WELDING RESEARCH

SUPPLEMENT TO THE *WELDING JOURNAL*, FEBRUARY 2000
Sponsored by the American Welding Society and the Welding Research Council

Improved Ferrite Number Prediction in Stainless Steel Arc Welds Using Artificial Neural Networks — Part 1: Neural Network Development

The development of a new neural network model for predicting Ferrite Number more accurately than existing constitution diagrams is described

BY J. M. VITEK, Y. S. ISKANDER AND E. M. OBLOW

ABSTRACT. Neural network modeling is a powerful nonlinear regression analysis method that is extremely useful in identifying behavioral trends. This methodology was applied to the problem of predicting Ferrite Number in arc welds as a function of composition. This paper describes the details of the development of the neural network model, named FNN-1999, including the identification of the optimum network architecture and network parameters. The model was trained on the same data as the WRC-1992 constitution diagram and covers a range of Ferrite Numbers from 0 to 117, with a corresponding wide range in composition. Results of the model are presented in Part 2. It is shown that the accuracy of the FNN-1999 model in predicting Ferrite Number is superior to the accuracy of other models that are currently available, including the WRC-1992 diagram.

Introduction

The ability to predict the ferrite content in stainless steel arc welds is essential for many reasons. To a large extent, the final ferrite content determines a weldment's properties. In addition, ferrite content is a useful indicator of the mode of solidification and the cracking propensity of austenitic stainless steels, which is strongly influenced by whether the alloy solidifies in the primary austenitic or ferritic mode. Manufacturers as well as consumers often use Ferrite Number as an alloy specification to ensure that weldments contain a desired minimum (or maximum) ferrite level.

Over the years, various attempts have been made to try to accurately predict the ferrite content in stainless steel arc welds. Constitution diagrams have been developed that convert the alloy composition into two factors, a chromium equivalent (Cr_{eq}) and a nickel equivalent (Ni_{eq}). The former contains alloying elements that influence the microstructure in the same way as chromium, *i.e.*, they are ferrite stabilizers, while the latter contains elements that behave like nickel, *i.e.*, as austenite stabilizers. The constitution diagrams display ferrite levels as a function of these chromium and nickel equivalents. An early version of such a diagram that has been used for welding is the Schaeffler diagram (Ref. 1). Since the introduction of the Schaeffler diagram, several modifications and improvements have been proposed (Refs. 2–7). Corresponding constitution diagrams for stainless steel castings have also been proposed (Refs. 8, 9). The various versions of constitution diagrams differ primarily in the coefficients that are used to convert the alloy composition into the Cr_{eq} and Ni_{eq}. An extensive review is given in Ref. 5. The most recent version of the constitution diagram is the WRC-1992 diagram (Ref. 7). The following are the equations for the Cr_{eq} and Ni_{eq} factors in the WRC-1992 diagram:

$$Cr_{eq} = Cr + Mo + 0.7\ Nb \qquad (1a)$$

$$Ni_{eq} = Ni + 35\ C + 20\ N + 0.25\ Cu \qquad (1b)$$

where the elemental symbols represent the weight percent of each element.

Recently, another approach (Function Fit model) has been proposed for predicting Ferrite Number in stainless steel weldments (Ref. 10). In this method, the difference in free energy between ferrite and austenite was calculated as a function of composition, and this was related to the Ferrite Number. A regression analysis was used to determine the coefficients associating the Ferrite Number with the free energy change. It was found

KEY WORDS

Neural Network
Ferrite Number
Constitution Diagram
WRC-1992
Predicting Ferrite
Stainless Steel

J. M. VITEK and E. M. OBLOW are with Oak Ridge National Laboratory, Oak Ridge, Tenn. Y. S. ISKANDER is a graduate student at the Georgia Institute of Technology, Atlanta, Ga.

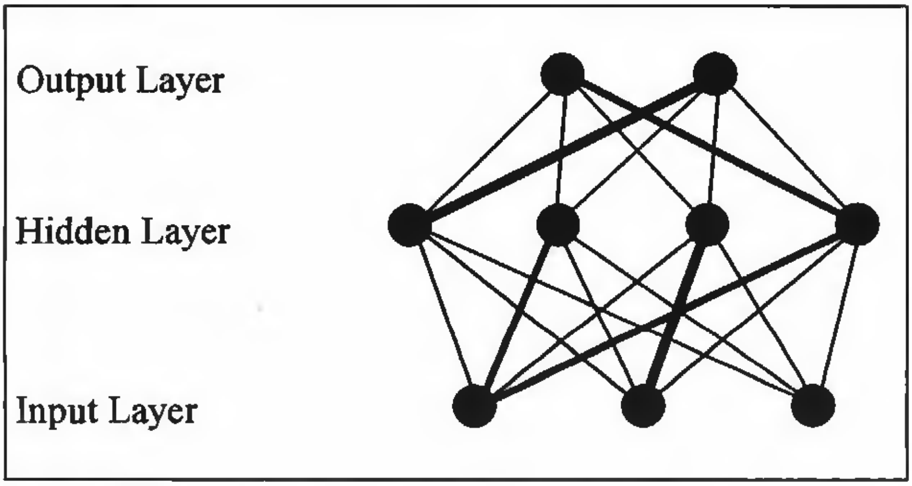

Fig. 1 — Schematic diagram showing the multiple-layer structure of a neural network and the interconnectivity between the nodes of the network.

that this approach was comparable in accuracy to that of the WRC-1992 diagram. The advantage of this new approach was that certain restrictions with regard to the alloying elements and their composition ranges were eased (Ref. 10).

The residual, room-temperature ferrite content in weldments is controlled by several factors and is the result of a series of microstructural transformations that take place during the welding process. Initially, the ferrite content is influenced by the solidification mode. This is determined by the alloy composition. Recently, many investigations have shown that the solidification conditions, and in particular the cooling rate, can also strongly influence the solidification mode. It was found that under rapid cooling conditions, alloys that normally solidify in the primary ferrite mode may, instead, solidify in the primary austenite mode (Refs. 11–17). Once solidified, the alloy undergoes a solid-state transformation of ferrite to austenite as the ferrite stability, relative to austenite, decreases with decreasing temperature, at least in the elevated temperature range of about 800 to 1100°C. The extent of this transformation depends on the nucleation and diffusion-controlled growth kinetics of austenite. As in the case of the solidification mode, this transformation is strongly influenced by several factors, including the cooling rate, the degree of segregation within the ferrite phase, and the amount of austenite already present after solidification. Thus, the relationship between overall alloy composition and final ferrite content can be expected to be quite complex. Simple relationships such as those in Equation 1 cannot be expected to take into account all of the critical factors. First, the relative influence of each alloying addition, given by that element's coefficient in the Cr_{eq} or Ni_{eq} expression, is likely to change over the full composition range. Furthermore, constitution diagrams that rely on simple linear expressions for the Cr_{eq} and Ni_{eq} factors ignore interactions between the elements. For example, the influence of manganese levels on ferrite content may vary depending on the chromium, nickel or carbon content (Refs. 18, 19). Since the coefficients are not composition dependant, such interactions are totally ignored. Some attempts at deriving composition-dependent parameters have been proposed (Refs. 2, S), but these are either valid for restricted composition ranges or they are not as accurate as the predictions of the WRC-1992 diagram.

Therefore, it is desirable and appropriate to describe the residual ferrite content in a more flexible manner so that elemental interactions can be factored into the relationships and, presumably, more accurate predictions would follow. Furthermore, it would be desirable to include process conditions as a variable since these often affect the final ferrite levels. Neural networks are ideally suited to improve the flexibility, robustness and accuracy of ferrite predictions because they make use of nonlinear regression methods. The identification of complex relationships between variables is straightforward with the use of neural networks, whereas they are extremely difficult or impossible to identify by standard regression analyses. A preliminary study by the authors (Ref. 20) developed a neural network model for ferrite prediction, and it was shown that the accuracy of the neural network model was better than conventional constitution diagrams. The present work describes a more advanced model that considers additional alloying elements. Part 1 of this paper describes in detail the methodology for the development of the neural network, including the identification of the optimum neural network architecture. In Part 2, the neural network results are described (Ref. 21). It is shown in Part 2 that the advanced neural network model (referred to as FNN-1999) is significantly more accurate than either the constitution diagrams or the earlier, simpler neural network model. Although process variables will not be considered in this study, it is assumed that a robust artificial neural network based on composition alone will provide a solid foundation for including process variables in the future.

Neural Network Theory

Neural networks are based on a simple scheme in which inputs are related to outputs by a system of interconnected nodes. A very elementary description of the concept behind neural networks is given below. There is extensive literature on the theory behind neural networks. The reader is referred to other publications for more details (Refs. 22, 23).

A network structure consists of layers of three types of nodes: input nodes, hidden nodes and output nodes. A schematic diagram of a simple neural network is shown in Fig. 1. The three layers are marked in the diagram. In principle, the hidden layer can consist of multiple sublayers, but in this study only one hidden layer was considered. The number of input and output nodes (three and two, respectively, in the schematic diagram of Fig. 1) represent the input and output variables. When applied to Ferrite Number prediction, the input nodes correspond to elemental concentrations while the output node represents the Ferrite Number. The number of hidden nodes is adjustable and is optimized for maximum accuracy, as described in the following pages. The nodes are "connected" to each other so that the value of one node will affect the value of another. The relative influence that a given node has on another one is specified by the "weight" that is assigned to each connection. The different weights for the node connections are shown schematically in Fig. 1 by different line thicknesses.

The network is based on the following sequence of steps:

1) Convert input values (V_i) to normalized inputs (I_i).

2) Sum all input contributions to each

hidden node (S_j).

3) Convert each sum to a hidden node value (H_j).

4) Sum all hidden node contributions to each output node (S_k).

5) Convert each sum to a normalized output value (O_k).

6) Convert normalized outputs to real output values (V_k).

The nomenclature for the variables uses the following conventions:

Real variables are denoted by V while neural network variables are given by I, H or O, for the input, hidden and output layers, respectively. The subscripts i, j and k refer to the individual nodes in the input, hidden and output layers, respectively. The sequence of steps outlined above is easily described mathematically. Details are provided in Appendix A.

The neural network analysis involves training of the network to achieve optimal accuracy. Training takes the form of adjusting the weights (connections) between the nodes in the different layers. Once the optimal neural network weights are determined, the process of calculating the outputs from a set of inputs can be expressed analytically and is readily converted into a spreadsheet format.

Neural network training is carried out with the use of a training data set that contains sets of inputs and corresponding outputs. The neural network is developed by comparing the predicted output values with the actual outputs and, using an optimization scheme, adjusting the weights to minimize the prediction error. Through the learning process, which involves thousands of iterations, a complicated set of empirical relationships between input and output variables is developed. Eventually, with minimal influence from the user, the network "learns" a scheme in which outputs are associated with the inputs. In the present analysis, a feed-forward network with a back-propagation optimization scheme was utilized (Ref. 23). A commercial software package (NeuralWorks Professional II/PLUS™) was used for the neural network development (Ref. 24).

Although terms corresponding to elemental interaction cross-terms in the chromium and nickel equivalent expressions (Equation 1) are not directly identifiable in the network (Fig. 1), the artificial neural network analysis allows for nonlinear relationships among the inputs and can describe such elemental interactions very well. Unlike standard regression analyses, the form of the elemental interaction is not specified in advance with the neural network approach.

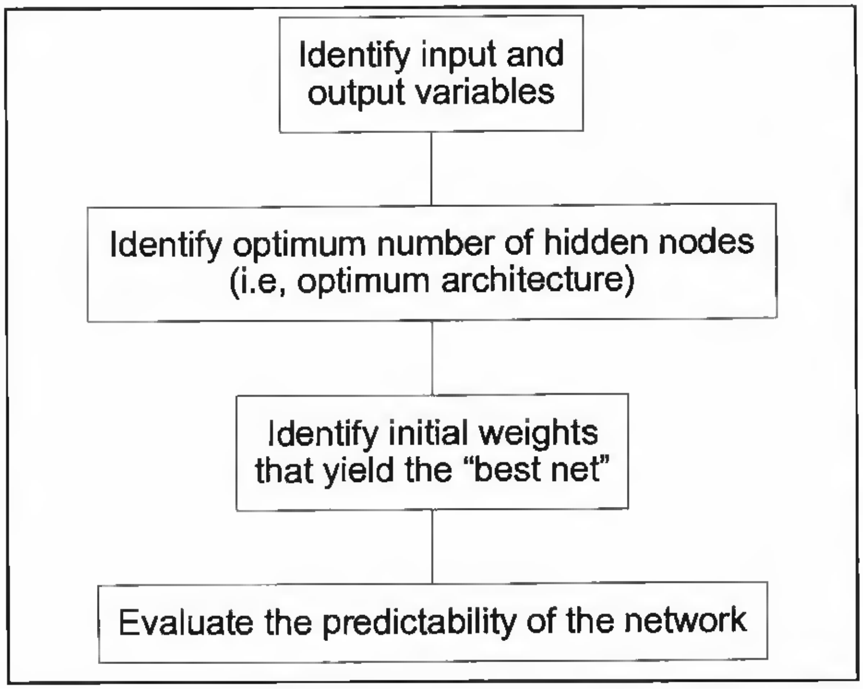

Fig. 2 — Flow chart showing the four basic steps to identify the optimum neural network architecture and the best neural network.

Experimental Ferrite Number Data

The same data that were used in the development of the WRC-1992 constitution diagram (Ref. 7) and the more recent Function Fit model (Ref. 10) were used in this study. The data consist of three compilations of composition and measured Ferrite Number (FN) with a total of 961 data points (Refs. 25–27). This data set is referred to in the text as the "complete training data set." In the course of the neural network development, this data set was subdivided into smaller training and testing data subsets, as described in the following pages. The composition and measured FN ranges for the data set are listed in Table 1.

As with the conventional constitution diagrams, Ferrite Number has been used as an indicator of the ferrite content. There are many advantages to using this quantity rather than a measurement of the actual volume percent of ferrite in the microstructure (Ref. 28). The data consisted of FN measurements and compositional analyses made by various laboratories on welds made by a variety of arc welding techniques. For this study, no attempt was made to include the welding process as an input variable. The effect of welding process on the ferrite content will be considered in a later investigation.

Since the experimental data came from several different sources, the composition analyses did not always include the same elements. Therefore, special consideration had to be given to the way in which missing composition data were handled. In the earlier neural network model (Ref. 20), it was decided to consider only those elements for which a

Table 1 — Ranges in Composition and FN for the Training and Supplemental Data Sets

	C	Cr	Ni	Mo	N	Mn	Fe	Si	Cu	Ti	Nb	V	Co	FN
							Complete Training Data Sets							
min.	0.008	14.74	4.61	0.01	0.01	0.35	45.60	0.03	0.0	0.0	0.0	0.0	0.0	0
max.	0.2	32	33.5	6.85	0.3	12.67	72.51	1.3	3.04	0.54	0.88	0.23	0.32	117
							Supplemental Data Set							
min.	0.013	16.25	8.4	0.0	0.018	0.1	56.42	0.21	0.0	0.0	0.0	0.0	0.0	0
max.	0.13	23.91	13	2.86	0.12	9.5	72.39	1.64	0.0	0.0	0.78	0.0	0.0	30.1

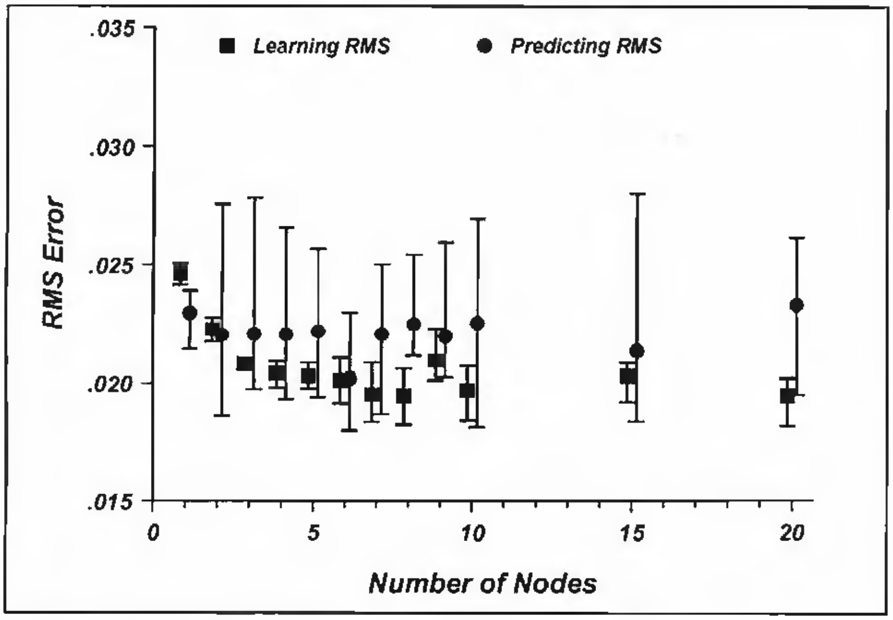

Fig. 3 — The variation of the learning RMS and the predicting RMS vs. the number of nodes in the hidden layer, based on average results for four train/test sets. The bars show the range of RMS values that were found. The data points are slightly offset horizontally to make them more easily visible.

chemical analysis was available for all the data. This limited the earlier neural network model to the consideration of only eight elements: C, Cr, Ni, Mo, N, Mn, Fe and Si, with eight corresponding input nodes. In the present analysis, five additional elements were considered: Cu, Ti, Nb, V and Co. As will be shown, the addition of these five elements led to a significant improvement in FN prediction accuracy. Exactly how the absence of chemical analyses for these five elements was handled by the network model is described in detail in Appendix B.

By using data from various sources, it is likely that some data were more reliable than others (Refs. 5, 27). However, no attempt was made to screen the data in advance, or to exclude data in any way. Naturally, the decision to use all of the data meant that considerable scatter in the training data was present. Often, data points with nearly identical compositions had different measured FN values. Some of this variability may be attributable to different welding processes that were used, but much of the scatter represents unavoidable experimental error and is due to variations from laboratory to laboratory in chemical analysis and in FN measurement. Whatever the source of the scatter, this scatter in the data limits the degree to which the neural network model can accurately predict Ferrite Number.

A second data set, referred to as the "supplemental data set," was used for testing of the neural network and was completely independent of any neural network training. This data set was used earlier to assess the accuracy of the WRC-1988 constitution diagram (Ref. 29). It consists of 265 data points and was compiled by H. Ornig (Ref. 30). The ranges of compositions and FN for this data set are also listed in Table 1. It can be seen that the element concentration ranges for this data set generally lie within the ranges of the training set. The only exceptions are for Mo, where the minimum is slightly lower in the supplemental set, Mn, where the lower limit is 0.1 compared to 0.35 in the training set, and Si, where the maximum is 1.64 compared to 1.3 in the training set. In theory, the neural network should be applied to data only within the range of the training set. Extrapolation to compositions outside the training set range is risky. However, in the present case, the small deviations are assumed to be insignificant. Perhaps more important is the fact that the supplemental data set ranges are considerably more restricted than the training data set ranges, both for the FN and for the Cr, Ni, Mo, N, Cu, Ti, V and Co elements. As a result, this supplemental data set is only able to assess the accuracy of the neural network predictions over a limited portion of the total range of applicability. This issue will be discussed in greater detail in Part 2, where the results are presented.

Neural Network Development

The development of a neural network involves four basic steps, as shown in Fig. 2. The first step requires the identification of the input and output node variables. In the present study, 13 input nodes representing 13 different element concentrations (C, Cr, Ni, Mo, N, Mn, Fe, Si, Cu, Ti, Nb, V and Co) were used. A single output node, corresponding to the Ferrite Number, was used. The range of validity of the neural network corresponds to the range in composition and Ferrite Number in the training data set, as listed in Table 1.

The next step is to determine the optimum architecture for the neural network and, in particular, the optimum number of hidden nodes. In a broad sense, increasing the number of hidden nodes is comparable to increasing the "order" in a regression analysis (2nd order, 3rd order, etc.). Optimization of the number of hidden nodes is required to achieve the best fit without overfitting the data. To achieve this goal, the complete training data set was subdivided into pairs of training/testing subsets, with 90% of the data in the former and 10% in the latter. Four such pairs were created, with no overlap of the data in the four test subsets. The division of data into training/testing pairs was done randomly. Different network architectures were created with different numbers of hidden nodes, varying from one to twenty. For each of these trial architectures, four different neural networks were trained, corresponding to each of the four training/testing pairs of data. The "learnability" of the neural networks was measured by calculating the root mean square (RMS) of the difference between the calculated and experimentally measured FN for the training subset (better learnability corresponds to a lower RMS). Similarly, the "predictability" of the neural networks was assessed by the RMS for the neural network predictions on the testing data subset. It is important to note that this predictability value is based on FN predictions for data that were not used in the training of the network, and so it is a true representation of the prediction accuracy.

Average RMS values (over the four training/testing subsets) were used to evaluate the learnability and predictability of the neural networks as a function of the number of hidden nodes. By using average RMS values, the dependence of the results on the particular subdivision of the complete data set was minimized. Plots of the average RMS errors for both the learning and testing data sets vs. the number of hidden nodes are shown in Fig. 3. It can be seen that the learning

RMS initially decreases with increasing number of hidden nodes, but beyond three nodes the learning RMS is relatively constant. In contrast, the predicting RMS has a very distinct minimum for six hidden nodes. To confirm that this behavior is real, and not an artifact, a more extensive analysis was performed for five, six and seven hidden nodes. Ten learning/testing pairs, with the total of all the testing subsets covering the entire training data set, were used to train ten neural networks for each of these three configurations. The same behavior, with a minimum in the predicting RMS for six hidden nodes, was found, although the degree of improvement at six nodes was considerably reduced. Since the lowest RMS corresponds to the minimum error between predicted and measured FN, a network with six hidden nodes was chosen as the optimal network architecture. The final, optimum neural network architecture is shown in Fig. 4.

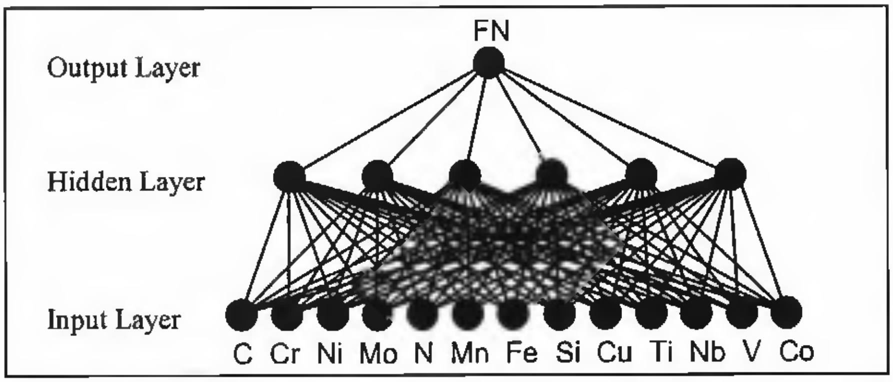

Fig. 4 — Final, optimum neural network architecture for predicting Ferrite Number.

Normally, one can expect the learning RMS error to decrease continuously with increasing number of hidden nodes. This simply reflects the fact that with more hidden nodes, the network can do a better job of fitting the network parameters to the data. This was not found in the present study. The results suggest that the initial decrease in learning RMS and the leveling off after three hidden nodes is a result of the scatter in the training data set. As noted above, the training data set is a compilation of data from several sources and for several arc welding techniques. Discrepancies due to the different processes or scatter in chemical analyses and ferrite measurements is an unavoidable consequence. Therefore, some of the data showed different FN values for basically identical compositions. Whether this is due to errors or to different welding techniques is immaterial as far as the network is concerned. The end result is that there is some degree of stochasticity in the data and the network is limited in the degree to which it can learn relationships between the inputs and outputs. Even under ideal circumstances, when the learning RMS error decreases continuously with increasing number of hidden nodes, the optimal network configuration is not the one with the lowest learning RMS error. Such a configuration represents a network that has memorized the training set very well. Rather, the best architecture is the one with the lowest predicting RMS error. This was the criterion that was used in this study. The fact that the learning RMS reaches a plateau very quickly is an indication that a good generalization scheme was found by the neural network and that the neural network analysis is appropriate for this application.

Referring to Fig. 2, the next step is to find the best neural network. With the optimum architecture identified, 80 neural networks with different, randomly selected initial weight values were developed. The neural networks were trained on the complete training data set. The learning RMS was calculated for each of these networks and the one with the minimum RMS was selected as the "best" network. It should be noted that an absolute best neural network is not likely to be identified. This is for two reasons. First, training can be continued indefinitely and it is always possible that a marginally better network will be produced. In the present study, training was discontinued when 200,000 additional iterations, checked every 10,000 iterations, did not yield any improvement in the learning RMS. Second, it is always possible to consider additional initial weight distributions and these may also lead to an improved network accuracy. However, the top five networks among the 80 tested in this study had RMS values that varied by only 4% (relative). This is a small difference and, therefore, it is expected that further testing and training would yield only marginally better networks, with insignificant improvements in FN prediction accuracy. The best network is defined by the values for the weights, which are given in the tables in Appendix B. These weights are the coefficients w that are used in Equations A-2 and A-4 in Appendix A.

The final step in the analysis is to evaluate the accuracy of the neural network in predicting the FN. This was accomplished by several means. First, the differences between the calculated FN and the experimentally measured FN for the complete training data set were compared to those using other FN prediction methods. Second, the independent, supplemental data set was used to test the best network. Finally, a small number of data points (10, or roughly 1% of the 961 total points) were removed from the complete training data set, and the neural network was trained on the remaining (nearly complete) data set using the optimal architecture and the optimal starting weights. In this way, a network comparable to the best network was created, and a prediction test was run on the few (ten) removed data points that were not used in the training. There are advantages and disadvantages of each method of accuracy evaluation, and these are discussed with the results presented in Part 2. However, all three methods indicated the best neural network was quite accurate in predicting the FN, with a significant improvement in accuracy over other available FN prediction methods, including the WRC-1992 constitution diagram.

Conclusions

A neural network analysis is applied for the prediction of Ferrite Number in stainless steel arc welds as a function of weld composition. The development of the neural network model is described in detail, including the identification of the inputs and outputs and the optimum network architecture. The neural network model (FNN-1999) uses 13 element concentrations as input (C, Cr, Ni, Mo, N, Mn, Si, Fe, Cu, Ti, Nb, V and Co). The final network weights that correspond to the "best" network are presented. In Part 2, the results of calculations for predicting Ferrite Number using the best neural network are given. It is shown in Part 2 that the new FNN-1999 model provides a simple and quick means for predicting Ferrite Number in weldments with a substantial improvement in the accuracy of the prediction compared to other models that are currently available. In addition, the added flexibility that is available with the neural network method in predicting elemental effects on Ferrite Number as a function of composition is demonstrated in Part 2.

Acknowledgments

This research was sponsored by the Division of Materials Sciences, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp. The authors wish to acknowledge the help of C. McCowan, National Institute of Standards and Technology, Boulder, Colo., for providing the supplemental data set in electronic form and F. Lake, ESAB Welding and Cutting Products, Inc., for providing the software for calculating the WRC-1992 Ferrite Number as a function of composition. The authors would also like to thank S. S. Babu for making the FNN-1999 model available on the World Wide Web. Finally, the authors wish to acknowledge S. S. Babu, D. J. Kotecki and T. A. Siewert for reviewing the manuscript and providing helpful comments.

References

1. Schaeffler, A. 1949. Constitution diagram for stainless steel weld metal. *Metal Progress* 56: 680–680B.
2. Hull, F. C. 1973. Delta ferrite and martensite formation in stainless steels. *Welding Journal* 52(5): 193-s to 203-s.
3. DeLong, W. T. 1974. Ferrite in austenitic stainless steel weld metal. *Welding Journal* 53: 273-s to 286-s.
4. Kakhovskii, N. I., Lipodaev, V. N., and Fadeeva, G. V. 1985. The arc welding of stable austenitic corrosion-resisting steels and alloys. *Avt. Svarka* 5: 55–57.
5. Olson, D. L. 1985. Prediction of austenitic weld metal microstructure and properties. *Welding Journal* 64(10): 281-s to 295-s.
6. Siewert, T. A., McCowan, C. N., and Olson, D. L. 1988. Ferrite number prediction to 100 FN in stainless steel weld metal. *Welding Journal* 67(12): 289-s to 298-s.
7. Kotecki, D. J., and Siewert, T. A. 1992. WRC-1992 constitution diagram for stainless steel weld metals: a modification of the WRC-1988 diagram. *Welding Journal* 71(5): 171-s to 178-s.
8. Schneider, H. 1960. Investment casting of high-hot strength 12% chrome steel. *Foundry Trade Journal* 108: 562–563.
9. Schoefer, E. A. 1974. Appendix to Mössbauer effect examination of ferrite in stainless steel welds and castings. *Welding Journal* 53(1): 10-s to 12-s.
10. Babu, S. S., Vitek, J. M., Iskander, Y. S., and David, S. A. 1997. New model for prediction of ferrite number of stainless steel welds. *Science and Technology of Welding and Joining* 2(6): 279–285.
11. Vitek, J. M., DasGupta, A., and David, S. A. 1983. Microstructural modification of austenitic stainless steels by rapid solidification. *Metallurgical Transactions A* 14A: 1833–1841.
12. Katayama, S., and Matsunawa, A. 1984. Solidification microstructure of laser welded stainless steels. *Proc. of ICALEO 84*, Boston, Mass., 44: 60–67.
13. David, S. A., Vitek, J. M., and Hebble, T. L. 1987. Effect of rapid solidification on stainless steel weld metal microstructures and its implications on the Schaeffler diagram. *Welding Journal* 66(10): 289-s to 300-s.
14. Bobadilla, M., Lacaze, J., and Lesoult, G. 1988. Influence des conditions de solidification sur le déroulement de la solidification des aciers inoxydables austénitiques. *Journal of Crystal Growth* 89: 531–544.
15. Elmer, J. W., Allen, S. M., and Eagar, T. W. 1989. Microstructural development during solidification of stainless steel alloys. *Metallurgical Transactions A* 20A: 2117–2131.
16. Lippold, J. C. 1994. Solidification behavior and cracking susceptibility of pulsed laser welds in austenitic stainless steels. *Welding Journal* 73(6): 129-s to 139-s.
17. Koseki, T., and Flemings, M. C. 1997. Solidification of undercooled Fe-Cr-Ni alloys: part III, phase selection in chill casting. *Metallurgical and Materials Transactions A* 28A: 2385–2395.
18. Szumachowski, E. R. and Kotecki, D. J. 1984. Effect of manganese on stainless steel weld metal ferrite. *Welding Journal* 63(5): 156-s to 161-s.
19. Self, J. A., Matlock, D. K., and Olson, D. L. 1984. An evaluation of austenitic Fe-Mn-Ni weld metal for dissimilar metal welding. *Welding Journal* 63(9): 282-s to 288-s.
20. Vitek, J. M., Iskander, Y. S., Oblow, E. M., Babu, S. S., and David, S. A. 1999. Neural network model for predicting ferrite number in stainless steel welds. *Trends in Welding Research*, eds. J. M. Vitek, S. A. David, J. A. Johnson, H. B. Smartt and T. DebRoy, pp. 119–124, ASM International, Materials Park, Ohio.
21. Vitek, J. M., Iskander, Y. S., and Oblow, E. M. 2000. Improved ferrite number prediction in stainless steel arc welds using neural networks — part 2: neural network results. *Welding Journal* 79 (2): 41-s to 50-s.
22. Rumelhart, D. E., Widrow, B., and Lehr, M. A. 1994. The basic ideas in neural networks. *Communications of the ACM* 37(3): 87–92.
23. Bishop, C. M. 1994. Neural networks and their applications. *Review of Scientific Instruments* 65(6): 1803–1832.
24. NeuralWare, Inc., Pittsburgh, Pennsylvania. 1995.
25. Kotecki, D. J. 1986. Ferrite control in duplex stainless steel weld metal. *Welding Journal* 65(10): 273-s to 278-s.
26. Kotecki, D. J. 1989. Heat treatment of duplex stainless steel weld metal. *Welding Journal* 68(11): 431-s to 441-s.
27. McCowan, C. N., Siewert, T. A., and Olson, D. L. 1989. Stainless steel weld metal: prediction of ferrite content. *WRC Bulletin* 342: 1–36.
28. Kotecki, D. J. 1997. Ferrite determination in stainless steel welds — advances since 1974. *Welding Journal* 76(1): 24-s to 37-s.
29. Kotecki, D. J. 1988. Verification of the NBS-CSM Ferrite Diagram. International Institute of Welding Document II-C-834-88.
30. Ornig, H. 1988. *Proceedings International Institute of Welding 41st Annual Assembly and International Conference*, Vienna, July.

Appendix A — Mathematical Relationships Used to Describe the Neural Network.

The sequence of steps from real input values to real output values that is described in the text can be expressed by a corresponding sequence of equations. First, the normalized input values are calculated, based on the minimum ($V_{i,min}$) and maximum ($V_{i,max}$) values for each input variable, and are given by

$$I_i = \frac{V_i - V_{i,\,min}}{V_{i,\,max} - V_{i,\,min}} \tag{A1}$$

where V_i is the real input values. With normalized input node values, I_i, for all i input nodes, the sum S_j at hidden node j is given by

$$S_j = \sum_i I_i\, w_{ij} + I_0 \tag{A2}$$

where w_{ij} is the weights for the connections between the input nodes i and hidden nodes j, and I_0 is a constant known as the bias. This sum is calculated for each of the j hidden nodes. The value of the hidden node is then calculated from the sum with the use of a transfer function. In the current study, sigmoid transfer functions were used. Therefore, the value of each hidden node, H_j, is given by

$$H_j = \frac{1}{1 + e^{-S_j}} \tag{A3}$$

where S_j is given by Equation A2. The next step is to sum the contributions from each of the hidden nodes to each of the output nodes, k. The sum at each output node is given by

$$S_k = \sum_j H_j\, w_{jk} + H_0 \tag{A4}$$

where w_{jk} is the weights connecting hidden nodes j with output nodes k. Once again, a constant bias term, H_0, is included. The values for H_j are given by Equation A3. As was done for the hidden nodes, the sum at the output node is converted to an output value, O_k, by means of another (sigmoid) transfer function

$$O_k = \frac{1}{1 + e^{-S_k}} \tag{A5}$$

where the values for S_k are obtained from Equation A4. Finally, the network output values, O_k, must be converted into real output values. When training the neural network, the output values were normalized within the range of 0.2 to 0.8 in order to utilize the linear range of the sigmoid function. This was accomplished by normalizing the (experimental) outputs as follows:

$$O_k = 0.2 + \frac{V_k - V_{k,\ min}}{V_{k,\ max} - V_{k,\ min}}\ 0.6 \qquad \text{(A6)}$$

where V_k is the real output values and $V_{k,min}$ and $V_{k,max}$ are the real minimum and maximum output values, respectively. From Equation A6, the "denormalization" of network outputs into real values can be derived and it is given by

$$V_k = V_{k,\ min} + (O_k - 0.2)(V_{k,\ max} - V_{k,\ min})/0.6 \qquad \text{(A7)}$$

In the present study, the calculated output is Ferrite Number, which cannot be negative. Therefore, if the calculated Ferrite Number (V_k in Equation A7) is negative, it was reset to 0.

With the sequence of operations defined by Equations A1 through A7, outputs can be calculated directly from input values. The sequence of steps can be easily incorporated into a spread sheet format. Thus, the prediction of Ferrite Number from alloy compositions can be simple and quick. The neural network weights, as well as the composition ranges used in the Ferrite Number analysis, are given in Appendix B.

Appendix B — Trained FNN-1999 Neural Network Parameters for Ferrite Number Prediction.

The FNN-1999 neural network described in this study used 13 inputs corresponding to 13 elemental concentrations. The minimum and maximum concentrations over which the network was trained, and over which it is expected to be valid, are given in Table B-1. Note that non-zero minima are specified for eight elements while concentrations of 0 are acceptable for Cu, Ti, Nb, V and Co. If the concentration for any of these latter five elements is not known (and is presumably at a residual level), then a value of 0 is to be used in the network calculations. The rationale for this procedure was as follows.

For Cu, Ti, Nb, V and Co, analyzed concentrations were not available for much of the data in the complete training data set, as well as for the supplemental data set. Since a concentration had to be used for all the data when developing and using the neural network model, a concentration had to be assigned in those cases where chemical analyses were unavailable. When concentrations were unknown, it was assumed that the element levels were residual, and therefore small. Although assigning a residual level is somewhat arbitrary and risky, it was necessary and a value of 0 was chosen. In order to minimize the potentially harmful impact of arbitrarily assigning concentrations for these elements when none were available, artificial maxima in the concentration ranges for these five elements were introduced. These maxima were used when normalizing the concentrations (Equation A1). With an artificial maximum that was four times as high as the real maximum, the difference between the "assigned" value of 0 and the real, unmeasured, residual (small) concentration was minimized when converting the concentration to a normalized value (Equation A1). It was found that if a higher artificial maximum, at ten times the real maximum value, was used, the resultant neural networks produced less accurate network predictions. The artificial maxima used in the normalization calculation of Equation A1 ($V_{i,max}$) are listed in Table B1. Additional calculations confirmed that the assignment of a 0 concentration vs. a small residual level led to a minimal, insignificant change in FN. For example, changing the Cu concentration from 0 to 0.3 (a large but still residual level) for an austenitic stainless steel base composition results in a

Table B-1 — Composition Ranges over which the FNN-1999 Network Was Trained and over which It Is Valid

	C	Cr	Ni	Mo	N	Mn	Fe	Si	Cu	Ti	Nb	V	Co
Node No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Min.	0.008	14.74	4.61	0.01	0.01	0.35	45.599	0.03	0.0	0.0	0.0	0.0	0.0
Max.	0.2	32	33.5	6.85	0.3	12.67	72.51	1.3	3.04	0.54	0.88	0.23	0.32
Artificial Max.									12.16	2.16	3.52	0.92	1.28

The Min. and Max. values are the ones to be used in the normalization of the alloy composition (see Equation A-1) except for Cu, Ti, Nb, V and Co, where the Artificial Max is used. All values are given in weight percent.

Table B-2 — FNN-1999 Neural Network Weights for Input Layer to Hidden Layer Connections

Input Node Number, Element and Bias	Hidden Layer Node Number 1	2	3	4	5	6
1, C	−0.77387279	0.018434081	0.82022685	0.11006859	0.73543739	0.4691014
2, Cr	−0.069039397	−2.5659544	−0.044000875	−2.9419811	−3.9367428	−1.1520522
3, Ni	2.6727166	7.9642334	1.5666518	−0.23706625	4.4325504	0.048408136
4, Mo	−0.44477677	−1.4908272	2.6978562	−1.1000808	1.0438977	−1.2577403
5, N	0.44379342	0.56137705	0.96918595	0.11622228	2.2909589	−0.92807078
6, Mn	−0.044638768	1.0530523	−0.047280665	0.39095843	0.28415173	−0.55964953
7, Fe	−1.3890632	0.0016147093	−1.5885487	1.3223488	−1.071898	−0.21818578
8, Si	−2.0870762	−0.12974143	0.21457124	2.4922497	0.59629947	−0.21896739
9, Cu	2.0396113	1.8920177	0.83045024	−1.3074456	0.4277178	−0.67360127
10,Ti	−1.5564938	−4.5376453	−0.2121184	1.1786785	−1.1766754	0.006974767
11, Nb	0.26728973	−0.080053128	−0.077407442	0.0086041111	−0.20916402	−0.028071323
12, V	−0.037870687	−2.5422843	−1.7820013	−1.2304662	−0.80585718	1.1488333
13, Co	0.99747533	−1.0812732	−1.4582157	−1.4943409	−1.4962931	0.71126705
Bias	0.87582731	1.9934373	0.34536767	−0.6946488	−0.34683022	−0.69136626

Table B-3 — FNN-1999 Neural Network Weights for Hidden Layer to Output Layer Connections

Hidden Node Number, Bias	Output Node Weight
1	2.4113946
2	−5.8647647
3	2.0065274
4	2.2444227
5	−2.1600745
6	−1.2715312
Bias	2.2027299

change in predicted FN from 5.2 to 5.3, which is negligible.

The Fe concentration is also treated as an input. This may seem to be redundant in that the Fe level is normally the remaining balance after all other concentrations are given. However, this was included in case some elements other than the 13 listed in Table B1 are present in the alloy. In other words, the sum of the concentrations for the 13 elements specified in Table B1 does not necessarily equal 100. In particular, S and P concentrations were not used in the calculations and the use of an Fe concentration accounts for the absence of these elements.

The final, best neural network is defined by the weights for the node connections. The network weights, w_{ij} (Equation A2) and w_{jk} (Equation A4), as well as the constant bias factors, I_0 and H_0, are given in Tables B2 and B3. Finally, the minimum and maximum Ferrite Numbers must be specified in order to convert the network outputs (O_k) into actual Ferrite Numbers (V_k, see Equation A7). The neural network analysis considered data over the range of FN from 0 to 117, and these are the values of $V_{k,\,min}$ and $V_{k,\,max}$ that are used in Equation A7. In addition, these are the ranges in Ferrite Number over which the FNN-1999 model predictions are expected to be valid.