Artificial Neural Networks and Multi-Class Support Vector Machines 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 for classifying measurements of plasma columns in Tokamak reactors for nuclear fusion applications. The proposed method exploits several concepts derived from soft computing theory. In particular, Artificial Neural Networks and Multi-Class Support Vector Machines 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 compared with earlier methods.

***Keywords*—**Tokamak, Classification, Artificial Neural Network, Support Vector Machines.

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 kind of Magneto-Hydro-Dynamic (MHD) equilibrium of the plasma changes during the discharge. This is the case, for example, of a 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 occurs) [3]. During this transition, some geometric parameters meet an abrupt discontinuity in the mapping first derivative that is traditional approaches use piecewise linear approximation aiming to separate the problem in more parts each competing to a particular plasma category. The identification problem can be formulated as the search of a suitable mapping between the set of available measurements (sampled by means of sensors located around the chamber contour) and the selected set of shaping parameters. The problem is the determination of a description of the plasma column evolving in the vacuum chamber of a Tokamak reactor, in terms of its position, shape and current profile parameters. Because of the non-linearity of the governing PDE for MHD equilibrium, the full equilibrium identification procedure is normally a computationally intensive task involving the iterative fitting of the raw data to trial equilibrium. For typically available processors, the required computational time is in the order of one second. Techniques which are able to fit this need have been proposed in the past, based on statistical approaches (Functional Parametrization, FP) or Artificial Neural Networks (ANN's).

In the recent years, the neural computing approach has emerged as a successful framework for fast analysis of multi-channel data in plasma shape recognition [4], [5], [6]. A different viewpoint of the identification approach is concerning about the possibility of selecting, among a pool of candidates, the optimal set of sensors in order to achieve the best accuracy of the identification step. For present experimental reactors, the selection of the most important sensors, aims to reduce the computational complexity for real time applications. For future reactors, the determination of the most important sensors can be used in the design phase. In this paper, we focus our attention on ITER configuration (Fig. 1) in which dislocation of outer, inner and divertor sensors

Manuscript received July, 15 , 2005. This work was supported in part by the Dipartimento di Informatica Elettronica Trasporti, DIMET, Facoltà di Ingegneria, Università "Mediterranea" degli Studi, Reggio Calabria, I-89100, Italy.

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).

take place around the vacuum vessel (Fig. 2). In order to reduce the computational complexity, we propose a novel approach, based on soft computing approach and designed by the MatLab® Toolboxes, to classify magnetic measurements (inner, outer and divertor probes). In particular, special ANNs have been exploited for our purpose. To improve the procedure, due to the fact that the inner sensors are inaccessible, an alternative approach is proposed by means a sort of equivalence between outer measurements and inner-divertor ones. That equivalence can be necessary because it is possible to replace a set of sensors that do not work by another equivalent one. Finally, Multi-Support Vector Machines (M-SVMs) have been taken into account to reduce the global error of the procedure.

The paper is organized as follows. Section II reports an overview of the exploited numerical database. After a short description on ANNs and M-SVM reported in section III, we describe the proposed approach for classifying magnetic measurements by means of ANNs (sections IV). Sections V and VI show the procedures to drive to the improvement of the results by means of ANNs and M-SVMs respectively. Finally, some conclusions are drawn.

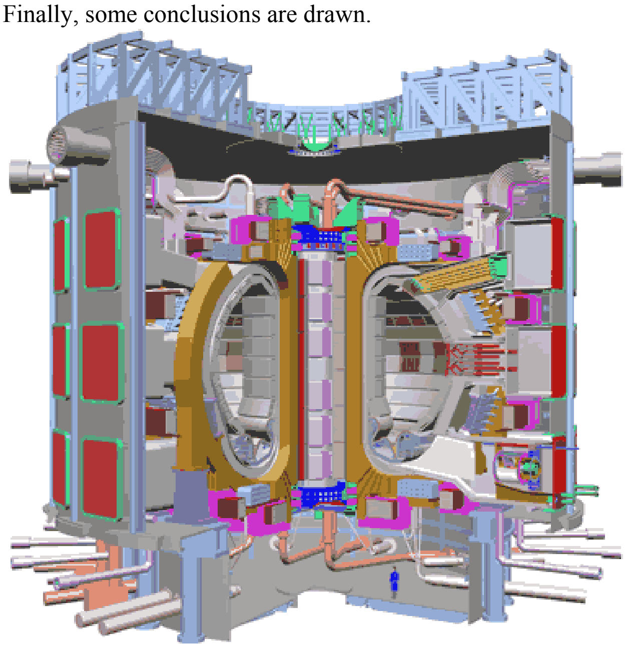

Fig. 1 Pictorial representation of ITER machine

II. The ITER Numerical Database: An Overview

The magnetic diagnostic system provides the main electromagnetic parameters of the plasma and includes sets of full and differential loops and tangential and normal pick-up coils on different poloidal contours given by the structural elements of the machine including the divertor cassette. Using the ITER coil and vessel geometry (Fig. 2), including the 6 dominant passive current eigenmodes, a database of 4848 lower single null equilibria (lower X-point) 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 [7], 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 centre of torus, is 5.3 Tesla. The magnetic measurements, deriving from the sensors located along the contour of the chamber, are subdivided 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.

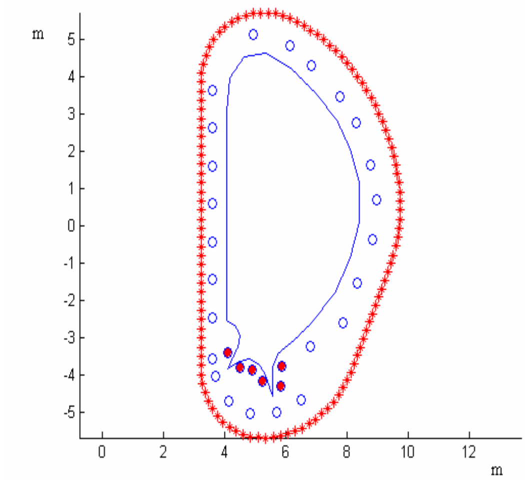

Fig. 2 The cross-section of the ITER Configuration with tentative location of the outer (star), inner (circles) and divertor (bold circles) sensors around the vacuum vessel

Our analysis considers three different configurations of inputs (Fig. 3). 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:** all the 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, the equivalence between outer measurements and inner-divertor ones is showed.

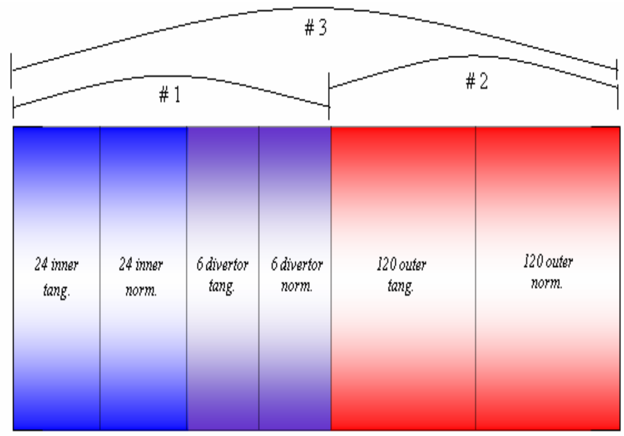

Fig. 3 Pictorial representation of the exploited database

III. Soft Computing Approach for Classifying Magnetic Measurements in Tokamak Reactors

A. The Artificial Neural Network Approach

Artificial Neural Network (ANN) implements a non linear function mapping one multidimensional space, $\{\vec{x}\}$, into another one, $\{\vec{z}\}$ [8]. 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 [9].

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

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

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

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;

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

$$\left(1+\exp(-a)\right)^{-1} \quad (3)$$

But other functions can be take into account [8]. 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:

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

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 consist of evaluating the derivatives of *E* with respect to the elements of the *WX* and *WY* matrixes and correct the unknown parameters using gradient descendent in the following way:

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

where δ is an appropriate learning rate parameter and *n* is the iteration number. Regarding our classification problem, we exploit Multi-Layer Perceptron (MLP) as reported in Section IV.

B. The Multi Support Vector Machines Approach

Support Vector Machine (SVM) was initially designed for binary (two class) problems where a support vector classifier attempts to locate a hyper-plane that maximizes the distance from the members of each class to the optimal hyper-plane. Assume that the training data with *k* number of samples is represented by $\{x_i, y_i\}$, i = 1, …, *k*, where $x \in \Re^n$ is an *n*-dimensional vector and $y \in \{-1, +1\}$ is the class label. These training patterns are linearly separable if a vector *w* (determining the orientation of a discriminating plane) and a scalar *b* (determine offset of the discriminating plane from origin) can be defined so that inequalities (6) and (7) are satisfied.

$$w * x_i + b \geq 1 \quad if \quad y_i = 1 \quad (6)$$

$$w * x_i + b \leq -1 \quad if \quad y_i = -1 \quad (7)$$

The aim is to find a hyper-plane which divides the data so that all the points with the same label lie on the same side of the hyper-plane. This amounts to finding w and b so that:

$$y_i (w * x_i + b) \geq 0 \qquad i = 1,...N \quad (8)$$

If a hyperplane satisfying (8), the two classes are linearly separable. In this case, it is always possible to rescale w and b so that:

$$\min_{1<i<k} y_i (w * x_i + b) \geq 1 \quad (9)$$

That is, the distance from the closest point to the hyper-plane is 1/||*w*|| . Then (8) can be written as:

$$y_i (w * x_i + b) \geq 1 \quad (10)$$

The hyper-plane for which the distance to the closest point is maximal is called the *Optimal Separating Hyper-plane* (OSH). If the data are not linearly separable, a slack variable $\xi_i \; i = 1,...k$ can be introduced with $\xi_i \geq 0$ [10] such that (10) can be written as:

$$y_i(w * x_i + b) - 1 + \xi_i \geq 0 \quad (11)$$

and the solution to find a generalized OSH, also called a soft margin hyper-plane, can be obtained using the conditions:

$$\min_{w,b,\xi_1,\ldots,\xi_k}\left[\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{k}\xi_i\right] \quad (12)$$

$$y_i(w * x_i + b) - 1 + \xi_i \geq 0 \quad (13)$$

$$\text{and } \xi_i \geq 0 \qquad i = 1,\ldots,k \quad (14)$$

The first term in (12) is same as in the linearly as in the not linearly separable case, and controls the learning capacity, while the second term controls the number of misclassified points. The parameter C is chosen by the user. Larger values of C imply the assignment of a higher penalty to errors. When it is not possible to have a hyper-plane defined by linear equations on the training data, the techniques described above for linearly separable data can be extended to allow for non-linear decision surfaces. A technique introduced in [11], maps input data into a high dimensional feature space through some nonlinear mapping. The transformation to a higher dimensional space spreads the data out in a way that facilitates the finding of linear hyper-planes. In this case, a ϕ function, called kernel function, is used to map x in the feature space $\phi(x)$, in order to early find the OSH. A number of kernel functions are used for support vector classifier [10], [11]. When dealing with several classes, an appropriate multi-class method is needed. In this study, One-against-all and one-against-one methods are presented.

1. One-Against-All and One-Against-One

The earliest used implementation for SVM multi-class classification is probably the *one-against-all method* [10], [11]. It constructs k SVM models where k is the number of classes.. The i^{th} SVM is trained with all of the examples in the i^{th} class with positive labels, and all other examples with negative labels. Thus given l training data $(x_1, y_1),\ldots,(x_l, y_l)$, where $x_i \in R^n$, $i=1,\ldots,l$ and $y_i \in \{1,\ldots,k\}$ is the class of x_i, the i^{th} SVM solves the following problem:

$$\begin{aligned}
&\min_{\omega^i, b^i, \xi^i} \frac{1}{2}(\omega^i)^T\omega^i + C\sum_{j=1}^{l}\xi_j^i \\
&(\omega^i)^T\phi(x_j) + b^i \geq 1 - \xi_j^i, \quad if \;\; y_j = i \\
&(\omega^i)^T\phi(x_j) + b^i \leq -1 + \xi_j^i, \quad if \;\; y_j \neq i \\
&\xi_j^i \geq 0, \qquad j = 1,\ldots,l
\end{aligned} \quad (15)$$

where the training data x_i are mapped to a higher dimensional space by the function Φ and C is the penalty parameter. Minimizing $0.5(\omega^i)^T\omega^i$ means that we would like to maximize $2/\|\omega^i\|$, the margin between two groups of data. When data are not linear separable, there is a penalty term $C\sum_{j=1}^{l}\xi_j^i$ which can reduce the number of training errors. The basic concept behind SVMs is to search for a balance between the regularization term $0.5(\omega^i)^T\omega^i$ and the training errors. There are k decision functions to resolve:

$$\begin{aligned}
&(\omega^1)^T\phi(x) + b^1 \\
&\ldots \\
&(\omega^k)^T\phi(x) + b^k
\end{aligned} \quad (16)$$

We say x is in the class which has the largest value of the decision function

$$\text{class of } x \equiv \text{argmax}_{i=1,\ldots k}\,((\omega^i)^T\phi(x_j) + b^i) \quad (17)$$

Practically we solve the dual problem of (18) whose number of variables is the same as the number of data in (18). Hence k l-variable quadratic programming problems are solved. Another major method is called the *one-against-one method* [10], [11]. This method constructs *k(k-1)/2* classifiers where each one is trained on data from two classes. For training data from the i^{th} and the j^{th} classes, we solve the following binary classification problem:

$$\begin{aligned}
&\min_{\omega^{ij}, b^{ij}, \xi^{ij}} \frac{1}{2}(\omega^{ij})^T\omega^{ij} + C\sum_t \xi_t^{ij} \\
&(\omega^{ij})^T\phi(x_t) + b^{ij} \geq 1 - \xi_t^{ij}, \quad if \;\; y_t = i \\
&(\omega^{ij})^T\phi(x_t) + b^{ij} \leq -1 + \xi_t^{ij}, \quad if \;\; y_t = j \\
&\xi_t^{ij} \geq 0
\end{aligned} \quad (18)$$

There are different methods for doing the future testing after all *k(k-1)/2* classifiers are constructed. After some tests, we decide to use the following voting strategy: if sign $((\omega^{ij})^T\phi(x_t) + b^{ij})$ says x is in the i^{th} class, then the vote for the ith class is added by one. Otherwise, the j^{th} is increased by one. Then we predict x is in the class with the largest vote. The voting approach described above is also called the "Max Wins" strategy. In case that two classes have identical votes, thought it may not be a good strategy, now we simply select the one with the smaller index. Practically we solve the dual of (14) whose number of variables is the same as the number of data in two classes. Hence if in average each class ha *l/k* data points, we have to solve *k(k-1)/2* quadratic programming problems where each of them has about *2l/k* variables.

IV. Multi-Layer Perceptron to Classify 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 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 codification:** reported in Table I. Each type of measurement is associated to a sequence of zero and unity (output of procedure).

TABLE I
MLP APPROACH: CODIFICATION 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]

The MLP configuration, visualized in Fig. 4, 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 as reported in Table I);
- 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 by means of three different database:

- Database A (so-called A Class): obtained by means of 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 taken into account.

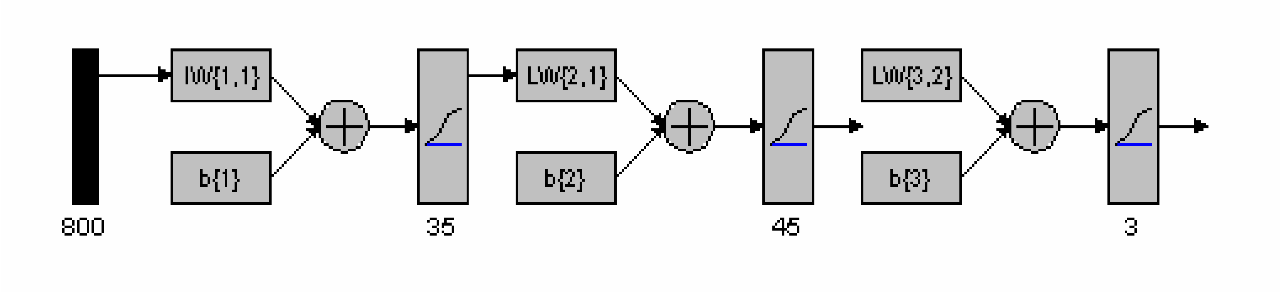

Fig. 4 Classification of magnetic measurements: the exploited MLP network

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. In particular, the classification of Inner and Divertor sensors is more difficult because their location inside the vacuum vessel is next to the plasma contour.

TABLE II
MLP APPROACH: SUMMARY OF OBTAINED RESULTS

		A CLASS		B CLASS		C CLASS	
# Errors		*100/300*		*133/300*		*134/300*	
# 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**	

Fig. 5 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 way, our attention is addressed to an alternative approach.

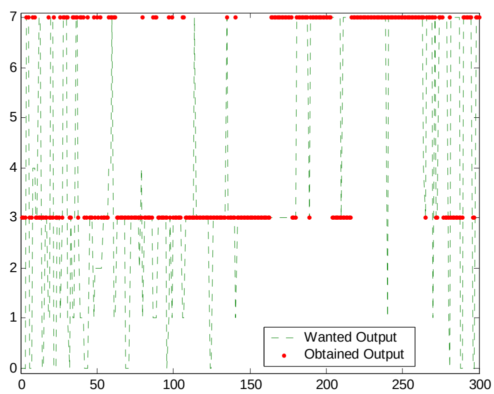

Fig. 5 Classification of magnetic measurements for A Class A: wanted values and obtained values

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

In plasma physics, outer variable is very important to control the position of plasma 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). This procedure is also suggested from the fact that, in ITER machine, outer sensors are not physically accessible. 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 associate outer sensors (normal and tangential) to inner-divertor ones that present similar behavior. This association has been carried out by means of the exploitation of a special MLP net and computing the 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 Fig. 6, 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;

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. In this case, by means of the alternative approach, the classification error is lower and, in addition, it is possible classify Inner and Divertor probes.

TABLE III
ALTERNATIVE APPROACH: SUMMARY OF OBTAINED RESULTS

		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**	

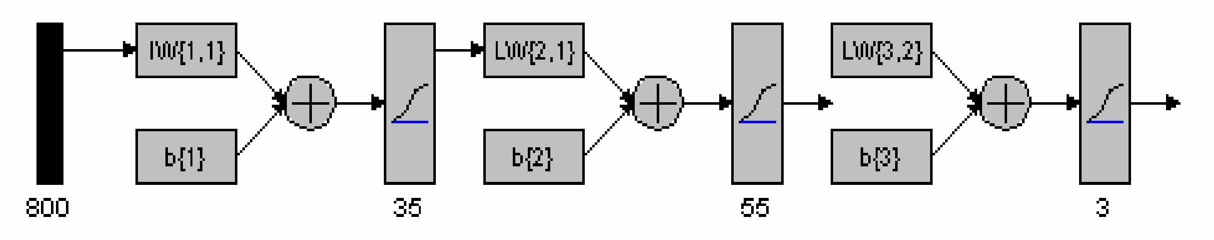

Fig. 6 ANN for classifying magnetic measurements by means of the alternative approach

Fig. 7 shows, for A Class, the obtained output (red points) and wanted ones. In this case, the reliability of the network is improved.

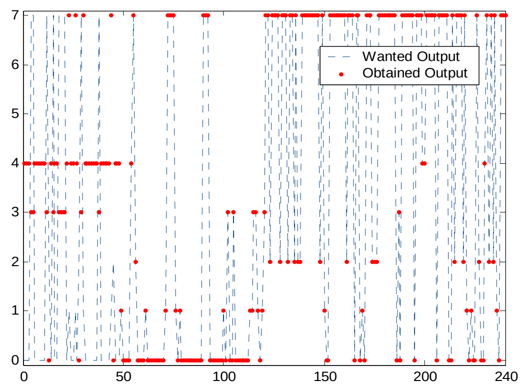

Fig. 7 Classification of magnetic measurements by means of alternative approach (A Class): visualization of the obtained results

VI. MULTI-SUPPORT VECTOR MACHINE TO CLASSIFY MAGNETIC MEASUREMENTS IN TOKAMAK REACTOR

In the previous sections, an ANN has been trained to classify magnetic probes in ITER configuration: its low performance has addressed us to conceive a novel method, based on a sort of equivalence among sensors, that reduces the global error. In this section, M-SVMs are exploited to improve the obtained results. As reported in section IV, we distinguish three phases: a) training, testing phases, where we exploit the third configuration of database (see Section II);
b) classification phase, in which A, B and C Classes have been used.

Table IV visualize the exploited codification for that case (output of the procedure).

The characteristics of the proposed M-SVM are reported in the following lines:

- polynomial kernel

$$K(x_i,x_j)=(\gamma\|x_i-x_j\|+C)^{Degree},\gamma>0 \qquad (16)$$

  which has given the best performance with the lowest computational load;
- γ, C: coefficients of the polynomial kernel;
- degree of the polynomial kernel;
- number of samples for pattern training;
- the labels of codifications;

- the value of the penalty function (slack variables);
- number of Support Vectors (*nSVs*) for each class *L* (*L*=1,…6);

In Table V, the parameters of classifier, referred to this approach, take place, where the associations between codification of each class and nSV are showed.

TABLE IV
MSVMs Approach: The Exploited Codification

Class	Codification
Inner Tangential	*1*
Inner Normal	*2*
Divertor Tangential	*3*
Divertor Normal	*4*
Outer Tangential	*5*
Outer Normal	*6*

TABLE V
MSVMs Approach: Parameters of Classifier

nLabel					
1	*6*	*4*	*5*	*2*	*3*

nSV					
Inner Tang	*Outer Norm*	*Div. Norm*	*Outer Tang*	*Inner Norm*	*Div. Tang*
23	*77*	*6*	*48*	*24*	*6*

In Table VI we report the results achieved by means of the M-SVM system described above concerning the classification of magnetic measurements in ITER configuration. Obviously, the classification by means of M-SVMs produces the best results with respect to ANN and alternative approach ones because the learning procedure generates an optimal number of Support Vector through automatic extraction from database (Fig. 8).

TABLE VI
M-SVM Approach: Summary of Obtained Results

		A Class		B Class		C Class	
# Errors		*5/300*		*12/300*		*49/300*	
# of variables correctly classify	In	*20*	*24*	*24*	*20*	*14*	*16*
	Div	*5*	*6*	*6*	*6*	*5*	*5*
	Out	*120*	*120*	*120*	*112*	*106*	*105*
% error		**1.7**		**4**		**16.3**	

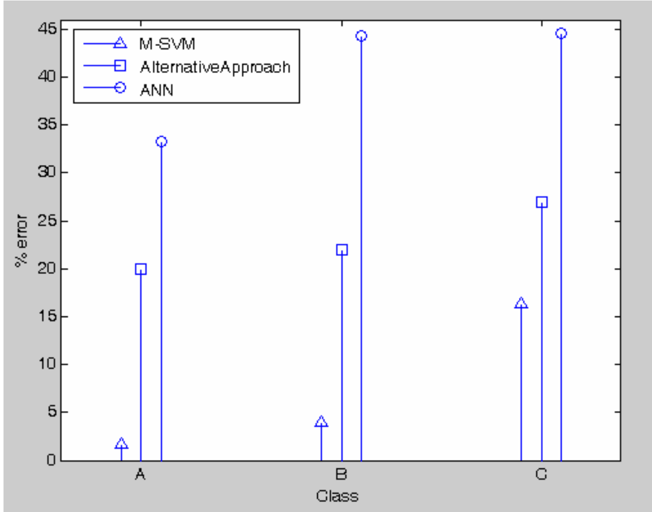

Fig. 8 Comparison of the achieved results in terms of percentage error. The alternative approach (squared point) performs better than the approach where no pre-processing takes place (circled point). M-SVM is the better method for this applications

VII. Conclusion

In this paper, we have proposed the use of ANNs and M-SVMs for classifying of magnetic measurement for ITER configuration in Tokamak Reactors. The study case is derived from the database which was made available through the the Plasma Data Analysis Group (PDAG), Physics Department, University College Cork, Association EURATOM-DCU. In particular, special MLP have been eused for that purpose. The improvement of the procedure has been carried out by means of a sort of equivalent between outer measurements and inner-divertor ones replacing set of sensors that could not work by another equivalent one. Finally, the exploitation of M-SVMs allows us to reduce the global error of the designed procedure.

Acknowledgement

We sincerely thank Association EURATOM-ENEA-CREATE-DCU and ITER Team of IPP of Plasma Data Analysis Group (PDAG), Physics Department, University College Cork, Ireland for kindly provide us the experimental database.

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, M. Versaci, "Fuzzy Inference Systems (FIS) for Classification Identification of Plasma Columns in Tokamak Reactors", International Journal of Chaos Theory and Applications, Vol. 5, N°1, pp. 11-40, 2000.
[4] P.J. McCarthy, F.C. Morabito, "Function Parametrization and Artificial Neural Networks for Equilibrium Parameters Recovery in ASDEX-Upgrade: a Comparison", Journal of Applied Electromagnetics and Mechanics, Vol. 9, pp. 1-31, 1997.
[5] F.C. Morabito, M. Campolo, "Are Hybrid Fuzzy-Neural Systems Actually Useful in Plasma Engineering?, Proc. Of the 9th Italian

Workshop on Neural Nets, M. Marinaro and R. Tagliaferri Eds., WIRN'97, Vietri S/M, pp. 22-24, 1997.
[6] F.C. Morabito, M. Versaci, "A Fuzzy-Neural Approach to Real Time Plasma Boundary Reconstruction in Tokamak Reactors", IEEE ICNN International Conference on Neural Networks, Houston, Texas, pp. 43-47, June 1997.
[7] F.C. Morabito et al. "Final report on EFDA Study Contract FU05 CT 2002-00162 (EFDA 02-1001)", 2002.
[8] C. M. Bishop, "Neural Network for Pattern recognition", Clarendon Press, Oxford, 1995.
[9] 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.
[10] B.Cortes, V.N. Vapnik, Support vector networks, Machine Learning Vol. 20, pp. 273-297, 1995.
[11] C. W. Hsu, C.-J. Lin, "A comparison of methods for Multi-class Support Vector Machines", Department of Computer Science and Information Engineering, National Taiwan University.