Forecast of TEXT Plasma Disruptions Using Soft X-Rays as Input Signal in a Neural Network

A. Vannucci,*
K.A. Oliveira,* and T. Tajima
Institute for Fusion Studies, The University of Texas at Austin
Austin, Texas 78712 USA

March 3, 1998

Abstract

A feed-forward neural network with two hidden layers is used in this work to forecast major and minor disruptive instabilities in TEXT discharges. Using soft X-ray signals as input data, the neural net is trained with one disruptive plasma pulse, and a different disruptive discharge is used for validation. After being properly trained the networks, with the same set of weights, is then used to forecast disruptions in two others different plasma pulses. It is observed that the neural net is able to predict the incoming of a disruption more than 3 ms in advance. This time interval is almost three times longer than the one already obtained previously when magnetic signal from a Mirnov coil was used to feed the neural networks with. To our own eye we fail to see any indication of an upcoming disruption from the experimental data this far back from the time of disruption. Finally, from what we observe in the predictive behavior of our network, speculations are made whether the disruption triggering mechanism would be associated to an increase of the $m = 2$ magnetic island, that disturbs the central part of the plasma column afterwards or, in face of the results from this work, the initial perturbation would have occurred first in the central part of the plasma column, within the $q = 1$ magnetic surface, and then the $m = 2$ MHD mode would be destabilized afterwards.

*Instituto de Fisica, Universidade de Sao Paulo, C.P. 66318, Sao Paulo, 05315-970, SP, Brazil

1 Introduction

Disruptive instabilities in tokamaks are still a major concern and considered a great obstacle in achieving thermonuclear fusion using a tokamak reactor. The next large fusion devices, because of the large current values and high plasma temperature to be achieved [1], will have to be very carefully projected and properly constructed in order to support the occurrence of a large number of such instabilities if some proper mechanism has not been discovered by then to avoid them or at least minimize their harmful effects. Neural network forecasting, in this respect, might be one possible solution to this problem if it proves to be able to alert the incoming of a disruption in a period of time sufficiently long in order to allow defensive mechanisms to be turned on.

Nowadays, it is widely accepted that the $m = 2$ component of the magnetohydrodynamic (MHD) activity of the plasma plays an important role in starting the disruptive process in tokamaks. Many papers have already suggested, for example, that a MHD mode coupling would be the main mechanisms for triggering both minor and major disruptions [2,7].

The first paper published exploring the idea of forecasting disruption using a neural network showed that major and minor disruptions could be forecasted more than 1 ms before they actually take place, when magnetic data obtained from one Mirnov magnetic coil was used to feed the neural net [8]. This period of time, however, was not considered to be sufficiently large to allow any proper controlling mechanism such as electron resonant heating [9,10], neutral beam [11], external resonant magnetic fields [12,13], or even impurity pellets injection [14] to be used in real time, in a possible attempt to suppress the upcoming disruption.

More recently, another paper was published reporting the use of neural network to estimate the high-β limit in DIII-D by combining signals from a large number of plasma diagnostics, resulting in a more accurate prediction of the high-β disruption boundary than

that provided by the traditional Troyon limit [15]. After the instability boundaries have been accurately mapped, the neural network was reported to be able to alert in real time the controlling systems that those boundaries were about to be reached. The limitation of this method of forecasting disruption in a future reactor, for example, appears to be the necessary large number of disruptions that would be necessary to occur beforehand in order to create a database that would then allows, afterwards, those instability boundaries to be properly set.

In this work, the result of forecasting disruptions in TEXT tokamak by feeding a neural networks with soft X-ray signals is presented. After being properly trained, the neural net was able to forecast the occurrence of disruptions more than 3 ms before it happens. This period of time is almost the triple of that obtained previously by using magnetic signal from a Mirnov coil [8].

2 Neural Networks

Artificial neural networks are computer algorithms which simulate, in a very simplified form, the ability that brain neurons have to process information. A typical perceptron neural net is constituted by a certain number of binary units organized in layers, usually 3 or more. In a feed-forward network, each neural unit is connected with all units from the previous layer and for each connection there is a specific statistical weight related to it [16,17]. Within each unit of the network, all the input weighted signals are summed and an excitatory or inhibitory signal is then fired to the next layer's units, depending whether the result of the sum has reached or not a certain threshold value, which is defined by the activation function chosen (Fig. 1). These weights are adjusted (or educated) to minimize error in prediction (back propagation [18]).

Observing the time delayed vector of a physical quantity X at time t, of the tokamak

data:

$$X_t = [x_t, x_{t-\tau}, x_{t-2\tau}, ..., x_{t-n\tau}] \tag{1}$$

then it is reasonable to suppose that the future state of the system, at time $t+\tau$, could be predicted by a smooth nonlinear function F:

$$x(t+\tau) = F(X_t) \tag{2}$$

where $\tau = 0.04\,\mathrm{ms}$ corresponds, in this work, to the sampling rate of the CAMAC acquisition system. If the dynamics of the system under study is a low dimensional one, it is hoped that such a function may be found even for stiff functions of dynamics such as the sudden onset of disruption.

However, since the function F is not known, the idea is to alternatively use a neural network to approximate F and, therefore, predict the future evolution of the system. In order to do its job, n in Eq. (1) must be smaller than the dynamical dimension d of the system, and the neural net must be trained first, that is, the correct set of weights for all connections must be found. When we start weights in tabula rasa, we need to "educate" these weights. The training process basically consists in feeding the neural net repeatedly with experimental data, from a given plasma pulse, and in comparing the output signal O (forecasted) with the real signal O (experimental data in our case). In this process, the back propagation algorithm [18] is used to minimize the error function for the current weights $\{W_{ij}\}$:

$$\varepsilon\{W\} = 1/M\Sigma\left(\mathcal{O}_j - \widehat{\mathcal{O}}_j\{W\}\right)^2 \tag{3}$$

which evaluates the efficacy of the training after each training epoch. The minimization process is basically that of the Newton method based on the gradient of ε with respect to the element of weight W_{ij}.

In a single-step process the soft X-ray data points are predicted one time step ahead only ($\tau = 0.04\,\mathrm{ms}$), while in multisteps predictions the predicted output at time $t+\tau$ is fed back

into the input and is used to predict a new output at time $t + 2\tau$, which is fed back into the input together with the values previously predicted, in order to predict a new output at time $t + 3\tau$, and so on.

The neural net architecture used in this work had the configuration $(15-9-3-1)$, that is, 15 neural units in the input layer, 9 units in the first hidden layer, 3 units in the second hidden layer and only one unit in the output layer. The activation functions were chosen to be $g(x) = \tanh(x)$ for all the hidden units and $g(x) = x$ for the output unit. The training of the network and the disruption prediction was carried out over the last 200 ms of the plasma discharges.

3 Disruption Forecast

In order to find the adequate weights for all the connections between the neural units, two different disruptive plasma pulses have been used: the training set and the validation set. The single and multistep forecasting processes are performed, afterwards, over two others disruptive pulses, distinct from the ones used for training and validation.

In Fig. 2 the basic experimental signals, related to the first one of the plasma disruptive discharges used for forecasting, are shown. This pulse corresponds to an $I_p \approx 170\,\mathrm{kA}$ plasma discharge that disrupted at $t \approx 470\,\mathrm{ms}$. About 18 ms before the major disruption a minor disruption occurred, as observed in the Mirnov magnetic signal [Fig. 2(c)], causing a significant drop on the average electron density [Fig. 2(f)] and electron temperature as observed through the X-ray emission signals [Fig. 2(d)]. This same feature, that is, the major disruption being preceded by a minor disruption just before the plasma current collapse, is also observed on the plasma discharges used for validation and for training the net.

The result of the forecasting process for this particular plasma shot in shown in Fig. 3. As it can be observed, for one time step-prediction [Fig. 3(a)] the result obtained from the neural networks adjusts almost perfectly with the experimental signal. By increasing the

forecast time interval, i.e., the number of timesteps, as showed in Fig. 3(b) for 25 timesteps, the shape of the sawteeth oscillations is observe to be somewhat deformed but the net is still capable of accurately predicting the instance of the minor disruption that takes place at $t \approx 452\,\mathrm{ms}$. The longest forecasting time achieved is obtained for 63 timesteps [Fig. 3(c)] for which the net still accurately predicts the occurrence of the minor disruption. This corresponds to a forecasting of the disruptive instability 2.52 ms before it actually takes place. For timesteps longer than that, the net continues predicting the disruption but now with a time shift delay, as shown in Fig. 2(d) for a 90 timesteps forecasting.

Examining in detail the result obtained for the 63 timesteps forecast [Fig. 3(c)], it can be observed that in order to make this prediction, the net used 15 experimental data points that were positioned around $t \approx 449.5\,\mathrm{ms}$. This particular time interval is located exactly in the oscillatory region of the last "typical" sawtooth oscillation, when our eyes are unable to perceive any peculiar occurring in the soft X-ray emission which would signal that an instability has started at that point (or before) and that a disruption in coming soon. Only after the crash of this sawtooth at $t \approx 450.0\,\mathrm{ms}$, when an strong fluctuation starts to build up afterwards, one can say that a disruptive instability indeed has been triggered.

As another test for the neural networks, exactly the same set of weights obtained and used above is now used to forecast the disruption that occurred in a second plasma discharge (Fig. 4). This also corresponds to an $I_p \approx 170\,\mathrm{kA}$ plasma discharge that disrupted at $t \approx 424\,\mathrm{ms}$. Differently from the first discharge analyzed, however, in this particular plasma pulse the major disruption was not preceded by any minor disruption.

The results of the multistep forecasting analysis done for this discharge are shown in Fig. 5. Once again, for a single timestep prediction the result of the net is observed to beautifully match the experimental signal [Fig. 5(a)]. For longer forecast intervals [Fig. 5(b)] the net still accurately predicts the disruption time up to timestep 78 [Fig. 5(c)]. This corresponds to a forecast of the occurrence of the major disruption 3.12 ms in advance,

value which is nearly triple the time that has been obtained previously using only magnetic data to feed the neural net [8]. For larger timesteps, once again a time shift in prediction appears between the experimental signal and the result provided by the network [Fig. 5(d)].

Analyzing in more detail the result obtained for timestep 78 [Fig. 5(c)], we can verify that the experimental data points used by the net to make this prediction correspond to an instance of time around $t \approx 420.6\,\mathrm{ms}$. This corresponds to about 0.5 ms earlier in time before the last sawtooth crash occurs, yielding the strong MHD precursor for the major disruption that follows.

Interestingly, in both prediction cases the experimental data points used by the net to accurately forecast the minor and major disruption [Figs. 3(c) and 5(c)] are located in time prior to the amplitude increase of the magnetic fluctuation in the corresponding magnetic signals, as can be seen in Figs. 6 and 7, respectively.

Since the neural networks was able to forecast disruptions using data points related to some particular instances of time before the observation of an increase in amplitude of the MHD activity, this observation might suggest that it is not the growing magnetic islands related to the $q = 2$ magnetic surface that would consequently disturb the inner island at $q = 1$ magnetic surface. The results obtained in this work suggest that it is probably the other way around, i.e., some disturbance phenomena would develop first around the central part of the plasma column (visualized by the soft X-ray central detector) and then the plasma region within the $q = 2$ magnetic surface would be affected, destabilizing the $m = 2$ MHD mode. This interpretation reminds us the works already done in tokamaks suggesting that the disruptive instabilities would be caused by a "cold bubble" moving towards the plasma center, as it could be experimentally observed in the electron temperature profile [19,20]. However, some more careful and further investigations must be done in order to give more confidence about the reality or not of this process.

4 Forecasting the Entire Plasma Discharge

Since the set of weights used to forecast disruptions in this work was obtained by training the neural network with the last 200 ms of the training and validation plasma discharges, it is reasonable to use only the last 200 ms of the forecasting plasma pulses for predicting the disruption time. Now, if that same set of weights previously obtained is applied to forecast the whole plasma discharge, instead of the last 200 ms only, some strong distortions in the net output signal are observed in comparison to the real signal. This result should be expected, however, once the experimental data corresponding to the ramping up phase of the soft X-ray signal is not used during the training section of the net.

In an attempt to improve the net performance over the whole plasma discharge, therefore, a new training section is performed by now using the entire soft X-ray experimental data of two disruptive plasma discharges. A third discharges is used for validation and finally a fourth one for forecasting. The best results obtained are showed in Fig. 8. For a single timestep prediction [Fig. 8(b)] the match with the experimental signal [Fig. 8(a)] can be considered perfect. For 13 timesteps prediction [Fig. 8(c)], the signal output from the neural net exhibits some oscillations during the ramping up phase of the signal and, very strongly, right after the occurrence of the minor disruption. However, by examining the 13 timestep result more closely in the neighborhood of the disruption time (Fig. 9) it can be observed that the net is still capable of determining the exact instance of time of disruption up to 0.52 ms in advance. Unfortunately, for longer timestep predictions a time shift occurs between the prediction by the net and the real disruption time. This indicates that the proper set of weights has yet to be found. Consequently, either a larger number of experimental data must be used to train the neural net, or the number of training epochs must be increased, or even a different net architecture might be necessary to be tested.

5 Conclusion

It has been shown that feed-forward neural networks can be effectively used to forecast both minor and major disruptive instabilities in tokamaks. It was demonstrated in this work that soft X-ray experimental signals are more appropriate than magnetic data from a Mirnov coil to be used as input signal to the net, for the purpose of disruption forecast. Our forecasting time of minor or major disruptions is almost three times the one based on magnetic data [8]. We also note that the future larger tokamaks have longer plasma time scales than the medium size machines such as TEXT. This opens up a possibility of using feedback controlled auxiliary systems such as electron cyclotron heating, neutral beam, pellet injection, external magnetic fields, etc., to avoid the occurrence of the upcoming disruption or at least to minimize its harmful effects. Also, the forecasting neural net attached to a tokamak can learn by itself through a series of discharges and become "smarter" to be able to steer the plasma away from an unstable domain of operation (a "neural tokamak") [21].

Finally, observing that the soft X-ray experimental data points used by the net in the best forecasting cases are located prior to the instance of the amplitude increase in the Mirnov magnetic signals, we might be able to suggest that the perturbation which triggers the disruptions first initiates in the central part of the plasma column where the $q = 1$ magnetic surface is located and only afterwards this instability would reach the outer part of the plasma column, destabilizing the $m = 2$ MHD mode. Speculations can be made whether this process could be related to some already reported explanations of disruptions that an "cold bubble," pushed towards the center of the plasma, would trigger the disruptive instabilities in tokamaks.

In conclusion, the neural net is capable of predicting the time evolution of tokamak plasmas even for a sudden and violent disruptive occurrence. This predictive capability may be rather surprising, as the prediction time is well before an obvious visual pattern (or

"omen") appears. This predictive capability may be useful for large tokamaks which have long time scales of plasma evolution. The predictive behavior of the net shows that some of the underlying physics of the disruption, as the soft X-ray signals, are better equipped than the magnetic signals as an indication of the possible origin of the disruptive instability. This could portend the discrimination of possible mechanisms of disruptions.

Acknowledgments

The first author would like to thank all the FRC and IFS staff, in particular Don Patterson and Jim Dibble, for the support received. Also the first author is grateful to FAPESP–Fundacao de Amparo a Pesquisa do Estado de Sao Paulo, for the financial support. This work was supported in part by the U.S. Dept. of Energy Contract No. DE-FG03-96ER-54346.

References

[1] STACEY, W.M., Nucl. Fusion **35** (1995) 1369.

[2] VANNUCCI, A., NASCIMENTO, I.C., CALDAS, I.L., Plasma Phys. Contr. Fusion **3** (1989) 147.

[3] McGUIRE, K.M. and ROBINSON, D.C., Phys. Rev. Lett. **44** (1980) 1666.

[4] VANNUCCI, A., and McCOOL, S.C., Nucl. Fusion **37** (1997) 1229.

[5] TSUJ, S., NAGAYAMA, Y., MIYAMOTO, K., *et al.*, Nucl. Fusion **25** (1985) 305.

[6] VANNUCCI, A. and GILL, R.D., Nucl. Fusion **31** (1991) 1127.

[7] PIETRZYK, Z.A., POCHELON A., BEHN R., *et al.*, Nucl. Fusion **32** (1982) 1735.

[8] HERNANDEZ, J.V., VANNUCCI A., TAJIMA, T., LIN, Z., HORTON, W., McCOOL, S.C., Nucl. Fusion **36** (1996) 1009.

[9] HOSHINO, K., MORI, M., YAMAMOTO, T., TAMAI, H., *et al.*, Phys. Rev. Lett. **69** (1992) 2208.

[10] KISLOV, D.A., ALIKAEV, V.V., ESIPCHUK, Yu.V., KAKURIN, A.M., *et al.*, Nucl. Fusion **37** (1997) 339.

[11] YOSHINO, R., NEYATANI, Y., HOSOGANE, N., WOLFE, S.W., *et al.*, Nucl. Fusion **33** (1993) 1599.

[12] HENDER, T.C., FITZPATRICK, R., MORRIS, A.W., CAROLAN, P.G., *et al.*,Nucl. Fusion **32** (1992) 2091.

[13] ARAUJO, M.S.T., VANNUCCI, A., CALDAS, I.L., Il Nuovo Cim. (1996).

[14] PAUTASSO, G., BUCHL, K.. FUCHS, J.C., GRUBER, O., *et al.*, Nucl. Fusion **36** (1996) 1291.

[15] WROBLEWSKI, D., JAHNS, G.L., LEUER, J.A., Nucl. Fusion **37** (1997) 725.

[16] HERTZ, J., KROGH, A., PALMER, R.G., *Introduction to the Theory of Neural Computation*, (Addison-Wesley, Reading, MA, 1991).

[17] HINTON, G., Scientific American, Sept. issue (1992) 145.

[18] RUMELHART, D.E., HINTON, G.E., WILLIAMS, R.J., Nature **323** (1986) 533.

[19] FREDRICKSON, E.D., McGUIRE, K.M., BELL, M.G., BUSH, C.E., *et al.*, Nucl. Fusion **33** (1993) 141.

[20] SCHULLER, F.C., Plasma Phys. Contr. Fusion **37** (1995) A135.

[21] YOSHINO, R., *et al.*, Fusion Technol. **30** (1996) 237.

FIGURE CAPTIONS

FIG. 1. Illustration of how the information is processed within one neuron unit of the neural net. All the input signal (V_i) from the units of the previous layer, after been properly weighted (W_{ij}), are summed and then an inhibitory or oscillatory signal is fired to the next layer's units as defined by the activation function $g(x)$.

FIG. 2. A TEXT disruptive discharge which was used for forecasting. The major disruption takes place at $t \approx 470\,\mathrm{ms}$ and was preceded by a minor disruption that occurs about 18 ms earlier. The signals corresponds to: (a) plasma current, (b) loop voltage, (c) Mirnov magnetic signal, (d) central soft X-ray intensity, (e) safety factor, (f) average electron density.

FIG. 3. Results of the multistep prediction carried out for the plasma discharge shown in Fig. 2. After properly trained, the neural network is capable of forecasting the occurrence of the disruption up to 2.52 ms ahead in time, corresponding to a 63 timestep prediction (c). For longer timesteps, a time shift is observed between the neural net result and the real soft X-ray signal (d).

FIG. 4. Plasma discharge that was also used for forecasting, with the same set of weights used previously. Note that, in this case, the major disruption is not preceded by any minor disruption. The nomenclature is the same as that for Fig. 2.

FIG. 5. Results of the multisteps prediction carried out for the plasma discharge shown in Fig 3. In this case, the neural net is able to predict the occurrence of the major disruption 3.12 ms ahead of time, corresponding to a 78 timestep prediction (c). Once again, a time shift between the result from the neural net and the experimental soft X-ray signal is observed for longer time predictions (d).

FIG. 6. Expanded view of the soft X-ray intensity (a) and Mirnov magnetic signal (b) for

the plasma discharge shown in Fig. 2, close to the disruption time ($t \approx 452\,\mathrm{ms}$). As can be observed, the experimental data used by the net in its best performance are located in time just before the MHD activity starts increasing in amplitude.

FIG. 7. Expanded view of the soft X-ray intensity (a) and Mirnov magnetic signal (b) for the plasma discharge shown in Fig. 4, close to the disruption time ($t \approx 424\,\mathrm{ms}$). Once again it is observed that the experimental data used by the net in its best performance for this discharge are located in time before the MHD activity starts increasing in amplitude.

FIG. 8. Using a different set of weights, obtained by choosing two disruptive pulses for training and a third one for validation, the neural networks was used to forecast the whole discharge over a fourth distinct plasma shot. The single step prediction (b) is observed to match very well the experimental signal (a) while for 13 timesteps (c), although the net is still capable of accurately predicting the exact instance the disruption occurs, some spurious oscillations appear in the beginning and, more strongly, just after the minor disruption at $t \approx 454\,\mathrm{ms}$.

FIG. 9. Expanded view of Fig. 8(c) close to disruption time showing the accuracy of the prediction performed by net for a 13 timesteps forecasting. For longer time predictions a time shift appears, probably indicating that the net was not yet adequately trained.