needed for process of calibration of the ANN](image)

It is seen in Fig. 3 that the numerical model for ANNs calibration is very simple (this is an easier, the drained test). It is assumed 25 hybrid FE-ANN axially-symmetric eight node elements with common ANN for each, inserted in the sense of the formulae quoted in Sec. 2. Oedometric data are taken from our own test of sandy clays from glacial moraine from Łódź. In Fig. 4, few of these data are drawn with pointed line. This kind of sandy clays is very typical for the geology of the region of Łódź, thus the calibrated ANN can be considered as a numerical representation of the data concerning a typical soil. In Fig. 4 we show a comparison of the measured data
Some aspects of application of artificial neural network for numerical modeling in civil engineering

Fig. 4. In Fig. a) the graph “executed” by very simple ANN of the geometry: 4751 (continuous line) trained with data from dynamic consolidation of composite structure. This graph is compared with the experimental observation (dotted line). The other graphs present a classical result of the oedometric drained test: thickness of the sample as a function of applied stress. Learning data are collected in Fig. b). In Figs. c) and d) – comparison of the autonomous ANN activity and the measured data is shown for two loading paths and the data resulting from “autonomous” network activity. It is to note, that the data inscribed in the hybrid FE-ANN element are qualitatively different: usually engineer asks for oedometric moduli, various for primary and secondary loading and unloading. The values of these moduli can be defined as finite quotients for chosen range of stress. In the proposed procedure, constitutive description is consisting of weights and biases of ANN that are impossible for interpretation, except – the interpretation shown in Fig. 4: by the graph of the results of the “virtual” experiment! It is seen that the qualitative agreement between measured and simulated displacements is quite good.

We stress that the presented above calibration should be performed “over” all the accessible experimental data. It can be treated as knowledge based soil description and corresponding software for geo-engineering. Creation of the new type intelligent data base and a suitable, postprocessor and preprocessor can be proposed for all “difficult” material as for example – concrete. The concrete soft modeling was at the beginning (from chronological point of view, [5]) of the history of the application of the ANN in civil engineering. All constitutive doubts concerning existing, adapted or revitalized structures in civil engineering should be arbitrated using the proposed methodology.

3. Example of solution of inverse problem using ANN in environmental geotechnics

An inverse problem is, in the author’s opinion, the second central problem of applications of ANN in civil engineering. Examples of some solutions of the inverse problem in calibration of constitutive law and in identification of parameters of the physical process are presented in the next sections of the paper.

The problem of identification of source position of the suspension that propagates in the domain of porous medium is very important. For example, in environmental geomechanics – identification of leakage from waste disposal basing on data from standard monitoring of the suspensions concentration in some piezometers – can be interpreted as the problem of identification of convective flow in a domain of porous media from observations of concentration of a solute in given points, using artificial neural networks and a numerical model. In this work, only virtual observations are used. Virtual observations consist of data read from numerical solutions of the problem of propagation of the solute in a porous medium. An artificial neural network is trained to solve the inverse problem with the data acquired from several direct solutions. The direct problem can be understood as a mapping between a given set of data (source position and its intensity, diffusivity vector and others, velocities etc.) and the set of elements of the corresponding solutions (e.g. the concentrations observed at the given points). On the contrary in this case, the ANN approximates the inverse mapping that attributes the source position, its intensity and the other input parameters – to the concentrations computed in some selected points.

It is well known that ANN has the properties of a universal approximating of a function or of an operator. If an unequivocal dependence of source positions (and other para-
meters of the convection problem) on the data observed at the test point exists, it is approximated with some tolerance by an ANN created for this problem. The uniqueness of the solution is one of the main issues of an inverse problem. In this context, the use of ANNs is advantageous since the response of the networks is always possible and usually it can be suitably interpreted [47, 48].

As an example we examine the problem of propagation of a pollutant in the soil. First we consider a 2D flow in a homogeneous medium in which the propagation is governed by advection and diffusion, a field of hydraulic head is given and the source of pollutant is prescribed. Direct solution for the steady state is obtained using Finite Differences method. Results of the direct solution – concentrations of pollutant in a few test points – are given at the input of the ANN; the coordinates of the position vector of the source (and its intensity if it changes) are given at the output of the ANN during training mode. In recall mode the network responds with the unknown leakage source position having at the input the given observations from the piezometers. The quality of the solution is discussed for identification of various flow parameters.

3.1. Convection in soils. Let us suppose that the convection problem is defined by the following equation, defined over a 2D unbounded domain \( \Omega \):

\[
\frac{\partial}{\partial x_i} \left( D_{ij} \frac{\partial c}{\partial x_i} \right) + v_i \frac{\partial c}{\partial x_i} = \frac{\partial c}{\partial t},
\]

where \( D \) is diffusivity, the concentration \( c \) and velocity \( v \) are defined as some averages of the corresponding 3D porous medium across a permeable stratum. Since we are interested in the inverse solution, we are not going to enter into details neither of physical meaning of the variables in (6) nor the mathematical solution in general, which is strongly marked by the proportion between the diffusivity vector and the given velocity field (Peclet number is small in the present application). Here we consider only a stationary case with constant, given velocity. It can be interpreted as the situation for \( t \geq t_s \) and \( v \) is given as, for example, for a river of a vein source of c starts and remains constant in time and the concentration is given as, for example, for a river of a vein source of c starts and remains constant in time and the constant velocity is given as, for example, for a river of a vein or an artesian aquifer). Furthermore, we suppose that the velocity field is derivable from a given hydraulic head \( h \). The set of stationary equations (Problem P1 – direct) writes as follows:

\[
\frac{\partial}{\partial x_i} \left( D_{ij} \frac{\partial c}{\partial x_i} \right) - k_{ij} \frac{\partial h}{\partial x_j} \frac{\partial c}{\partial x_i} = \frac{\partial c}{\partial t},
\]

with

\[
\mathbf{x}_s = \left( x_s, y_s \right) \in \Omega \quad c(\mathbf{x}_s) = c_s
\]

\[
\mathbf{x}_\infty \in \Omega : |\mathbf{x}_\infty - \mathbf{x}_s| \to \infty \Rightarrow c(\mathbf{x}_\infty) \to 0.
\]

In (7) \( k \) is a matrix of Darcy’s permeability coefficients. At infinity the concentration goes asymptotically to zero. For numerical computations we must cut the domain \( \Omega \) in some distance from the source of concentration. This distance should be large enough. Finite difference method allows us to obtain a family of solution of equations (7). In Fig. 5 a typical solution (represented by a map of concentration \( c \)) is shown. The presence of the velocity field superimposed on the source activity can make the identification even very difficult. The most difficult case seems to be the identification of the velocity field having only the data concerning the concentration field. As defined by equation (6), direct convection model attributes a field of concentration to a set of given data: boundary condition for \( c \), given concentration sources and velocity field treated as a given parameter. In this case, the inverse problem P2 can be defined as follows:

For a given \( n \) points of observation \( \mathbf{x}_{pi}, i = 1..n \), (called piezometers), for given \( c_i = c(\mathbf{x}_{pi}) \) and \( h_i = h(\mathbf{x}_{pi}) \),

Find \( \mathbf{x}_s = (x_s, y_s) \in \Omega \) and \( c(\mathbf{x}_s) \) such that \( c(\mathbf{x}) \) verifies equations (6) (direct problem P1).

In this preliminary research we limit ourselves to propose an algorithm that works well in a typical, engineering situation. We do not discuss thus the problem of non-uniqueness of the solution.
3.2. ANN for solving the inverse problem. Since the ANN has properties of a universal approximate, we can approximate with it the inverse relations defined in problems $P2$. We put, as the ANN target, the values of $x_s$ and $c_s$ corresponding to the hydraulic head and concentrations in observation points (known from the direct solution or from piezometers in a real situation) presented at the input of the ANN. Weights of the ANN are shaped to obtain a required output signal. We train thus the ANN to approximate the solution of the inverse problem $P2$ that we have defined above.

Figure 6 reports the interpretation of input and output nodal values suitable for the problem considered. The sequence of activation values for nodes in the input (or output) layer is called “the input-output pattern”. Hidden node activation has no physical meaning.

3.3. Examples of inverse solution: identification of source of the pollutant and its intensity. In practice of porous media, for example in environmental geotechnics, velocities are difficult to be measured in situ. It is natural, indeed, to express the velocity field in the domain of observation by gradient of hydraulic head. Hydraulic head is perfectly observable in piezometers. Numerical experiment of identification problem $P2$ in the case when data are collected in real conditions of various and unknown velocity field requires the ANN with number of nodes at the input layer twice than the number of piezometers: values of head and concentration in each of them. The minimal reasonably correct ANN in the case for which the velocity vector and source intensity are unknown is $10833$ (concentrations and head from 5 piezometers at 10 input nodes, 8 and 3 nodes in hidden layer and coordinates of the sources $x_c$, $y_c$ and intensity $c_s$ at three output nodes).

Quality measures are the following in this case (RMS error means: Root Mean Square error):

- Learning set RMS error: 0.00669
- Learning set correlation: 0.9996
- Test set RMS error: 0.00691
- Test set correlation: 0.9989

Quality measures are the following in this case (RMS error means: Root Mean Square error):

- Learning set RMS error: 0.00669
- Learning set correlation: 0.9996
- Test set RMS error: 0.00691
- Test set correlation: 0.9989

Fig. 7. Blue line indicates the correct values of the x coordinate of the source, red points – network output in learning mode. The output is well aligned with the target (ANN 10 83 3)

Fig. 8. Blue line indicates the correct values of the y coordinate of the source, red points – network output in learning mode. The output is well aligned with the target (ANN 10 83 3)

Fig. 9. Blue line indicates the correct value of the intensity of the source, red points – network output in learning mode. Small oscillation are visible. These errors are admissible from practical point of view. (ANN 10833)
For 5% perturbation the new network output is illustrated in Fig. 10. Blue points – network output are the ANN answers in recall mode (ANN 10 83 3).

We conclude that the presented identification procedure is not very complicated from numerical point of view and, at the same time – very reliable. We observe also that the relatively complex problems, like simultaneous identification of source position and the source intensity, or identification of the intensity of source is given. For unknown intensity the precision of the predicted source position is very low.

4. Conclusions

In situ material testing (properties of concrete or steel of existing structures that are to be reinforced, revitalized or adapted) and, first of all, properties of soils and interpretation of geotechnical tests seem to be the most important field of application of ANN in civil and geotechnical engineering.

With some details, the application of the ANN for constitutive materials description, namely for soils, was presented in this paper. The preferred form of this representation is the hybrid FE-ANN code. An incremental form of the constitutive relation is privileged from the point of view of the adopted strategy. Corresponding structure of the ANN for this application is shown. We can conclude that the proposed representation is not very complicated from numerical point of view and, at the same time – very reliable. We observe also that the proposed strategy is known in the literature, it becomes more and more popular and it probably will lead to a qualitatively new description of materials behavior in form of bases of knowledge.

Only five experiments used in the training of the ANN in the example have been sufficient to obtain good representation of the considered soil. In the opinion of the author of his paper, it should be created a common system of constitutive data description (similar to the one proposed in this paper) and corresponding data base accorded with the presented, hybrid ANN-FEM numerical treatment of BVP in civil and geotechnical engineering.

The presented approach is particularly efficient if nonlinear material behavior dominates the static work of the structure, as it is usually in soil mechanics. The loop of hysteresis for loading and unloading can be easily taken into account.

The most important advantage of the proposed approach is the following: the constitutive representation is constructed automatically, without any qualitative “a priori” consideration that can constrain the solution. These considerations need usually a great deal of scientific creativity which is impossible at the engineering level of application.

Important element of the proposed computational strategy is a solution of the inverse problem. The exemplary solution has been presented to introduce the ANN of the special structure that works well in this context. The ANN based solution of the inverse problem is insensitive on some defects of the problem formulation like for example redundant or repeated experimental data.

It is seen also that the relatively complex problems, like for example – simultaneous identification of source position...
vector and source intensity — is not very difficult and can be correctly done with ANN.

What is very important in the use of ANN to solve inverse problem in civil engineering — the ANN is always the numerical representation of the inverse relation and thus — it can be used in the same application as many times as needed, replacing typical “ad hoc” back calculation for any new set of experimental data.

Until recently, engineering knowledge has been deposited mainly in the heads of people. In some countries, applications of the ANN for creating various data bases of engineering knowledge enables the specialist to access know-how and engineering data, stored in a structured way. Some data bases contain now thousands of ‘experiences’ that are of particular importance. It enables engineers to use former experience in future realization in the field of civil engineering. More and more designers adopt the knowledge contained in this ‘collective brain’ each day via Internet.

Acknowledgements. The scientific research reported here has been carried out as a part of the Project “Innovative recourses and effective methods of safety improvement and durability of buildings and transport infrastructure in the sustainable development” at the Technical University of Łódź, Poland, financed by the European Union from the European Fund of Regional Development based on the Operational Program of the Innovative Economy.

REFERENCES

[27] Z. Waszczyżyn, “Neural networks in plasticity: some new results and prospects of applications”, Proc. of Eur. Cong. on...


