

Real-time Control of a Tokamak Plasma Using Neural Networks

Christopher M. Bishop[‡], Paul S. Haynes[†], Mike E. U. Smith[†],
Tom N. Todd[†], and David L. Trotman[†]

[‡] Current address: Microsoft Research,
7 J J Thomson Avenue, Cambridge, CB23 0FB, U.K.
cmbishop@microsoft.com
<http://research.microsoft.com/~cmbishop>

[†] AEA Technology, Culham Laboratory,
Oxfordshire OX14 3DB
(Euratom/UKAEA Fusion Association)

Published in *Neural Computation* **7**, No. 1 (1995) 206–217.

Abstract

In this paper we present results from the first use of neural networks for real-time control of the high temperature plasma in a tokamak fusion experiment. The tokamak is currently the principal experimental device for research into the magnetic confinement approach to controlled fusion. In an effort to improve the energy confinement properties of the high temperature plasma inside tokamaks, recent experiments have focussed on the use of non-circular cross-sectional plasma shapes. However, the accurate generation of such plasmas represents a demanding problem involving simultaneous control of several parameters on a timescale as short as a few tens of microseconds. Application of neural networks to this problem requires fast hardware, for which we have developed a fully parallel custom implementation of a multilayer perceptron, based on a hybrid of digital and analogue techniques.

1 Introduction

Fusion of the nuclei of hydrogen provides the energy source which powers the sun. It also offers the possibility of a practically limitless terrestrial source of energy. However, the harnessing of this power has proved to be a highly challenging problem. One of the most promising approaches is based on magnetic confinement of a high temperature ($10^7 - 10^8$ Kelvin) plasma in a device called a tokamak (from the Russian for ‘toroidal magnetic chamber’) as illustrated schematically in Figure 1. At these temperatures the highly

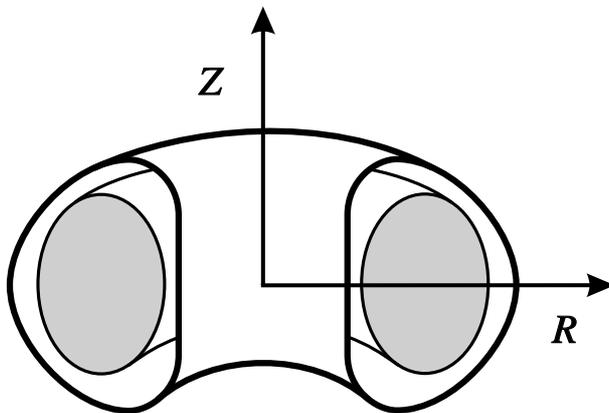


Figure 1: Schematic cross-section of a tokamak experiment showing the toroidal vacuum vessel (outer D-shaped curve) and plasma (shown shaded). Also shown are the radial (R) and vertical (Z) coordinates. To a good approximation, the tokamak can be regarded as axisymmetric about the Z -axis, and so the plasma boundary can be described by its cross-sectional shape at one particular toroidal location.

ionized plasma is an excellent electrical conductor, and can be confined and shaped by strong magnetic fields. Early tokamaks had plasmas with circular cross-sections, for which feedback control of the plasma position and shape is relatively straightforward. However, recent tokamaks, such as the COMPASS experiment at Culham Laboratory, as well as most next-generation tokamaks, are designed to produce plasmas whose cross-sections are strongly non-circular. Figure 2 illustrates some of the plasma shapes which COMPASS is designed to explore. These novel cross-sections provide substantially improved energy confinement properties and thereby significantly enhance the performance of the tokamak.

Unlike circular cross-section plasmas, highly non-circular shapes are more difficult to produce and to control accurately, since currents through several control coils must be adjusted simultaneously. Furthermore, during a typical plasma pulse, the shape must evolve, usually from some initial near-circular shape. Due to uncertainties in the current and pressure distributions within the plasma, the desired accuracy for plasma control can only be achieved by making real-time measurements of the position and shape of the boundary, and using error feedback to adjust the currents in the control coils.

The physics of the plasma equilibrium is determined by force balance between the thermal pressure of the plasma and the pressure of the magnetic field, and is relatively well understood. Particular plasma configurations are described in terms of solutions of the Grad-Shafranov equation (Shafranov, 1958), given by

$$R \frac{\partial}{\partial R} \left(\frac{1}{R} \frac{\partial \Psi}{\partial R} \right) + \frac{\partial^2 \Psi}{\partial Z^2} = -R J(\Psi, R) \quad (1)$$

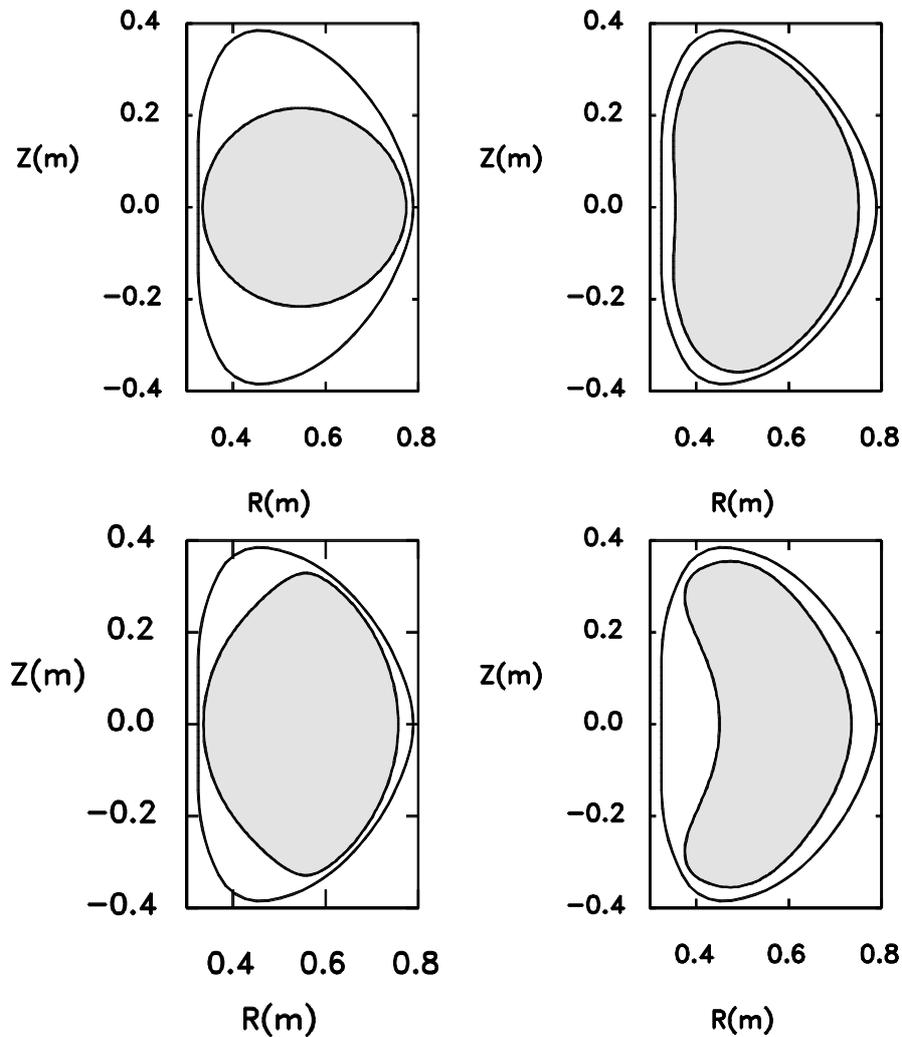


Figure 2: Cross-sections of the COMPASS vacuum vessel showing some examples of potential plasma shapes. The solid curve is the boundary of the vacuum vessel, and the plasma is shown by the shaded regions. Again, R and Z are the radial and vertical coordinates respectively in units of meters.

where the coordinates R and Z are defined in Figure 1, the function Ψ is called the poloidal flux function, and the plasma boundary corresponds to a surface of constant Ψ . The function $J(\Psi, R)$ specifies the plasma current density, and for the work reported here we have chosen the following representation

$$J(\Psi, R) = b \left\{ \beta R + \frac{(1 - \beta)}{R} \right\} (1 - \Psi^{\alpha_1})^{\alpha_2} \quad (2)$$

which is motivated by plasma physics considerations. Here b is a constant, β controls the ratio of plasma pressure to magnetic field energy density, and the parameters α_1 and α_2 are numbers ≥ 1 which can be varied to generate a variety of current profiles. Fortunately, the plasma configurations obtained by solution of the Grad-Shafranov equation are relatively insensitive to the precise choice of representation for the function $J(\Psi, R)$.

Due to the non-linear nature of the Grad-Shafranov equation, a general analytic solution is not possible. However, for a given current density function $J(\Psi, R)$, the Grad-

Shafranov equation can be solved by iterative numerical methods, with boundary conditions determined by currents flowing in the external control coils which surround the vacuum vessel. On the tokamak itself it is changes in these currents which are used to alter the position and cross-sectional shape of the plasma. Numerical solution of the Grad-Shafranov equation represents the standard technique for post-shot analysis of the plasma, and is also the method used to generate the training dataset for the neural network, as described in the next section. However, this approach is computationally very intensive and is therefore unsuitable for feedback control purposes.

For real-time control it is necessary to have a fast (typically $\leq 50\mu\text{sec.}$) determination of the plasma boundary shape. This information can be extracted directly from a variety of diagnostic systems, the most important being local magnetic measurements taken at a number of points around the perimeter of the vacuum vessel. Most tokamaks have several tens or hundreds of small pick up coils located at carefully optimized points around the torus for this purpose. We shall represent these magnetic signals collectively as a vector \mathbf{m} .

The position and shape of the plasma boundary can be described in terms of a set of geometrical parameters such as vertical position and elongation, which we collectively denote by y_k . These parameters are illustrated in Figure fig:geometry, and will be discussed in more detail in the next section. The basic problem which has to be addressed, there-

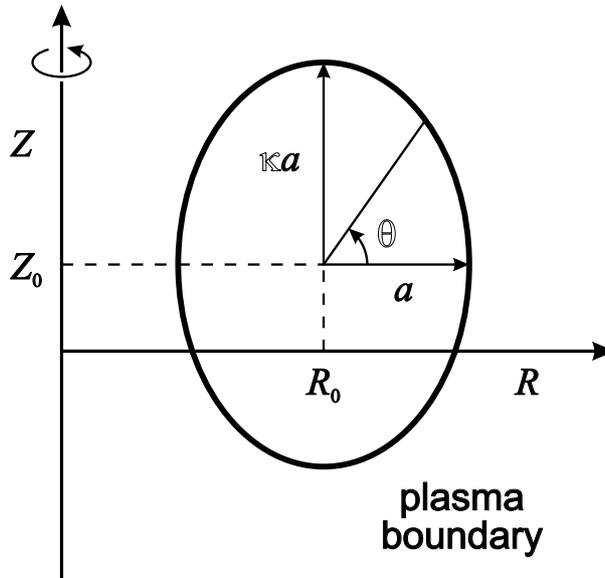


Figure 3: Schematic illustration of a cross-section of the toroidal vacuum vessel showing the definitions of various coordinates and parameters. The elliptical curve denotes the plasma boundary, whose center is at $R = R_0$, $Z = Z_0$. The parameter κ describes the elongation of the plasma, and θ is called the poloidal angle. The triangularity δ (not shown) describes the departure of the plasma boundary from a simple ellipse. (Values of $\kappa = 1$ and $\delta = 0$ correspond to a circular plasma boundary).

fore, is to find a representation for the (non-linear) mapping from the magnetic signals \mathbf{m} to the values of the geometrical parameters y_k , which can be implemented in suitable hardware for real-time control.

The conventional approach presently in use on many tokamaks involves approximating the mapping between measured the magnetic signals and the geometrical parameters by a single linear transformation. However, the intrinsic non-linearity of the mappings suggests

that a representation in terms of feedforward neural networks should give significantly improved results (Lister and Schnurrenberger, 1991; Bishop et al., 1992; Lagin et al., 1993). Figure 4 shows a block diagram of the control loop for the neural network approach to tokamak equilibrium control.

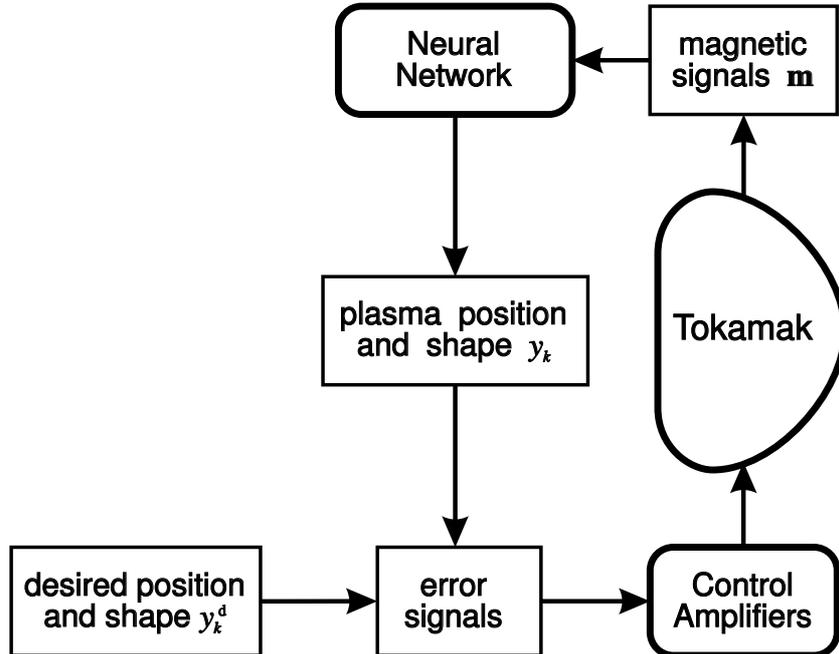


Figure 4: Block diagram of the control loop used for real-time feedback control of plasma position and shape. The neural network provides a fast non-linear mapping from the measured magnetic signals onto the values of a set of geometrical parameters y_k (illustrated in Figure 3) which describe the position and shape of the plasma boundary. These parameters are compared with their desired values y_k^d , and the resulting error signals used to correct the currents in a set of feedback control coils using standard linear PD (proportional-differential) controllers.

2 Software Simulation Results

The dataset for training and testing the network was generated by numerical solution of equation 1 using a free-boundary equilibrium code. This code contains a detailed description of the COMPASS hardware configuration, and allows the boundary conditions to be expressed directly in terms of currents in the control coils. The data base currently consists of over 2,000 equilibria spanning the wide range of plasma positions and shapes available in COMPASS. Each configuration takes several minutes to generate on a fast workstation.

For a large class of equilibria, the plasma boundary can be reasonably well represented in terms of a simple parameterization, governed by an angle-like variable θ , given by

$$\begin{aligned} R(\theta) &= R_0 + a \cos(\theta + \delta \sin \theta) \\ Z(\theta) &= Z_0 + a\kappa \sin \theta \end{aligned} \quad (3)$$

where we have defined the following parameters

| | |
|----------|--|
| R_0 | radial distance of the plasma center from the major axis of the torus, |
| Z_0 | vertical distance of the plasma center from the torus midplane, |
| a | minor radius measured in the plane $Z = Z_0$, |
| κ | elongation, |
| δ | triangularity. |

Thus, for instance, if the triangularity parameter δ is zero, the boundary is described by an ellipse with elongation κ . These parameters (except for the triangularity) are illustrated in Figure 3. Each of the entries in the database has been fitted using the form in equation 3, so that the equilibria are labelled with the appropriate values of the shape parameters.

On the COMPASS experiment, there are some 120 magnetic signals which could be used to provide inputs to the network. Since each input could either be included or excluded, there are potentially $2^{120} \simeq 10^{36}$ possible sets of inputs which might be considered. In order to find a computationally tractable procedure for selection of a suitable subset of inputs, we have used forward sequential selection (Fukunaga, 1990), based on a simple linear mapping (discussed shortly) to provide a selection criterion. Simulations aimed at finding a network suitable for use in real-time control have so far concentrated on 16 inputs, since this is the number available from the initial hardware configuration.

It is important to note that the transformation from magnetic signals to flux surface parameters involves an exact linear invariance. This follows from the fact that, if all of the currents are scaled by a constant factor, then the magnetic fields will be scaled by this factor, and the geometry of the plasma boundary will be unchanged. It is important to take advantage of this prior knowledge and to build it into the network structure, rather than force the network to learn it by example. We therefore normalize the vector \mathbf{m} of input signals to the network by dividing by a quantity proportional to the total plasma current. A scaling of the magnetic signals by a common factor then leaves the network inputs (and hence the network outputs) unchanged. Compared with learning by example, this explicit use of prior knowledge brings three distinct advantages: (i) the network exhibits exact, rather than approximate, invariance to rescaling of the currents; (ii) the relative output accuracy can be maintained over a wide range of plasma current (which typically varies from a few kA to a few 100kA during the plasma pulse); (iii) the network training can be performed with a smaller dataset than would otherwise be possible, which can be generated for just one value of total plasma current. Note that the normalization has to be incorporated into the hardware implementation of the network, as will be discussed in Section 3.

The results presented in this paper are based on a multilayer perceptron architecture having a single layer of hidden units with ‘tanh’ activation functions, and linear output units. Networks are trained by minimization of a sum-of-squares error using a standard conjugate gradients optimization algorithm, and the number of hidden units is optimized by measuring performance with respect to an independent test set. Results from the neural network mapping are compared with those from the optimal linear mapping, that is the single linear transformation which minimizes the same sum-of-squares error as is used in the neural network training algorithm, as this represents the method currently used on a number of present day tokamaks. This minimization can be expressed in terms of a set of linear equations whose solution can be found efficiently and robustly using the technique of singular value decomposition (Press et al., 1992). Note that the same normalization of the inputs was used here as in the neural network case.

Initial results were obtained on networks having 3 output units, corresponding to the values of vertical position Z_0 , major radius R_0 , and elongation κ ; these being parameters

which are of interest for real-time feedback control. The smallest normalized test set error of 11.7 is obtained from the network having 16 hidden units. By comparison, the optimal linear mapping gave a normalized test set error of 18.3. This represents a reduction in error of about 30% in going from the linear mapping to the neural network. Such an improvement, in the context of this application, is very significant.

For the experiments on real-time feedback control described in Section 4 the currently available hardware only permitted networks having 4 hidden units, and so we consider the results from this network in more detail. Figure 5 shows plots of the network predictions for various parameters versus the corresponding values from the test set portion of the database. Analogous plots for the optimal linear map predictions versus the database val-

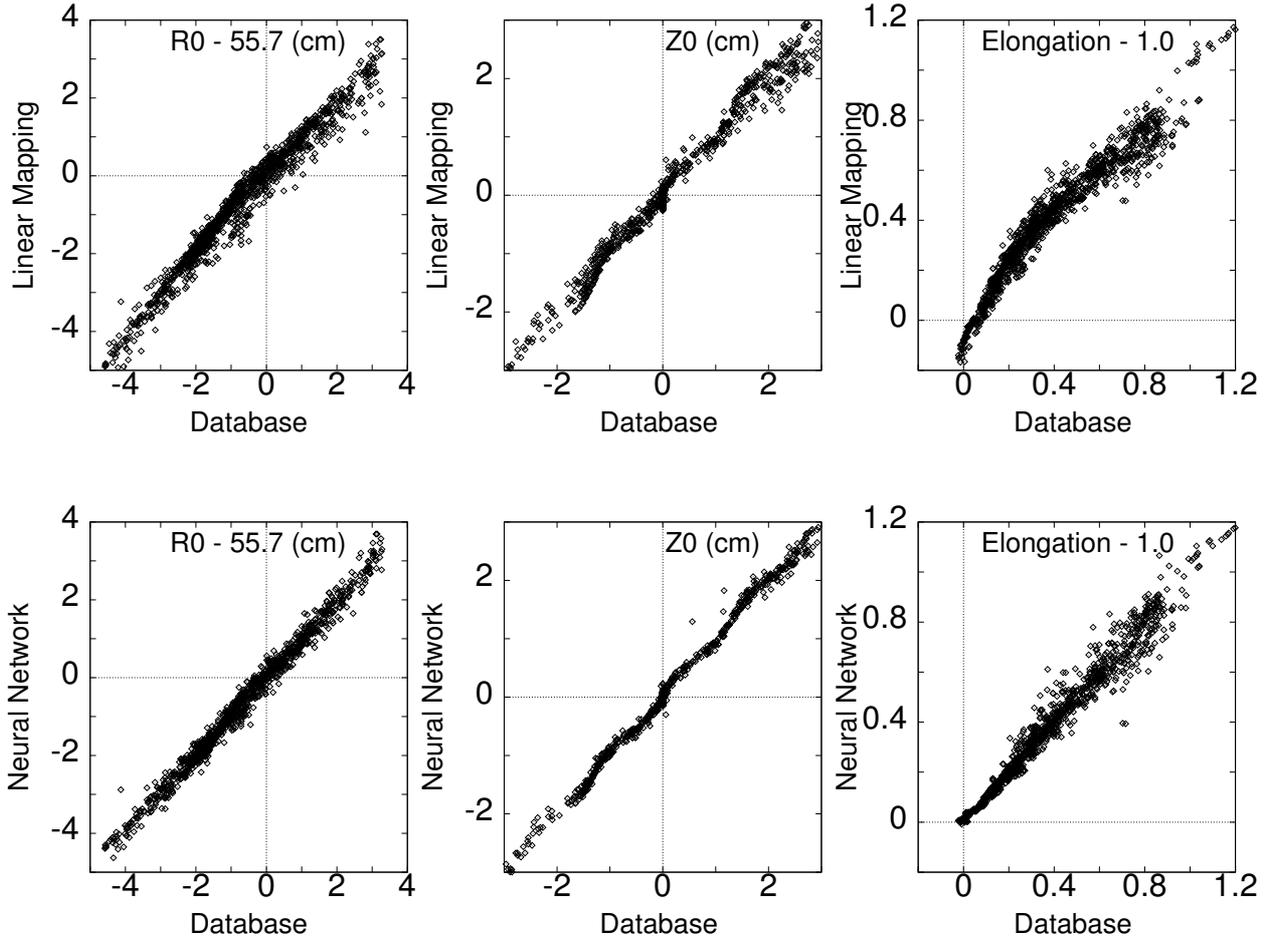


Figure 5: Plots of the values from the test set versus the values predicted by the linear mapping for the 3 equilibrium parameters, together with the corresponding plots for neural network with 4 hidden units.

ues are also shown. Comparison of the corresponding figures shows the poorer predictive capability of the linear approach, even for this sub-optimal network topology.

3 Hardware Implementation

The hardware implementation of the neural network must have a bandwidth of ≥ 20 kHz in order to cope with the fast timescales of the plasma evolution. It must also have

an output precision of at least 8 bits in order to ensure that the final accuracy which is attainable will not be limited by the hardware system. We have chosen to develop a fully parallel custom implementation of the multilayer perceptron, based on analogue signal paths with digitally stored synaptic weights (Bishop et al., 1993a). A VME-based modular construction has been chosen as this allows flexibility in changing the network architecture, ease of loading network weights, and simplicity of data acquisition. Three separate types of card have been developed as follows:

- Combined 16-input buffer and signal normalizer

This provides an analogue hardware implementation of the input normalization described earlier. For future flexibility this makes use of an EPROM (erasable programmable read-only memory) to provide independent scaling of groups of 8 inputs by an arbitrary function of an external reference signal. In the present application the reference signal is taken to be the plasma current (determined by a magnetic pick-up loop called a Rogowski coil) and the function is chosen to be a simple inverse proportionality.

- 16×4 matrix multiplier

The synaptic weights are produced using 12 bit frequency-compensated multiplying DACs (digital to analogue converters) which can be configured to allow 4-quadrant multiplication of analogue signals by a digitally stored number. The weights are obtained as a 12-bit 2's-complement representation from the VME backplane. Note that the DACs are being used here as digitally controlled attenuators, and not in their usual role of converting digital signals into analogue signals. Synaptic weights are downloaded (prior to the plasma pulse) via the VME backplane from a central control computer, using an addressing technique to label the individual weights.

- 4-channel sigmoid module

There are many ways to produce a sigmoidal non-linearity, and we have opted for a solution using two transistors configured as a long-tailed-pair, to generate a 'tanh' sigmoidal transfer characteristic. The principal drawback of such an approach is the strong temperature sensitivity due to the appearance of temperature in the denominator of the exponential transistor transfer characteristic. An elegant solution to this problem has been found by exploiting a chip containing 5 transistors in close thermal contact. Two of the transistors form the long-tailed pair, one of the transistors is used as a heat source, and the remaining two transistors are used to measure temperature. External circuitry provides active thermal feedback control, and stability to changes in ambient temperature over the range 0°C to 50°C is found to be well within the acceptable range. A separate 12-bit DAC system, identical to the ones used on the matrix multiplier cards but with a fixed DC input, is used to provide a bias for each sigmoid.

The complete network is constructed by mounting the appropriate combination of cards in a VME rack and configuring the network topology using front panel interconnections. The system includes extensive diagnostics, allowing voltages at all key points within the network to be monitored as a function of time via a series of multiplexed output channels.

4 Results from Real-time Feedback Control

Figure 6 shows the first results obtained from real-time control of the plasma in the COMPASS tokamak using neural networks. The evolution of the plasma elongation,

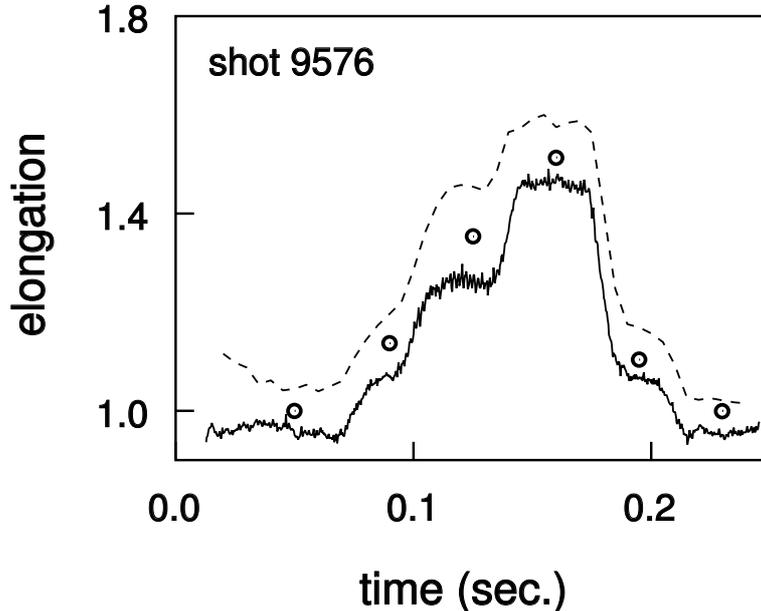


Figure 6: Plot of the plasma elongation κ as a function of time during shot no. 9576 on the COMPASS tokamak, during which the elongation was being controlled in real-time by the neural network. The solid curve shows the value of elongation given by one of the network outputs. The dashed curve shows the post-shot reconstruction of the elongation obtained from a simple ‘filament’ code, which gives relatively rapid post-shot plasma shape reconstruction but with limited accuracy. The circles denote reconstructions obtained from the full equilibrium code, which gives closer agreement with the network predictions.

under the control of the neural network, is plotted as a function of time during a plasma pulse. Here the desired elongation has been preprogrammed to follow a series of steps as a function of time. The remaining 2 network outputs (radial position R_0 and vertical position Z_0) were digitized for post-shot diagnosis, but were not used for real-time control. The graph clearly shows the network responding and generating the required elongation signal in close agreement with the reconstructed values. The typical residual error is of order 0.07 on elongation values up to around 1.5. Part of this error is attributable to residual offset in the integrators used to extract magnetic field information from the pick-up coils, and this is currently being corrected through modifications to the integrator design. An additional contribution to the error arises from the restricted number of hidden units available with the initial hardware configuration. While these results represent the first obtained using closed loop control, it is clear from earlier software modelling of larger network architectures (such as 32–16–4) that residual errors of order a few % should be attainable. The implementation of such larger networks is being pursued, following the successes with the smaller system.

Neural networks have already been used with great success for fast interpretation of the data from tokamak plasma diagnostics to determine the spatial and temporal profiles of quantities such as temperature and density (Bishop et al., 1993b; Bishop et al., 1993c; Bartlett and Bishop, 1993). There is currently considerable interest in extending

these techniques to allow real-time feedback control of the profiles to give more complete determination of the plasma configuration than is possible by boundary shape control alone. For such applications, neural networks appear to offer one of the most promising approaches.

Acknowledgements

We would like to thank Peter Cox, Jo Lister and Colin Roach for many useful discussions and technical contributions. This work was partially supported by the UK Department of Trade and Industry.

References

- Bishop C M, Cox P, Haynes P S, Roach C M, Smith M E U, Todd T N and Trotman D L, 1992. A neural network approach to tokamak equilibrium control. In *Neural Network Applications*, Ed. J G Taylor, Springer Verlag, 114–128.
- Bishop C M, Haynes P S, Roach C M, Smith M E U, Todd T N, and Trotman D L. 1993a. Hardware implementation of a neural network for plasma position control in COMPASS-D. In *Proceedings of the 17th. Symposium on Fusion Technology*, Rome, Italy. **2** 997–1001.
- Bishop C M, Roach C M and von Hellerman M. 1993b. Automatic analysis of JET charge exchange spectra using neural networks. *Plasma Physics and Controlled Fusion* **35** 765–773.
- Bishop C M, Strachan I G D, O'Rourke J, Maddison G and Thomas P R. 1993c. Reconstruction of tokamak density profiles using feedforward networks. *Neural Computing and Applications* **1** 4 – 16.
- Bartlett D V, and Bishop C M. 1993. Development of neural network techniques for the analysis of JET ECE data. In *Proceedings of the 8th. International Workshop on ECE and ECRH (EC8, 1992)*.
- Fukunaga K. 1990. *Statistical Pattern Recognition*, 2nd. Edition, Academic Press, San Diego.
- Lagin L, Bell R, Davis S, Eck T, Jardin S, Kessel C, Mcenerney J, Okabayashi M, Popyack J and Sauthoff N. 1993. Application of neural networks for real-time calculations of plasma equilibrium parameters for PBX-M, In *Proceedings of the 17th. Symposium on Fusion Technology*, Rome, Italy. **2** 1057–1061.
- Lister J B and Schnurrenberger H. 1991. Fast non-linear extraction of plasma parameters using a neural network mapping. *Nuclear Fusion*. **31**, 1291–1300.
- Press W H, Flannery B P, Teukolsky S A and Vetterling W T. 1992. *Numerical Recipes in C: The Art of Scientific Computing* 2nd. Edition, Cambridge University Press.
- Shafranov V D. 1958. On magnetohydrodynamical equilibrium configurations. *Sov. Phys. JETP*, **8** 710.