Journal of Intelligent Manufacturing (1992) **3**, 229–235

Neural network methods for the modeling and control of welding processes

GABOR KARSAI, KRISTINN ANDERSEN, GEORGE E. COOK and R. JOEL BARNETT

Department of Electrical Engineering, Vanderbilt University, PO Box 1824, Station B, Nashville, TN 37235, USA

Received June 1990 and accepted August 1991

While welding processes are of great importance in manufacturing, their modeling and control is still subject of research. The highly nonlinear, strongly coupled, and multivariable nature of these processes renders the use of analytical tools practically impossible. In this article a novel approach is presented which employs networks of simple nonlinear units: a neural network. A widely used welding process, the Gas Tungsten Arc Welding is presented and the problem of its modeling and control is exhibited. A very brief introduction to neural networks is followed by presenting the experimental results for modeling the static and dynamic behavior of the process, as well as some practical recommendations regarding the use of the neural network techniques for controlling these processes.

Keywords: neural networks, arc welding, modeling and control, nonlinear systems

1. Introduction

Welding is the most frequently used method in manufacturing for joining metal assemblies. The quality of a weld is of extreme importance, especially in the case of high-risk systems, like nuclear reactors or space systems. To ensure this quality one has to maintain a high level of control of the welding process itself.

Control for welding has traditionally referred to the control of individual components of the welding system. The control tasks are typically accomplished using linear feedback controllers of the single-input/single-output variety. Examples include the control of torch travel speed, control of the welding voltage or current, and other instances where a single process variable is forced to track a constant or varying setpoint signal. A relatively limited amount of work has been carried out in studying and implementing schemes for controlling more global aspects of the arc welding processes.

In general, any arc welding process can be viewed as a multiple-input/multiple-output system. The input variables to the process are generally those parameters which can be controlled by the welder, such as the various equipment parameters (arc current, voltage, torch travel speed, etc.). These parameters affect the finished weld to a varying degree, and using the notation of Cook (1980), they will be referred to as Indirect Weld Parameters (IWPs). At the output end of the process another set of parameters is defined to characterize the end result, and therefore the success, of the welding. Examples of such parameters include descriptors of the weld bead shape, its width and depth, the various strength measures of the weld, and the presence or absence of defects. These parameters, which characterize the finished weld, are referred to as Direct Weld Parameters (DWPs). As a whole, the process can therefore be described as a system which transforms any set of applied IWPs into a corresponding set of DWPs. Generally the multivariable system which constitutes an arc welding process is not well suited to traditional linear control for the following reasons:

(1) The system parameters are tightly coupled, i.e. each of the output parameters is strongly dependent on a number of the input parameters, and each input parameter affects a number of the output parameters as well.

(2) The entire system is usually not well known or defined in terms of the mathematical formalization necessary for proper controller design.

(3) The process is nonlinear, i.e. each of the output parameters is nonlinearly related to one or more of the input parameters.

These characteristics, which are common to all arc welding processes, complicate the design of the overall multivariable arc welding controller.

Traditionally, the individual parameters of the arc welding processes have been selected by human operators. Successful welding operation has relied on the skills and experience of the welder, combined with trial-and-error approaches as necessary. By aiding or replacing the human welding experts with a multivariable controller which selects and maintains the indirect weld parameters at the values required to obtain a given set of direct weld parameters, the overall efficiency and reliability of welding can be improved. The approach outlined in this paper uses artificial neural networks to select and control the parameters of an arc welding process. It is used in a continuing project carried out for the NASA Marshall Space Flight Center, aimed at improving the Gas Tungsten Arc Welding process (Andersen *et al.*, 1991). This welding process is used in a variety of applications, including welding on the NASA Space Shuttle Main Engine (SSME). The artificial neural network approach has been compared with more conventional methods, and it has been demonstrated to be reliable and well suited for modeling and controlling arc welding, as discussed later in this paper.

2. The gas tungsten arc welding process

Arc welding is generally categorized into a number of processes, each of which is based on unique techniques and therefore uses its own type of equipment and materials. Gas tungsten arc welding (GTAW) is one of these processes, frequently used for precision welding. Unlike some other arc welding processes the electrode in this process is not consumable, i.e. it does not melt and add material to the welded joint. The arc is maintained between the electrode tip and the welded joint (refer to Fig. 1). An inert shielding gas, such as argon or helium, is routed around the electrode so that it covers the arc and the molten weld pool. This gas has two purposes:

(1) It provides atoms to be ionized and thus form the arc plasma which carries the arc current.

(2) It shields the weld pool from the oxygen in the atmosphere, and thus it prevents undesirable oxidation of the welded metal.

The electric power source is operated as a current source rather than voltage source, and the welding current is typically on the order of 50–300 A. The arc voltage is therefore determined by the electrical impedance of the arc, which is primarily affected by the length of the arc. The length of the GTAW arc (defined as the distance between the electrode tip and the surface of the weld pool) is usually on the order of 0.030-0.100 in, with the corresponding arc voltage roughly in the range of 8–12 V, respectively, when argon is used for shielding. As a result, the arc voltage is usually set and maintained constant through a servo control mechanism (automatic voltage control – AVC) which adjusts the torch height and the arc length accordingly. Frequently, external metal is added to the molten spot under the arc through the feeding of filler wire. The torch holding the electrode, the shielding gas cup, and the wire feed mechanism if one is used, are moved along the welded joint at a controlled travel speed, typically on the order of 5–20 in/min.

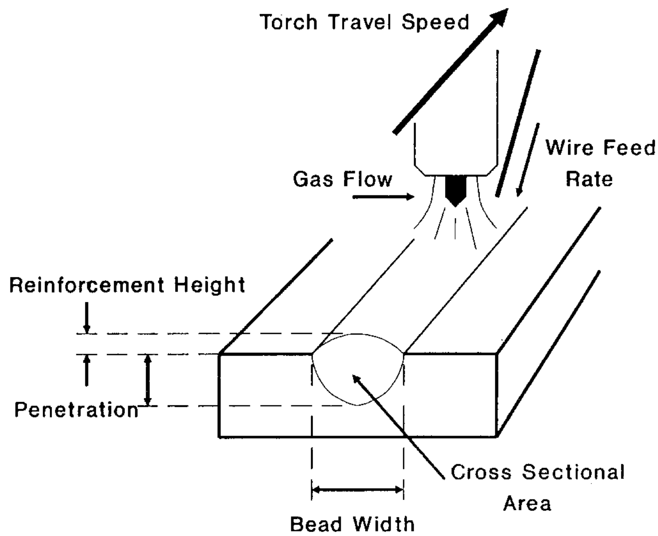

Fig. 1. The gas tungsten arc welding process.

A given joint, depending on the thickness of the welded materials, may only require one pass to complete the welding, or it may require a number of passes back and forth along the joint until the pieces are adequately welded together. In the case of a multiple-pass weld the welded joint is usually prepared as a groove, and filler material is used to fill the groove, layer by layer. During the first pass it is usually desirable to maintain the pool depth just about adequate to melt through the bottom of the joint, in which case surface tensions suffice to prevent the pool from falling through the joint. This is usually referred to as full penetration. Full penetration may be used with or without backing material, which serves to hold the molten pool in the joint and prevent it from falling through. Partial penetration, on the other hand, is obtained when the pool depth is less than the thickness of the welded material.

A number of control variables, or IWPs, are available to control the direct weld parameters of the GTAW process. These include the arc current (which may be constant, pulsed, or varied in other ways), arc voltage, torch travel speed, and wire feed rate. Additional parameters are static during welding but may be selected before the arc is initiated, such as electrode diameter and tip angle, feed wire diameter and composition, etc. The human welder usually selects the control parameters based on past experience, which in turn may be based on trial-and-error procedures. The arc current is the primary variable for controlling heat input, while the arc voltage controls the heat to a lesser extent. This is largely because the current can be varied over a relatively much larger range than the arc voltage.

It may appear that the number of available input variables gives the designer of the weld process controller a wide latitude in obtaining the desired set of direct weld parameters. In practice, however, the range of process input parameters is limited by the fact that many of these parameters need to be interrelated to maintain an acceptable weld. Welding current and travel speed, for example, are related in the sense that only a limited range of travel speeds is allowed for a given current level, assuming that other parameters of the process are fixed. Generally, the travel speed is increased as the current is increased, so as to maintain appropriate heat input per unit time to the weld. Either too much heat input or too little heat input is undesirable. Extremely significant variations are totally unacceptable. Therefore, the process controller must be aware of such limits and observe them as it selects the IWPs of the process.

The relations between weld bead geometry and weld equipment parameters have been studied extensively by a number of researchers. The physical phenomena governing some of these relationships are discussed in *The Physics of Welding* (Lancaster, 1986), where both nonlinearities and parameter coupling examples are illustrated in various contexts. Additional information can be found in Connor (1987).

3. Neural networks: mapping networks

Recent successes in employing artificial neural network (ANN) models for solving various computationally difficult problems have inspired revival of research in the area. Early work by McCullogh and Pitts (1943), and Widrow and Hoff (1960) focused largely on mathematical modeling while more recent research has augmented theoretical analysis with computer simulations and implementation demonstrations. Numerous variants of pattern classifiers using ANNs have been studied by Hopfield (1982), and Lippmann *et al.* (1988). Introductory texts to the subject may be found in Rumelhart *et al.* (1986), IEEE (1988), and Lippmann (1987).

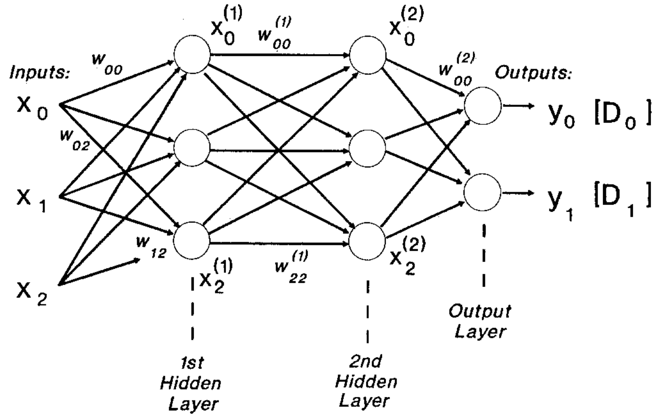

Fig. 2. The feedforward neural network.

The most frequently used neural network architecture is the multilayer mapping network, pictured in Fig. 2. It can be proven (Kolmogorov, 1957) that this network is capable of approximating any nonlinear and multivariable mappings. It achieves this capability through an incremental training process, which gradually adjusts the parameters, the weights of the network, to approximate the mapping in a least-square sense (Lippmann, 1987). The training process should present input/output pairs of examples, i.e. vectors of real numbers, of the desired mapping to the network (Lapedes and Farber, 1988). Having trained the system, entirely new input data can be presented to ANN, which, in turn, predicts new outputs based on the transfer characteristics learned during the training. If these new data are obtained from the same local region of operation of the process as during the training phase, data from the input/output relations should be governed by the same underlying process and the ANN should perform adequately.

Formally, the neural network model presented above is capable of learning arbitrary mappings of the form:

$$F: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

where n and m denote the dimensionality of the input

Table 1. Neural network performances for static models in terms of standard deviation.

Class	*Bead width (%)*	*Penetration (%)*	*Reinforcement height (%)*	*Cross-sectional area (%)*
Training data	2.86	7.75	7.13	6.72
New data	8.01	20.18	3.97	7.10

and output vectors. This feature can be utilized to facilitate the IWP-to-DWP mapping as follows. Experimental data can be considered as samples of the mapping and they can be used to train a network to approximate the mapping. The training algorithm reduces the mapping error at the points below a certain limit. Between the points, the network will interpolate for the continuity of the transfer function of the individual units. The quality of this interpolation can naturally be increased through finer sampling, i.e. more data.

4. Experimental results with mapping networks

To examine the performance of a neural network, numerical data from actual GTAW experiments were used. The objective was to train neural networks to map vectors of IWPs to the corresponding DWPs. A set of 31 distinct welds was used, each produced by a specific arc current, voltage, travel speed, and wire feed rate (IWPs). Analysis and cross-sectioning of each weld revealed the corresponding bead width, bead penetration, reinforcement height, and bead cross-sectional area (DWPs). The training of the neural network was carried out by teaching the network to approximate the measured IWP to DWP mapping. The procedure was repeated for each of the 31 training welds. This was continued until the combined error for the entire set of data reached a given threshold. Various network topologies were tried and one with two hidden layers and 18 nodes per layer was found to converge fastest without compromising overall accuracy. The stopping condition for the training process was set to 0.09 for the sum of squared DWP errors across the training data set. Once the neural network has been trained with the 31 welds, its performance on 11 randomly selected welds out of the training set was tested. An additional 11 welds, not from the training set, were given to the network and its performance in predicting the DWPs from the IWPs of this new set of data was evaluated for comparison. The results are shown in Table 1. Based on these results and other tests, we considered the trained neural network robust enough for practical purposes.

5. Neural networks: dynamic networks

The modeling of a dynamic system using adaptive techniques can be considered as the problem of teaching an adaptive system to follow arbitrary space–time trajectories. There have been many suggestions for teaching a neural network to do this. Pineda (1988) generalizes the backpropagation algorithm for networks containing recurrent feedback; thus the network becomes an adaptive *dynamic* system. Jordan (1988) extends a feedforward network with a feedback path, which goes from the output units to a group of input units, representing past states of the system; the structure was shown to learn sequences. Robinson and Fallside (1988) use a similar structure: the output of a feedforward network is fed back to its input through a delay line. Pearlmutter (1988) derives a gradient search algorithm for a general continuous recurrent network which can be used to teach the system to follow state space trajectories. Most of these systems have been devised for applications in cognitive psychology, and, according to our knowledge, have rarely been used for modeling physical systems.

In our case the following approach was used. Suppose that the general, nonlinear, multivariable dynamic system can be modeled according to the following (autoregressive) equation:

$$y(n) = F(\, x(n), \ldots, x(n-k+1), y(n-1), \ldots, y(n-l))$$

where $y(n)$ is the output vector of p variables of the system at time point n, $x(n)$ is the input vector of q variables to the system at time point n, and F is a nonlinear mapping:

$$F\colon \mathbf{R}^{k\times q+l\times p} \to \mathbf{R}^{p}$$

That is to say, by knowing

(1) the output of the system for the last l samples;
(2) the input to the system for the last $k-1$ samples; and
(3) the current input $x(n)$,

we can compute the output of the system $y(n)$ at time n.

The feedforward multilayer network presented above may be taught to approximate F, a nonlinear, multivariable mapping. The teaching can be done using the backpropagation learning algorithm. Our assumption is that if the modeling equation presented above is valid, and the network is able to approximate the mapping, a network equipped with appropriate tapped delay lines on its input can serve as a teachable model for a nonlinear multivariable dynamic system. The architecture is visualized in Fig. 3.

As in the case of the static mapping network, there remains a set of 'free parameters', which should be determined empirically. These include:

(1) number of layers and nodes in the network;
(2) 'length' of the tapped delay line for input variables;
(3) 'length' of the tapped delay line for output variables;
(4) the value of the learning constant.

In our work we determined these values by using experimental data from the GTAW process.

The architecture suggested here was intended for use in modeling mainly the transient behavior of a system. In

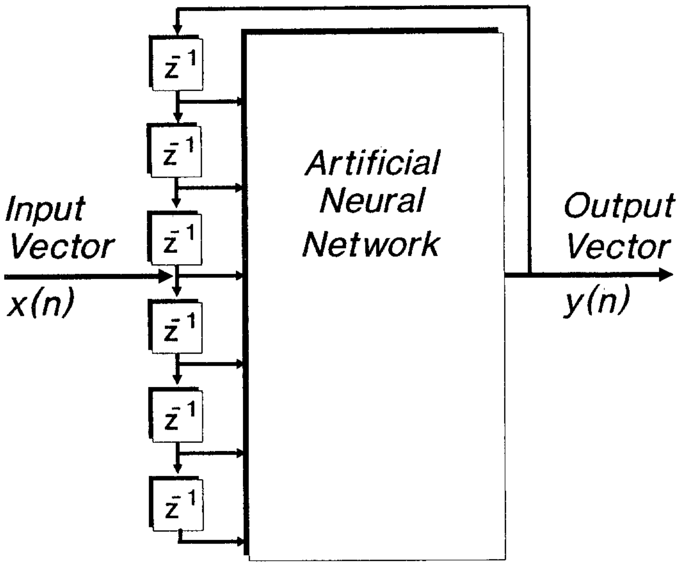

Fig. 3. Neural network model with delay lines.

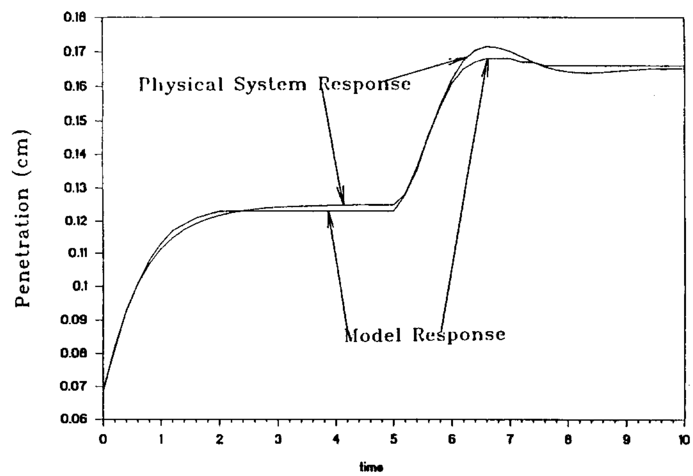

Fig. 4. Actual response versus model response to change in current.

welding it is typical that the system has many fixed points, because for example, the IWPs take their values from a finite set, based on the possible equipment settings. During the actual welding, IWPs are maintained constant most of the time; thus the IWP to DWP mapping is relatively fixed. However, there can be sudden changes in the IWPs, which take the system along a particular trajectory from one setpoint to another one. Our task was to teach the neural network to imitate these transitions.

6. Experiments with dynamic networks

For dynamic modeling, various experiments have been performed (Andersen *et al.*, 1988). Data were recorded from the physical process to determine how DWPs change in time as IWPs change. In one case, for example, the arc current was stepped from 100 A to 150 A followed by another step from 150 A to 200 A, and the effect of this on the penetration depth was measured. The measurements were made by stepping the current and observing how the weld geometry changed. Note that the torch travels along the weld, and the behavior of the process in the time-domain can be observed in spatial dimension. The physical process exhibited a first-order behavior for the first step and second-order behavior for the second step. These data were used to train the network. The result of the experiment can be seen in Fig. 4. The network had three layers, with four units in the hidden layer, two input delays and three output delays. Further experiments are currently being conducted to test the feasibility of the modeling technique.

7. Welding applications of artificial neural networks

Controllers for welding systems are currently designed as assemblies of subsystems (Cook, 1981). Each subsystem is concerned with the control of traditional equipment parameters, such as current and voltage. To obtain the desired results from the welding process, the human operator interactively adjusts these equipment parameters until he is satisfied with the weld. Artificial neural networks give the welding system designer the potential to approach the entire problem from a more global viewpoint. Instead of requiring the welder to rely on experience and experimentation to adjust the equipment parameters for the desired results, the system may apply neural networks to aid in selecting these parameters, or, ideally, to control them without any intervention from the welder. In that case, the welder would only prescribe the desired properties of the weld, e.g. bead dimensions, and the system would select and control the equipment parameters throughout the weld to obtain the desired results. The general modeling capabilities of artificial neural networks appear to provide the means for the design of such systems.

Having identified the arc welding process as a multivariable system, the questions of control methodology arises. All control methodologies require a formal description of the system to be controlled. In the case of arc welding such description is usually difficult to formulate accurately. This is mainly due to the limited knowledge of the physics of the arc welding process, and also because the process is frequently influenced by factors that are not taken into account or known by the control system designer.

Controlling a multivariable system such as the GTAW process is not a trivial task without an adequate model. In broad terms, weld process models are either derived from the physics of heat transfer or they are constructed from empirical data. The models derived from heat transfer physics frequently assume that the arc can be modeled as a heat source of a given form (a point source, a disk-shaped source of heat, etc.) and then the three-dimensional heat equation is applied to calculate the temperatures at various points in the workpiece. Because of the numerous assumptions and simplifications necessary to derive the analytical models, these models are usually fairly inaccurate. On the other hand they offer some insight into the mechanisms of the weld process and may illustrate qualitatively how some of the individual process inputs and outputs are related.

Contrasted against the physics-based models are the empirically derived models. These models may simply be a set of equations relating the process outputs to the process inputs, and derived by obtaining a best fit of experimental data to the equation form. In such extreme cases the models are derived without any consideration to the underlying physics of the process. A number of weld process models can be placed between the two extremes of purely physics-derived and purely empirical models. Frequently, physics-based models are derived using the necessary approximations and then various empirical constants and other unspecified variables are tuned until the model adequately agrees with experimental welding data.

There are various roles for models in a generic weld process controller. Firstly, a model can be used in defining the initial equipment parameters of the process. The welder specifies the desired DWPs, such as weld bead width, penetration, etc., and the model can be used to arrive at suitable IWPs, such as welding current, travel speed, etc. This is illustrated in Fig. 5, where an artificial neural network serves to select equipment parameters required to obtain the desired bead geometry.

Secondly, a model can be implemented in parallel with the actual process and provide calculations for DWPs that cannot be measured directly in real-time. Thus, a weld model can provide the controller at any time with an estimate of the weld bead penetration, although it may not be measurable in real-time. The real-time application of neural nets for welding control is shown in Fig. 6. A traditional proportional-plus-integral (PI) controller is used in this case, and it provides the controller

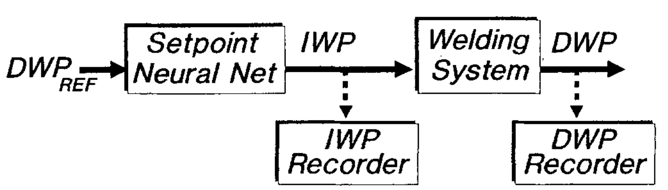

Fig. 5. An artificial neural network used as a weld parameters selector.

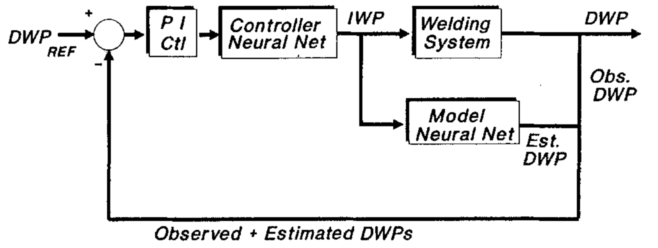

Fig. 6. Artificial neural networks used in a real-time weld controller.

neural network with scaled DWP values which are transformed to the appropriate indirect weld parameters. Note that if the controller network approximates the inverse of the plant, then in steady state the output of the PI equals the expected output of the plant. It should also be noted that the multivariable PI controller does not have to take cross-couplings into account, as decoupling of the IWPs from the DWPs takes place in the controller neural network.

8. Evaluation

When compared to other modeling methodologies ANNs have certain drawbacks as well as advantages. For the drawbacks the most notable one is the lack of direct relation to the physics of the process. Relating the qualitative effects of the network structure or parameters to the process parameters is usually impossible. On the other hand most physical models resort to substantial simplifications of the process and therefore trade accuracy for comprehensibility. The advantages of ANN-based models include relative accuracy and generality. If the training data for a neural network are general enough, spanning the entire range of process parameters, the resulting model will capture the complexities of the process, including nonlinearities and parameter cross-couplings, over the same ranges. Model development is much simpler than for most other models. Instead of theoretical analysis and development for a new model the neural network tailors itself to the training data. The network can be refined at any time with the addition of new training data. Finally, the neural network can calculate its results relatively quickly, as the input data are only propagated once through the network in the application mode.

9. Conclusions

Novel methods for the modeling and control of welding processes, based on artificial neural network techniques, have been presented. Due to their learning and selftuning capabilities, neural nets are excellent tools for modeling and controlling processes which are not readily analyzed and quantified, and thus not well suited for traditional methods. Specific tests and simulations have shown neural networks to yield worst-case errors in weld parameter predictions on the order of 20% and typical errors are well within 10%.

Acknowledgement

This work has been partly funded by the National Aeronautics and Space Administration SBIR contract No. NAS8-37401. Gratitude is due to Mr Clyde S. Jones, III, and Dr Arthur C. Nunes, Jr, both of the Materials and Processes Laboratory at the NASA Marshall Space Flight Center, for their support of this work.

References

Andersen, K., Barnett, R. J., Cook, G. E., Ramaswamy, K. and Prasad, T. (1988) Intelligent gas tungsten arc welding control – phase I final report, National Aeronautics and Space Administration, Contract No. NAS8-37306.

Andersen, K., Barnett, R. J., Springfield, J. F. and Cook, G. E. (1991) Intelligent gas tungsten arc welding control – phase II final report, National Aeronautics and Space Administration, Contract No. NAS8-37401.

Connor, L. P. (1987) *Welding Handbook, Volume 1: Welding Technology*, American Welding Society, Miami, Florida.

Cook, G. E. (1980) Feedback control of process variables in arc welding, in *Proceedings of the Joint Automated Control Conference*, Vol. II, IEEE Publ. 80CH1580-0, New York, The Institute of Electrical and Electronic Engineers, Inc., pp. FA7-B-1-10.

Cook, G. E. (1981) Feedback and adaptive control in automated arc welding systems. *Metal Construction*, **13**(9), 551–556.

IEEE (1988) Special issue on neural networks in control, *IEEE Control Systems Magazine*, **8**.

Jordan, M. L. (1988) Supervised learning and systems with excess degrees of freedom, in *Proceedings of the 1988 Connectionist Models Summer School*, Morgan Kaufmann Publishers, CA, pp. 62–75.

Hopfield, J. J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, USA, **79**, 2554–2558.

Kolmogorov, A. N. (1957) On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Doklady Akademii Nauk SSSR*, **114**, 953–956.

Lancaster, J. F. (ed.) (1986) *The Physics of Welding*, International Institute of Welding, Pergamon Press, 2nd Edition.

Lapedes, A. and Farber, R. (1988) How neural nets work, in *Neural Information Processing Systems*, American Institute of Physics, New York, pp. 442–456.

Lippmann, R. P. (1987) An introduction to computing with neural nets, *IEEE ASSP Magazine*, pp. 4–22.

Lippmann, R. P., Gold, B. and Malpass, M. L. (1988) A comparison of Hamming and Hopfield neural nets for pattern classification, MIT Lincoln Laboratory Technical Report, TR769.

McCullogh, C. A. and Pitts, W. (1943) A logical calculus of the ideas imminent in nervous activity. *Bulletin of Mathematical Biophysics*, **5**, 115–133.

Pearlmutter, B. (1988) Learning state space trajectories in recurrent neural networks, in *Proceedings of the 1988 Connectionist Models Summer School*, Morgan Kaufmann Publishers, CA, pp. 113–117.

Pineda, F. J. (1988) Generalization of backpropagation to recurrent and higher-order neural networks, in *Neural Information Processing Systems*, American Institute of Physics, New York, pp. 602–611.

Robinson, A. J. and Fallside, F. (1988) Static and dynamic error propagation networks with application to speech coding, in *Neural Information Processing Systems*, American Institute of Physics, New York, pp. 632–641.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986) *Parallel Distributed Processing*, **12**, MIT Press, Cambridge, MA.

Widrow, B. and Hoff, M. E. (1960) Adaptive switching circuits, *1960 IRE WESCON Conv. Record, Part 4*, pp. 96–104.