Artificial Neural Networks for Classifying Magnetic Measurements in Tokamak Reactors

A. Greco, N. Mammone, F.C. Morabito, and M.Versaci

Abstract—This paper is mainly concerned with the application of a novel technique of data interpretation to the characterization and classification of measurements of plasma columns in Tokamak reactors for nuclear fusion applications. The proposed method exploits several concepts derived from soft computing theory. In particular, Artifical Neural Networks have been exploited to classify magnetic variables useful to determine shape and position of the plasma with a reduced computational complexity. The proposed technique is used to analyze simulated databases of plasma equilibria based on ITER geometry configuration. As well as demonstrating the successful recovery of scalar equilibrium parameters, we show that the technique can yield practical advantages compares with earlier methods.

Keywords—Tokamak, Sensors, Artificial Neural Network.

I. INTRODUCTION

TOKAMAK [1] are experimental devices aiming to demonstrate the technical feasibility and practical relevance of controlled thermonuclear fusion via magnetic confinement. A critical issue both for design and operation of a Tokamak machine is the real control of the plasma ring in the chamber during the discharge [2]. For this reason, one needs a fast identification tool of the plasma position and shape starting from a set of measurements, usually given by magnetic probes and loops located in the proximity of the chamber wall. The task is difficult, especially if the plasma cross-section is non-circular, since there are more parameters to be estimated in order to completely characterize the equilibrium. The problem becomes even more difficult if the

Antonino Greco is with the Dipartimento di Informatica Elettronica Trasporti, DIMET, Facoltà di Ingegneria, Università "Mediterranea" degli Studi, Reggio Calabria, I-89100, Italy (corresponding author to provide phone: +390965875285; fax: +390965875220; e-mail: ngreco@ ing.unirc.it).
Nadia Mammone is with the Dipartimento di Informatica Elettronica Trasporti, DIMET, Facoltà di Ingegneria, Università "Mediterranea" degli Studi, Reggio Calabria, I-89100, Italy (corresponding author to provide phone: +390965875353; fax: +390965875220; e-mail: nadia.mammone@ ing.unirc.it).
Francesco Carlo Morabito is with the Dipartimento di Informatica Elettronica Trasporti, DIMET, Facoltà di Ingegneria, Università "Mediterranea" degli Studi, Reggio Calabria, I-89100, Italy (corresponding author to provide phone: +390965875397; fax: +390965875220; e-mail: morabito@ unirc.it).
Mario Versaci is with the Dipartimento di Informatica Elettronica Trasporti, DIMET, Facoltà di Ingegneria, Università "Mediterranea" degli Studi, Reggio Calabria, I-89100, Italy (corresponding author to provide phone: +390965875273; fax: +390965875220; e-mail: versaci@ ing.unirc.it).

kind of Magneto-Hydro-Dynamic (MHD) equilibrium of the plasma changes during the discharge. This is the case, for example, of plasma which passes from a Limiter configuration (where the plasma boundary is defined by the outermost magnetic flux line before touching any metallic wall) to an X-point configuration (where the plasma boundary is defined by the flux line where a null point and consequently a field bifurcation occur).

In the domain of plasma control, the location of plasma column (in terms of shape and position) evolving in the chamber represents a crucial step. In particular, the problem can be formulated as the search of a suitable mapping between the set of available measurements (sampled by means of sensor located around the chamber contour) and the selected set of shaping parameters.

In this paper, we have focused our attention on ITER configuration (Figure 1) in which we deal with inner, outer and divertor sets of sensors. In order to reduce the computational complexity and due to the fact that the inner sensors are inaccessible, we propose a Neural Network approach (NN) to classify measurements (inner, outer, divertor).

To improve the obtained performance, we propose an approach that provides equivalence between outer sensors and inner-divertor ones.

The paper is organized as follows. Section II reports an overview of the exploited numerical database. After a short description on Artificial Neural Networks (ANNs) reported in section III, we describe the proposed approach for classification problem (section IV). Section V shows the procedure to drive to the reduction of the computational complexity. Finally, we draw some conclusions.

II. THE ITER NUMERICAL DATABASE: AN OVERVIEW

Using the ITER coil and vessel geometry (Fig. 1), including the 6 dominant passive current eigenmodes, a database of 4848 lower single null equilibria has been generated by the Plasma Data Analysis Group (PDAG), Physics Department, University College Cork, Association EURATOM-DCU.

The equilibria were generated using a Database Generation and Analysis Package (DGAP) which has been developed by PDAG.

The core equilibrium calculation in DGAP is performed by the Garching Equilibrium Code (GEC). The magnetic parameters of database built with B-tangential and B-normal signals simulated of gaussian noise (average= zero, standard

deviation=magnitude of the simulated measurement noise) of 10 mTesla [3], are referred to lower X-point plasma which Plasma Current (IPLA) is 15 MAmpere, the toroidal field (Bo), referred to 6.2 meters from the center of torus, is 5.3 Tesla.

The magnetic measurements, deriving from the sensors located along the contour of the chamber, are subvided as follows:

- 24 B_ Tangential signals on the Vacuum Vessel Inner Skin Contour;
- 24 B_ Normal Signals on the vacuum vessel Inner Skin Contour;
- 6 B_ tangential signals below the Divertor Contour;
- 6 B_ normal Signals below the Divertor Contour;
- 120 B_ tangential Signals on the Vacuum Vessel Outer Skin Contour;
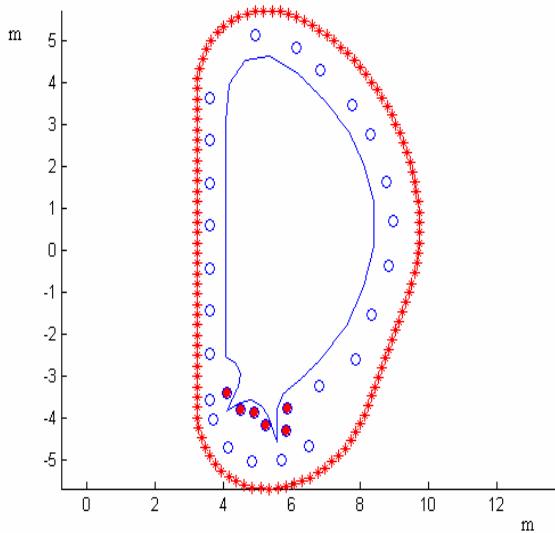- 120 B_ normal signals on the Vacuum Vessel Outer Skin Contour.

Fig. 1 The cross-section of the ITER Configuration with a schematic dislocation of the outer (star), inner (circles) and divertor (bold circles) sensors around the vacuum vessel

The analysis considers three different configurations of inputs. In particular:

1) first configuration: inner + divertor (60 parameters)
- 24 B_ Tangential signals on the Vacuum Vessel Inner Skin Contour;
- 24 B_ Normal Signals on the vacuum vessel Inner Skin Contour;
- 6 B_ tangential signals below the Divertor Contour;
- 6 B_ normal Signals below the Divertor Contour;

2) second configuration: Outer skin contour (240 parameters)
- 120 B_ tangential Signals on the Vacuum Vessel Outer Skin Contour;
- 120 B_ normal signals on the Vacuum Vessel Outer Skin Contour.

3) third configuration: the whole of magnetic signals (300 parameters).

Starting from third configuration, by means of the proposed approach, a classification of measurements is carried out. Finally, exploiting first and second configurations, a sort of equivalence between outer measurements and inner-divertor ones is showed.

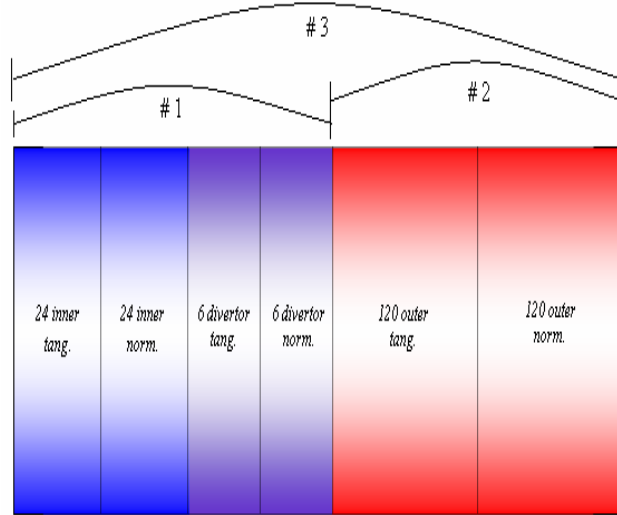A pictorial representation of the three configurations is showed in Figure 2.

Fig. 2 Pictorial representation of the database configuration

III. THE ARTIFICIAL NEURAL NETWORK APPROACH

Artificial Neural Network (ANN) implements a non linear function mapping one multidimensional space, $\{\vec{x}\}$, into another one, $\{\vec{z}\}$ [4]. This function has a predefined structure but contains several parameters which are going to be determined during the training phase which consists in the evaluation of the parameters which minimize the differences between the target output $\vec{t}$ and the network output, $\vec{z}$. Among several possible structures of the network, we use a, so called, feed-forward multilayer perceptron model.

This kind of network is known to approximate arbitrarily any continuous multi dimensional mapping [5].

The hth-component of the vector output (h=1, nz), can be written as

$$z_h = F(\sum_{i=1}^{ny} WY_{hi} y_i) \qquad (1)$$

$$y_i = F(\sum_{j=1}^{nx} WX_{ij} x_j) \qquad (2)$$

Where:

y_i is the ith-component of the output of the first layer;

nx, ny and nz are the dimension of the input vector, the number of the hidden neurons and the dimension of the network output respectively;

297

F is a non-linear function. Typically, it can be a sigmoidal function:

$$(1 + \exp(-a))^{-1} \qquad (3)$$

But other functions can be take into account. In each layer, the input variable to the specific layer is transformed first linearly, by means of a matrix (*WX* and *WY* for the first and the second layer respectively) and then by a non- linear function. The values of the (nx*ny+ny*nz) unknown elements of the matrixes *WX* and *WY* are found by minimizing an error function of the type:

$$E = 0.5 * \sum_{k=1}^{N} [\vec{z}(\vec{x}_{(k)}, WX, WY) - \vec{t}_{(k)}]^2 \qquad (4)$$

in which the sum is extended to the whole training set. A slow but reliable method to minimize the above equation is known as back-propagation algorithm [6] and consists of evaluating the derivatives of *E* with respect to the elements of the *WX* and *WY* matrixes and corrects the unknown parameters using gradient descendent in the following way:

$$WX_{ij}^{(n+1)} - WX_{ij}^{(n)} = -\delta \frac{\partial E}{\partial WX_{ij}} \qquad (5)$$

where δ is an appropriate learning rate parameter and *n* is the iteration number. Regarding our classification problem, we exploit Multi-Layer Perceptron (MLP). Next section explains the proposed approach.

IV. MULTI-LAYER PERCEPTRON TO CLASSYIFY MAGNETIC MEASUREMENTS IN TOKAMAK REACTOR

Multi-Layer Perceptron (MLP) is useful for classification problem [7] optimizing the solution by means of back-propagation algorithm. The goodness of the achieved results can be evaluated, for example, computing the Root Means Square Error (RMSE). The approach is designed according the following procedure:

1. Training phase: the set of input variables is represented by a reduced sub-set extracted from third configuration (see section II) that takes into account

TABLE I
CODIFY OF OUTPUT

Inner Tangential	$[0 \ 0 \ 0]$	Inner Normal	$[0 \ 0 \ 1]$
Divertor Tangential	$[0 \ 1 \ 0]$	Divertor Normal	$[1 \ 0 \ 0]$
Outer Tangential	$[0 \ 1 \ 1]$	Outer Normal	$[1 \ 1 \ 1]$

300 variables and 800 cases;

2. Validation and testing phases: two databases (300x800) extracted from third configuration.

3. The classification is carried out by means of a codify reported in Table I. Each kind of measurement is associated to a sequence of zero and unity.

The MLP configuration, visualized in Figure 3, has given the better performance. Its characteristics are reported in the following lines:

- Input layer of 800 neurons;
- Output layer of 3 neurons (array of zero and unity);
- 2 hidden layer of 35 and 45 neurons respectively;
- Non-linear function is sigmoid;
- Learning rate lr= 0.01;
- Minimum gradient min_grad= 10^{-22};
- Epochs= 500;
- Goal= 10^{-10};

The goodness of the results are tested with three different database:

- Database A (so-called A Class): obtained with a random permutation of the input vector exploited in training phase;

- Database B (so-called B Class): adding some new variables to Database A;

- Database C (so-called C Class): a new Database has been take into account.

Table II reports a summary of obtained results in terms of RMSE in which, for each class, the left side is referred to tangential variables, whereas the right side to normal variables. Figure 4 reports, for A Class, the obtained output (red points) and wanted ones. Notwithstanding the low value of convergence, the reliability of the nets is very poor.

In addition, from any dataset, the classifier is not able to extract information concerning the kind of inner and divertor

TABLE II
SUMMARY OF OBTAINED RESULTS BY MEANS OF MLP APPROACH

		A CLASS		B CLASS		C CLASS	
# Errors		100/300		133/300		134/300	
# of variables correctly classify	In	0	0	0	0	0	0
	Div	0	0	0	0	0	0
	Out	106	94	88	79	84	82
% error		33.3		44.3		44.6	

sensors. In this way, our attention is addressed to an alternative approach.
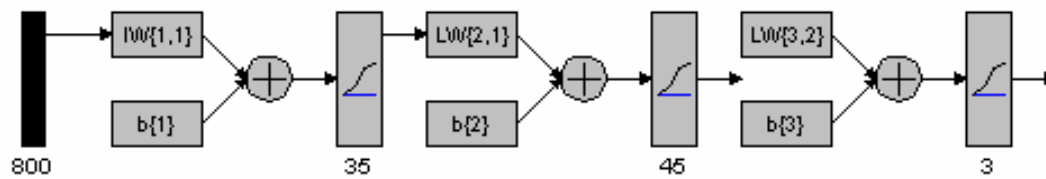
Fig. 3 Classification of magnetic measurements: the exploited MLP network

Fig. 4 Classification of magnetic measurements for A Class:
wanted versus obtained values

the better performance. Its characteristics are reported in the following lines:

- Input layer of 800 neurons;
- Output layer of 3 neurons (array of zero and unity);
- 2 hidden layer of 35 and 45 neurons respectively;
- Non-linear function is sigmoid;

Table III reports a summary of obtained results in terms of RMSE in which, for each class, the left side is referred to tangential variables, whereas the right side to normal variables.
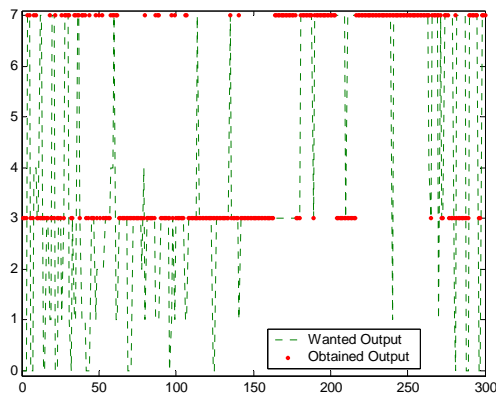
Figure 6 shows reports, for A Class, the obtained output (red points) and wanted ones. Notwithstanding the low value of convergence, the reliability of the nets is very poor. In addition, from any dataset, the classifier is not able to extract information concerning the kind of inner and divertor sensors. In this case, the reliability of the net is improved.

TABLE III
SUMMARY OF OBTAINED RESULTS BY MEANS OF ALTERNATIVE APPROACH

		A CLASS		B CLASS		C CLASS	
# Errors		48/240		53/240		65/240	
# of variables correctly classify	In	15	17	14	17	13	16
	Div	6	4	6	4	6	3
	Out	64	86	60	86	52	85
% error		20		22		27	

IV. AN ALTERNATIVE APPROACH FOR CLASSIFYING MAGNETIC MEASUREMENTS IN TOKAMAK REACTORS

In plasma physics, outer variable is very important to control the position of plasma in terms of its shape and position inside the vacuum vessel whereas inner and divertor variables are found in the inaccessible location. In this section of the paper, we propose a NN approach to stored outer variables as inner or divertor ones (normal and tangential). First step of the alternative approach consists to select only the outer parameters reducing the size of database from 300 to 248 variables. Then, we have associated outer sensors (normal and tangential) to inner-divertor ones that present similar behavior. This association has been carried out by means of computation of Root Means Square (RMS) on each possible couple of variables (tangential inner- tangential outer, tangential divertor-tangential outer, normal inner- normal outer, normal divertor – normal outer); if RMS is a minimum value, then that outer variables can be stored as inner/divertor variable. Once the transformation takes place, we exploit the procedure described above.

The MLP configuration, visualized in Figure 5, has given

V. CONCLUSIONS

In this paper, NNs for classifying magnetic measurements in Tokamak reactors are presented. Particularly, addressing our attention on ITER configuration, we have exploited MLP nets in order to solve the problem under study. The exploited approach shows a very strong adaptability of NNs with respect to the originally database. The improvement of results is carries out transforming some outer variables in inner or divertor ones (normal and tangential). Figure 7 shows the comparison of the results in terms of RMSE percentage.
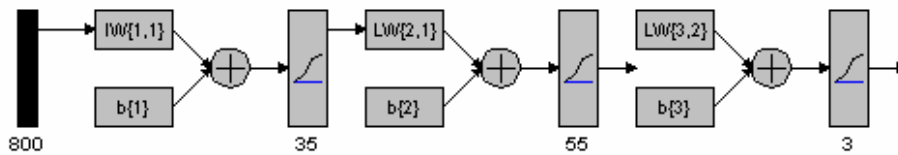
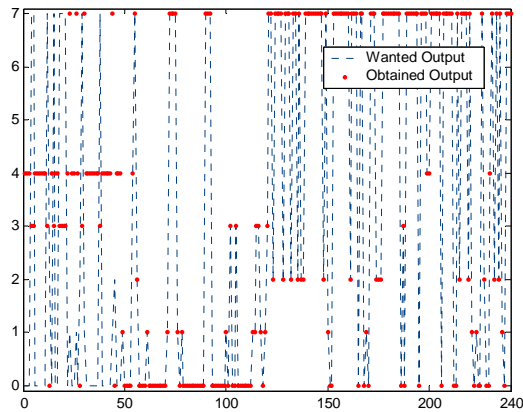Fig. 5 Neural Network alternatively used

Fig. 6 Classification of magnetic measurements by means of
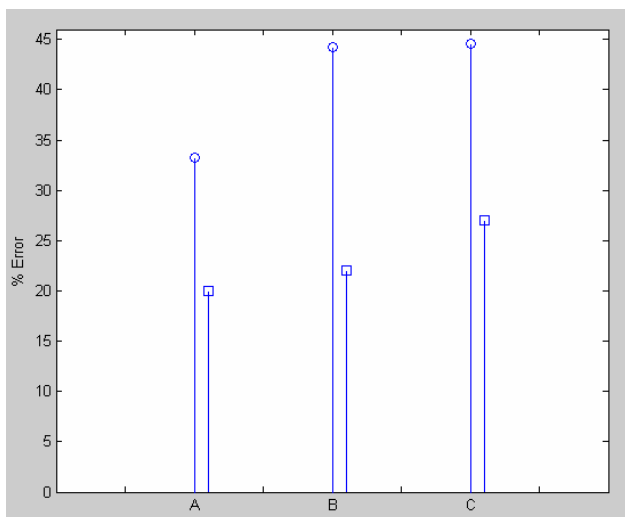alternative approach: visualization of the obtained results

Fig. 7 Comparison of the results in terms of RMSE. The alternative
approach (squared point) performs better than the approach where no
pre-processing takes place

.

REFERENCES

[1] J. Wesson, "Tokamaks", Oxford Science Pub., 1987.
[2] M. Matsukawa, et al., "Application of Regression Analysis to Deriving
     Measurements Formulas for Feedback Control of Plasma Shape in JT-
     60", *Plasma Physics and Controlled Fusion*, vol. 34, N° 6, pp. 907-921,
     1992.
[3] F.C. Morabito et al. "Final report on EFDA Study Contract FU05 CT
     2002-00162 (EFDA 02-1001)", 2002.
[4] C. M. Bishop, "Neural Network for Pattern recognition", Clarendon
     Press, Oxford, 1995.
[5] F.C. Morabito, et al., "On Line Plasma Shape Identification in a
     Tokamak Reactor via Neural Network", Proc. Of V Italian Workshop on
     Neural Networks, Word Scientific Publishing, p. 349, 1992.