

Neural network analysis of strength and ductility of welding alloys for high strength low alloy shipbuilding steels

E. A. Metzbower, J. J. DeLoach, S. H. Lalam, and H. K. D. H. Bhadeshia

There are considerable demands for the development of weld metals for high strength low alloy steels. To assist in meeting such demands, a neural network was trained and tested on a set of data obtained on weld metals for steels of the type used for shipbuilding. The input variables for the network were the chemical elements and the weld cooling rate. The outputs consisted of the yield and ultimate tensile strengths, elongation, and reduction of area. Many models were created using different network configurations and initial conditions. An appropriate committee of models was then assembled by testing the ability of the models to generalise on unseen data. The neural network technique used is due to MacKay, with a Bayesian framework, and hence allows the estimation of error bars, which warn the user when data are sparse or locally noisy. The method revealed significant trends describing the dependence of mechanical properties on weld composition and cooling rate.

Dr Metzbower is with the Physical Metallurgy Branch, US Naval Research Laboratory, Code 6320, Washington, DC 20375-5320, USA. Dr DeLoach is with the Welding and Non Destructive Evaluation Branch, Carderock Division, US Navy Surface Warfare Center, West Bethesda, MD 20817-5700, USA. Mr Lalam and Professor Bhadeshia are in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK. Manuscript received 15 April 2000; accepted 15 July 2000.

© 2001 IoM Communications Ltd.

INTRODUCTION

The US Navy has been working on a programme to replace high yield (HY) and high strength low alloy (HSLA) steels for surface ship construction. The compelling reason for this substitution is the cost saving associated with the reduction or elimination of preheating during welding operations. Currently, HSLA steels are welded using filler metal developed for the HY steels. Ideally, the weld metal should be designed specifically for HSLA steels with a view to reducing or eliminating the necessity for preheating, and to expand the envelope of safe welding conditions. There is, as always, a demand for a reduction in the risk from hydrogen cracking. Assuming that these goals can be achieved, the costs of welding would be reduced through higher productivity, reduced energy costs, reduced manpower requirements (both time and skill), and less post-weld inspection. A programme was instituted and developed to provide an integrated approach to the development and certification of improved filler metals for welding

550–690 MPa yield strength steels for surface ships and submarines. The objective of the programme was to develop, optimise, and certify improved filler metals for welding HSLA and HY steels.¹ The mechanical properties required for the filler metal are given in Table 1.

The above research programme to develop a suitable welding consumable has generated significant quantities of experimental data, which provide an opportunity for creating a quantitative model for the estimation of weld mechanical properties. Such a model could be useful in understanding the effect of each variable on the tensile properties of HSLA weld metals, and might be used to assess the behaviour of the welds over a wider range of parameters than that covered by the experimental data. It should in principle be possible to improve on the current or proposed welding consumables.

The method used in the present work is neural network modelling. A neural network is capable of replicating a great variety of non-linear relationships. Data are presented to the network in the form of input and output parameters, and the optimum non-linear relationship found by minimising the difference between the measured value and the predicted value. As in regression analysis, the results then consist of a series of coefficients (known as weights) and a specification of the type of function which, in combination with the weights, relates the inputs to the output. To prevent overfitting, MacKay²⁻⁵ has developed a Bayesian framework to control the complexity of a neural network. This framework also provides quantified error bars on the predictions of the network and renders it possible to identify automatically which of the many possibly relevant input variables are in fact important factors in the regression. This specific technique has been reviewed elsewhere^{5,6} and hence will not be discussed further except in the context of the present results.

DATABASE

The weld metal data generated in the research programme have been used to create the neural network. They include the chemical composition of the as deposited weld beads, the measured yield strength (YS) and ultimate tensile strength (UTS), and the cooling rate at 538°C. The cooling rate was determined from the welding parameters (voltage, current, welding speed), preheat, and the thickness of the plate.^{7,8} The input data included alloys that did not satisfy the acceptance criteria for YS or UTS in the context of ship

Table 1 Mechanical property objectives for improved filler metals

Yield strength, MPa	Charpy V notch impact testing value, J	
	At -1°C	At -51°C
607–690	70	47

construction. Nevertheless, all data were included so that the greatest predictive capability over the widest chemistry and cooling rate ranges could be developed. Table 2 gives the minimum, maximum, mean, and standard deviation of each of the variables, calculated from the 189 in the assembled database. The welds were made in plates with 60° included angle, with single V notch or double V notch preparation for the 25.4 and 50.8 mm thickness plates respectively.

The complete chemical compositions of the as deposited weld beads had not been determined in every instance; in particular, the oxygen and nitrogen concentrations were frequently omitted. Since these two elements are significant, it is not safe to set them to zero values; instead, average values were substituted for 60 of the 189 instances where the oxygen and nitrogen data were unavailable.

The predominant welding process used was gas metal arc welding in both spray and pulsed modes. Approximately 30 flux cored welds were included in the data. No attempt to relate the properties to the weld deposit type was made.

The major alloying additions, i.e. carbon, manganese, silicon, nickel, and molybdenum, affect the transformation products and, consequently, the strength and toughness. Impurity elements (sulphur, phosphorus, aluminium, nitrogen, oxygen) are included because of their known tendency to embrittle or because of their importance in the formation of non-metallic inclusions in welds. The isolated effects of some of the above elements on the strength of steel plates are discussed in Ref. 9.

Throughout the present paper, the elongation and reduction of area (RA) are expressed as percentages. The elongation is measured relative to a starting gauge length of 50.8 mm, and the RA relative to the initial specimen diameter of 12.7 mm.

NEURAL NETWORK

The aim of the neural network is to be able to estimate the YS and UTS as a function of the variables given in Table 2. Both the input and output variables were normalised within the range ±0.5 as follows

$$x_N = \frac{x - x_{min}}{x_{max} - x_{min}} - 0.5 \dots \dots \dots (1)$$

where x_N is the normalised value of x , which has maximum and minimum values given by x_{max} and x_{min} . This normalisation is not required for the analysis, but facilitates a more straightforward subsequent comparison of the significance of each of the variables in influencing the output.

Eighty neural network models were created using the data. The models varied in terms of the number of hidden units and the random seeds used to initiate the network. For a given number of hidden units, five different sets of random initial values were used. The number of hidden units varied from 1 to 16 for the 80 different models.

Linear functions of the inputs x_j multiplied by the weights w_{ij} are operated on by a hyperbolic tangent transfer function

$$h_i = \tanh \left(\sum_j w_{ij}^{(1)} x_j + \theta_i^{(1)} \right) \dots \dots \dots (2)$$

so that each input contributes to every hidden unit, where N is the total number of hidden units. The bias is designated θ and is analogous to the constant that appears in linear regression. The strength of the transfer function is in each case determined by the weight w_{ij} . The transfer to the output y is linear

$$y = \sum_i w_i^{(2)} h_i + \theta^{(2)} \dots \dots \dots (3)$$

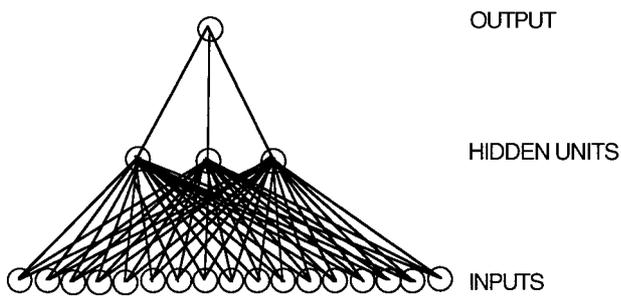
The output y is therefore a non-linear function of x_j , the function usually selected being the hyperbolic tangent because of its flexibility.

This specification of the network structure, together with the set of weights, is a complete description of the formula relating the inputs to the output. The weights were determined by training the network; the details are described elsewhere.²⁻⁵ The training involves a minimisation of the regularised sum of the squares. The complexity of the models is controlled by the number of hidden units and the values of the regularisation constants – one associated with each input variable, one for biases, and one for all weights connected to the output. The form of a typical network used in the analysis is illustrated in Fig. 1.

Table 2 Input and output variables for neural network models: concentrations in wt-% except where indicated

Variable	Minimum	Maximum	Mean	Standard deviation
C	0.008	0.062	0.0291	0.011
Mn	1.05	3.44	1.4063	0.2544
Si	0.05	0.4	0.2519	0.0581
Cr	0	0.21	0.0578	0.0487
Ni	1.66	5.63	3.5482	1.0141
Mo	0	1.23	0.4852	0.2255
Cu	0	0.48	0.0964	0.0789
S	0.001	0.012	0.004	0.0023
P	0.001	0.015	0.005	0.0035
Al	0.001	0.082	0.008	0.009
Ti	0.0008	0.3	0.014	0.0311
Nb	0	0.069	0.0041	0.0066
V	0	0.031	0.0038	0.0063
B	0	0.01	0.0007	0.0018
O, wt-ppm	109	627	191.0169	76.5035
N, wt-ppm	6	135	26.1008	21.9551
Cooling rate, K s ⁻¹	1.32	76.1667	26.9934	25.1342
Yield strength, MPa	485.8	918.162	677.3273	103.3283
Ultimate tensile strength, MPa	594.064	977.846	747.5553	87.7612
Elongation, %	3.5	29.2	20.8411	5.237
Reduction of area, %	7	84	64.6325	17.9575

Carbon concentration data apply to yield strength model. For ultimate tensile strength, the minimum, maximum, mean, and standard deviations for carbon were 0.002, 0.062, 0.0289, and 0.0113 wt-% respectively. For ductility parameters, the corresponding values for carbon are 0.001, 0.062, 0.0289, and 0.0114 wt-% respectively.



1 1 Schematic diagram of structure of network, showing input nodes, hidden units, and output node

A potential difficulty with the use of powerful regression methods is the possibility of overfitting data. To avoid this, the experimental data can be divided into two sets, namely, a training dataset and a test dataset. The model is produced using only the training data. The test data are then used to verify that the model generalises well when presented with previously unseen data. The test error will be large if the model is overly simple, so that it is unable to replicate the complexity in the data accurately, or if the model is so complex that it is erroneously modelling the noise in the experimental data. The optimum model corresponds to some case in between these two undesirable states.

The training process involves a search for the optimum non-linear relationship between the input and the output data and is computer intensive. Once the network is trained, estimation of the outputs for any given inputs is very rapid.

For the sake of brevity, the series of plots describing how the perceived level of noise σ_v and the test error vary with the different types of model is not included in the present paper. However, the performances of the most successful models in predicting the training and test data are presented in Fig. 2 for each of the mechanical properties studied. For both elongation and RA, it is evident that the single best models generalise poorly on unseen data; this emphasises the necessity for formation of committees. It would certainly not be reasonable to use the single models to make predictions.

For each property, therefore, a committee of best models is used to make predictions. It is possible that a committee of models can make a more reliable prediction than an individual model.⁵ The best models are ranked using the values of the test errors. Committees are then formed by combining the predictions of the best L models, where $L=1, 2, \dots$: the size of the committee is therefore given by the value of L . A plot of the test error of the committee versus its size L gives a minimum that defines the optimum size of the committee. Using this procedure, the optimum size of the YS committee was found to consist of 11 of the highest ranking models, whereas the committee for the UTS consisted of six of the highest ranking models. The optimum size of the elongation committee was found to consist of four of the highest ranking models, whereas the committee for the RA consisted of 41 of the highest ranking models. Once the optimum committee is selected, it is retrained on the entire dataset without changing the complexity of each model, with the exception of the inevitable, although relatively small, adjustments to the weights.

RESULTS AND DISCUSSION

Strength

Results

The correlations between the strength measurements and data predicted using the committee models are shown in Fig. 3. The committees evidently perform rather well. One point to note is that the error bars that are plotted here represent the fitting error, the magnitude of which depends on the position in the input space. The additional error σ_v is not plotted, but is constant and according to the most pessimistic assessment, can be taken as the highest value of σ_v for any member in the committee of models. Thus, for the YS and UTS models, the σ_v values for use with the committees are 0.076 and 0.061 respectively. Only the fitting error, which is a function of position in the input space, is plotted on all the diagrams included in the present paper.

The minimum required yield strength is 607 MPa. Examination of Fig. 3 shows that about one-quarter of the data points generated in the programme of experiments do not satisfy the design. Although there does not seem to be a requirement specified for the UTS, discussions below will show that the UTS should be monitored because of its connection with ductility.

The significance σ_w , which is plotted in Fig. 4, is a measure of the extent to which an input parameter can be correlated with variations in the output. In that sense it is rather similar to a partial correlation coefficient in linear regression analysis. It is therefore worth emphasising that whereas σ_w gives an indication of the correlation, it does not imply how sensitive the output is to the input – that information is in the weights.

In Fig. 4, there are several values of σ_w plotted for each input, since there are many models in each committee. Thus, the σ_w values from the top five (or fewer if the committee is smaller) models are displayed. The advantage of this procedure is that a strong dependence of the σ_w value on the model used indicates an uncertain relationship between that variable and the output, rendering the use of an individual model, as opposed to a committee of models, less reliable. An example is the very large variation in the perceived significance of manganese for the UTS. Generally, however, a strong dependence of σ_w on the model is not observed for most of the other inputs.

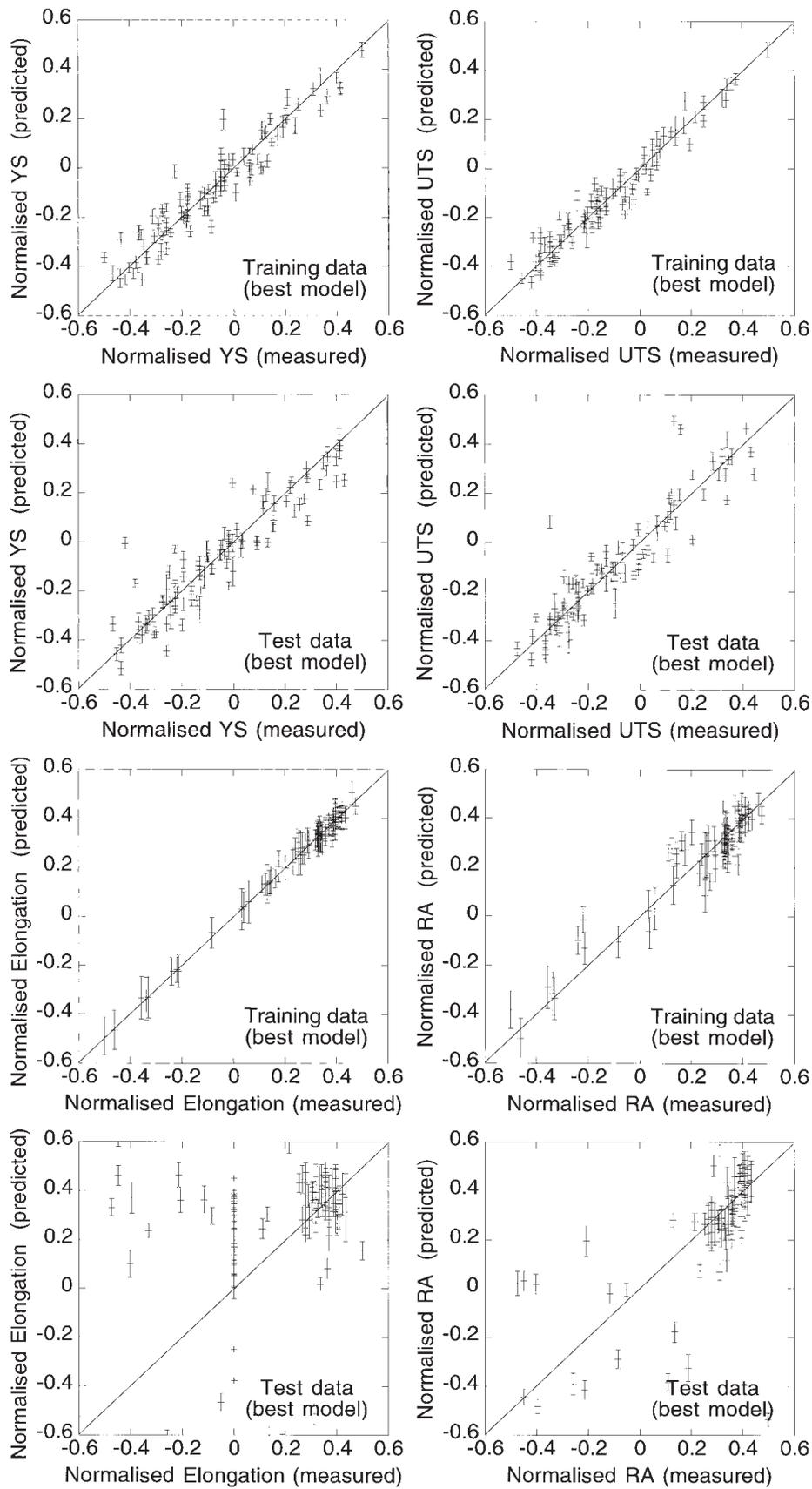
It is interesting and expected that carbon, manganese, silicon, nickel, molybdenum, titanium, and the cooling rate have a major role in determining the YS. Chromium does not, presumably because of the restricted range of concentration in the dataset. Furthermore, it is not surprising that sulphur does not affect the YS since it should be present mostly in the form of sulphide inclusions. Nitrogen is known to influence the YS via strain age hardening effects, but such effects should vanish following plastic deformation. Therefore, the σ_w values for nitrogen are significant for the YS but not for the UTS.

Predictions

The best method of examining the performance of the models is to use them to make predictions regarding trends as a function of each of the inputs, for a reasonable selection of baseline inputs. The set of baseline inputs is given in Table 3. The cooling rate was set at 61.2 K s^{-1} . An alloy such as this should have a mixed

Table 3 Standard set of inputs (concentrations) used to observe trends, wt-% (except where indicated)

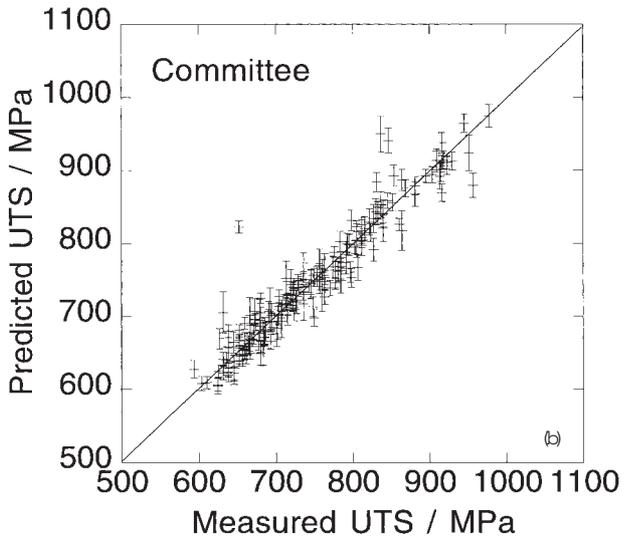
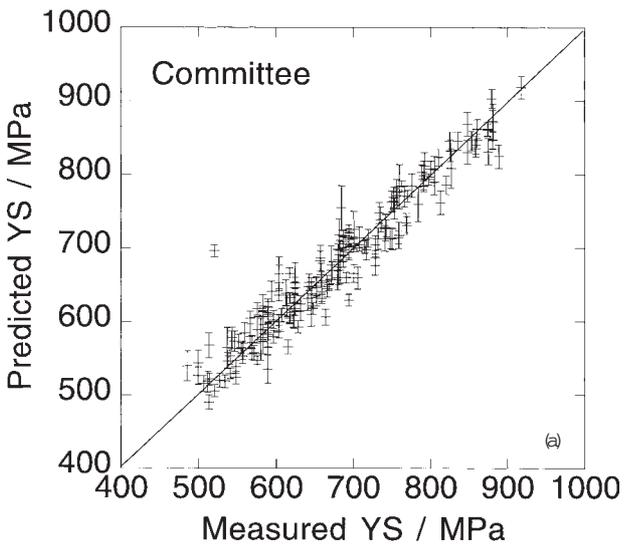
C	Si	Mn	P	S	Cr	Mo	Ni	Al	B	Cu	N, wt-ppm	Nb	Ti	V	O, wt-ppm
0.036	0.31	1.72	0.001	0.004	0.03	0.59	2.63	0.005	...	0.061	8	...	0.007	0.001	181



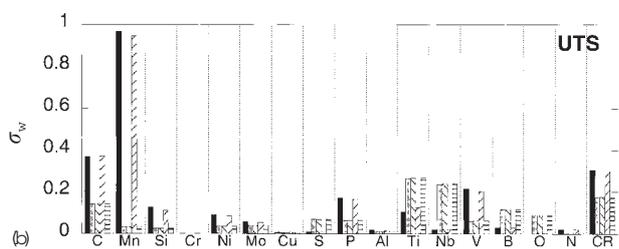
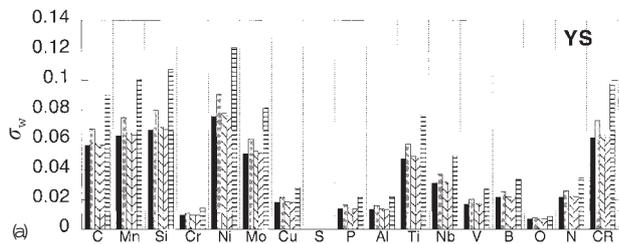
2 2 Predicted versus measured normalised values of yield strength (YS), ultimate tensile strength (UTS), elongation, and reduction of area (RA) for training and test datasets

microstructure of bainite and some martensite given the combination of high manganese and nickel concentrations in the presence of molybdenum.

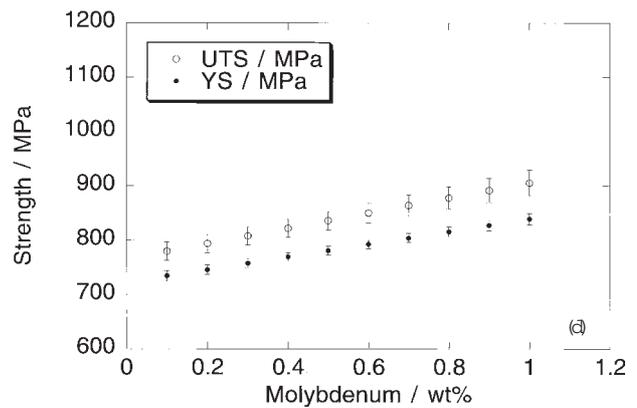
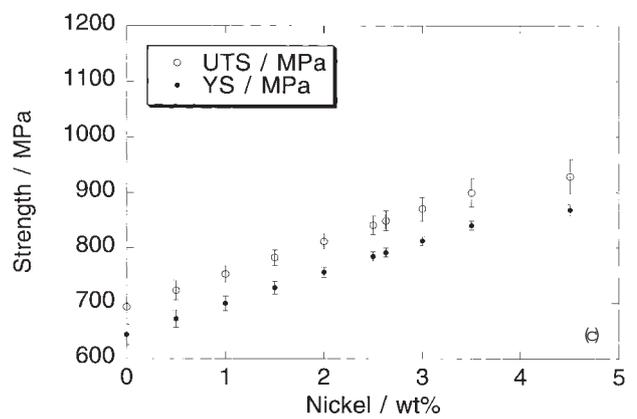
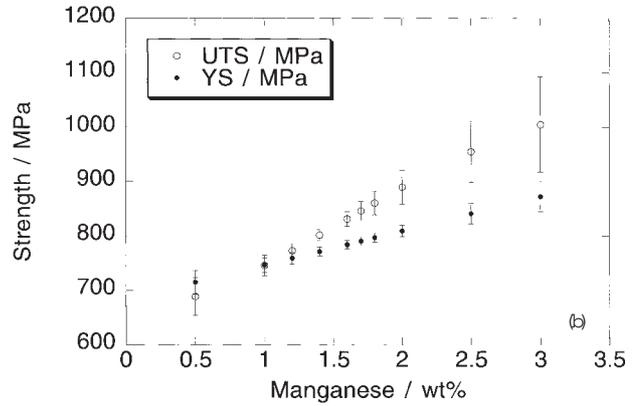
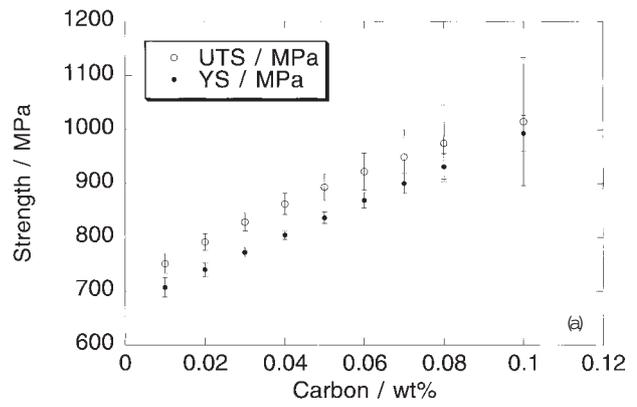
Figure 5 shows the effects of varying the carbon, manganese, nickel, and molybdenum concentrations on the YS and UTS. All the plots have the same vertical scale.



3 Comparison of predicted and measured values of *a* YS and *b* UTS for as deposited weld metal: calcu-

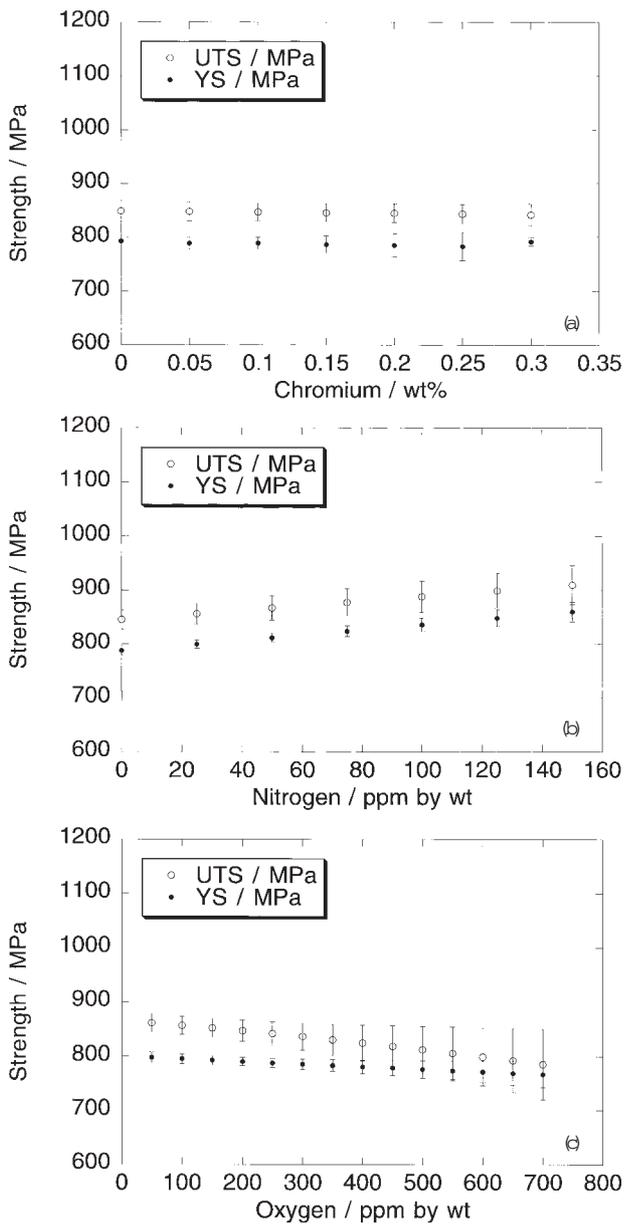


4 Bar graph showing perceived significance σ_w of each input in influencing variations in *a* YS and *b* UTS: data are for five highest ranking models in each committee (CR denotes cooling rate)



5 Variations in YS and UTS, predicted using committee models, as function of *a* carbon, *b* manganese, *c* nickel, and *d* molybdenum concentrations: error bars represent $\pm 1\sigma$ values

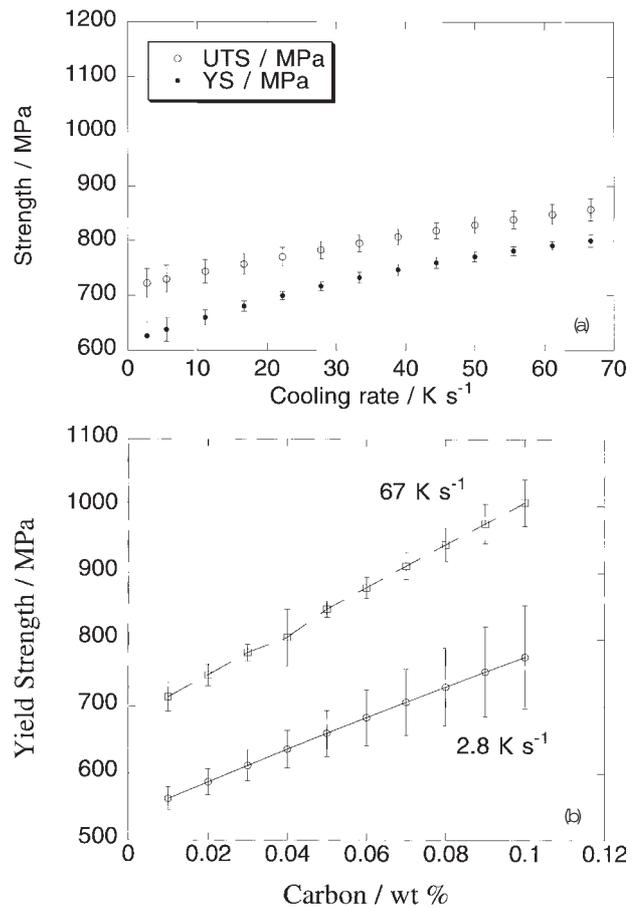
As expected, carbon has a strong effect on strength, although all the elements make significant contributions. For manganese, it is evident that a reduction in concentra-



6 6 Variations in YS and UTS, predicted using committee models, as function of *a* chromium, *b* nitrogen, and *c* oxygen concentrations: error bars represent $\pm 1\sigma$ values

tion below that indicated in Table 3 causes the difference between the YS and UTS to decrease, which may not be desirable from the point of view of fatigue. Note also that the YS is predicted to be higher than the UTS when the manganese concentration is reduced to 0.5 wt-%, but a statement such as this is not justified since it neglects the error bars, which can be seen to overlap. It is also interesting that the model predicts greater uncertainty at high manganese concentrations, in effect giving a warning that there is a lack of information in that region of the input space.

Figure 6 shows corresponding plots for chromium, nitrogen, and oxygen, again for the baseline data given in Table 3. Chromium, over the concentration range studied, has no significant influence. As expected, nitrogen causes strengthening, presumably via the strain aging effect described above. An increase in the oxygen concentration causes a greater decrease in the UTS, compared with a small effect on the YS; this is expected since the UTS is a plasticity



7 7 Variations in YS and UTS, predicted using committee models, as function of *a* weld cooling rate and *b* carbon concentration and cooling rate: error bars represent $\pm 1\sigma$ values

dominated property, and oxides nucleate voids which reduce the plastic strain before failure.

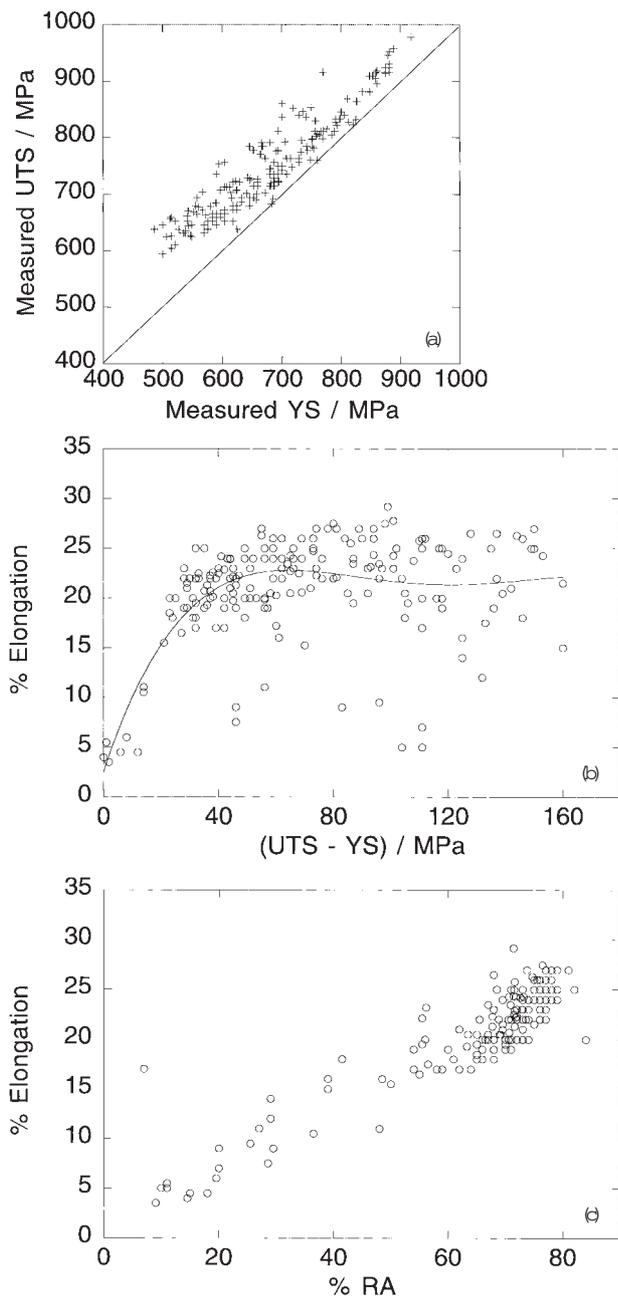
It is not surprising that an increase in the weld cooling rate causes an increase in strength (Fig. 7a). The high hardenability of the alloys considered (Table 3) must lead to a greater amount of martensite as the cooling rate increases. This is consistent with the observed higher sensitivity of the YS to the carbon concentration at high cooling rates (Fig. 7b) since the effect of carbon on hardness is greatest when it is in solid solution in martensite. Apart from the sensitivity to carbon, Fig. 7b also shows that for any given carbon concentration, the YS is much greater at the higher cooling rate; this can only occur if the higher cooling rate leads to a harder microstructure, i.e. one containing a greater fraction of martensite.

Finally, it is worth noting that Fig. 7b shows an important feature of non-linear modelling, namely, that it captures the interdependence of the variables. Thus, the predicted behaviour is different at high cooling rates.

Ductility

Naturally, strength alone is not a sufficient indicator of weld performance. The ductility and toughness must also meet design specifications. In the present work neural network models are developed for the tensile elongation and RA (models for the toughness will be considered in a future study).

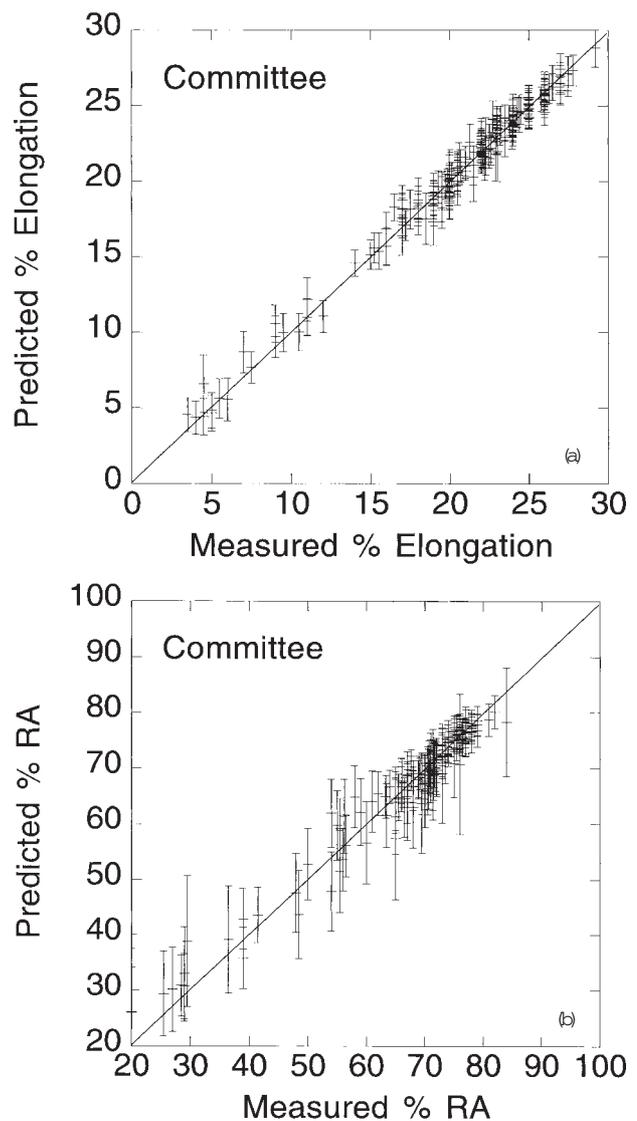
The tensile ductility is a complex function of the work hardening behaviour and of features such as inclusions, which initiate voids; the voids eventually link and failure results when this occurs on a sufficiently large scale. In a



8 8 *a* measured UTS as function of measured YS, *b* measured percentage elongation versus difference between UTS and YS, and *c* correlation between measured values of elongation and RA

very elementary interpretation, the work hardening tendency might be indicated by the difference between the YS and the UTS. Figure 8*a* shows a plot of the measured UTS versus the measured YS. It is evident that there are many instances where the difference (UTS–YS) is actually fairly small. Figure 8*b* illustrates that the elongation does indeed depend on the value of (UTS–YS), but only as the latter approaches zero. There is a great deal of unexplained variation, justifying a more detailed analysis with the inclusion of many more variables.

The two measures of ductility, i.e. elongation and RA, are not necessarily equivalent. Whereas both depend on a period of uniform strain followed by heterogeneous deformation, the elongation is measured over a large distance so it is possible to envision circumstances where the necking contributes little to elongation. In contrast, the RA is actually measured at the point on the sample where

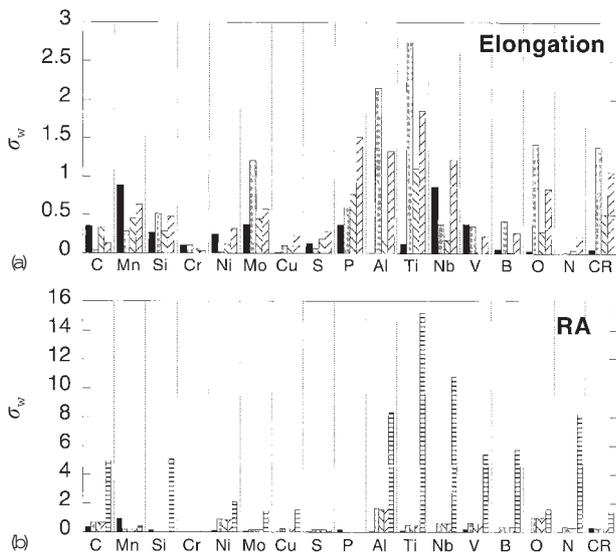


9 9 Comparison between predicted and measured values of *a* elongation and *b* RA for as deposited weld metal: calculations were carried out using committee models (error bars represent $\pm 1\sigma$ values)

the most intense deformation is focused. Nevertheless, Fig. 8*c* shows that for the experimental data considered here, there is a strong correlation between the two measures of ductility, which can be investigated using the neural network method. It is probable that the level of noise associated with the RA measurements should be greater than that in the elongation data, and analysis should be able to reveal such information.

Results

The correlations between the experimental measurements and data predicted using the committee models are shown in Fig. 9. Again, the committees evidently perform rather well. It is emphasised once more that the error bars that are plotted here represent the fitting error, the magnitude of which depends on the position in the input space. The additional error σ_v is taken as the highest value of σ_v for any member of the committee of models. Thus, for the elongation and RA models, the σ_v values for use with the committees are 0.036 and 0.14 respectively. The rather large perceived noise in the RA measurements is not surprising given both the smaller dimensions that are measured and the dependence of the process of final fracture on localised



10 10 Bar graph showing perceived significance of each input in influencing variations in *a* elongation and *b* RA: data are for up to five highest ranking models in each committee

heterogeneities. Only the fitting error, which is a function of position in the input space, is plotted on all the diagrams included in the present paper.

Some significance (σ_w) values are plotted in Fig. 10; the plot for the RA should not be considered to be very meaningful since there are 41 models in the committee, whereas only the data for the five highest ranking models are presented. The plot nevertheless shows the importance of using a committee since the different models evidently have different perceptions of significance. The elongation data do show the importance of the oxide forming elements aluminium, titanium, and oxygen.

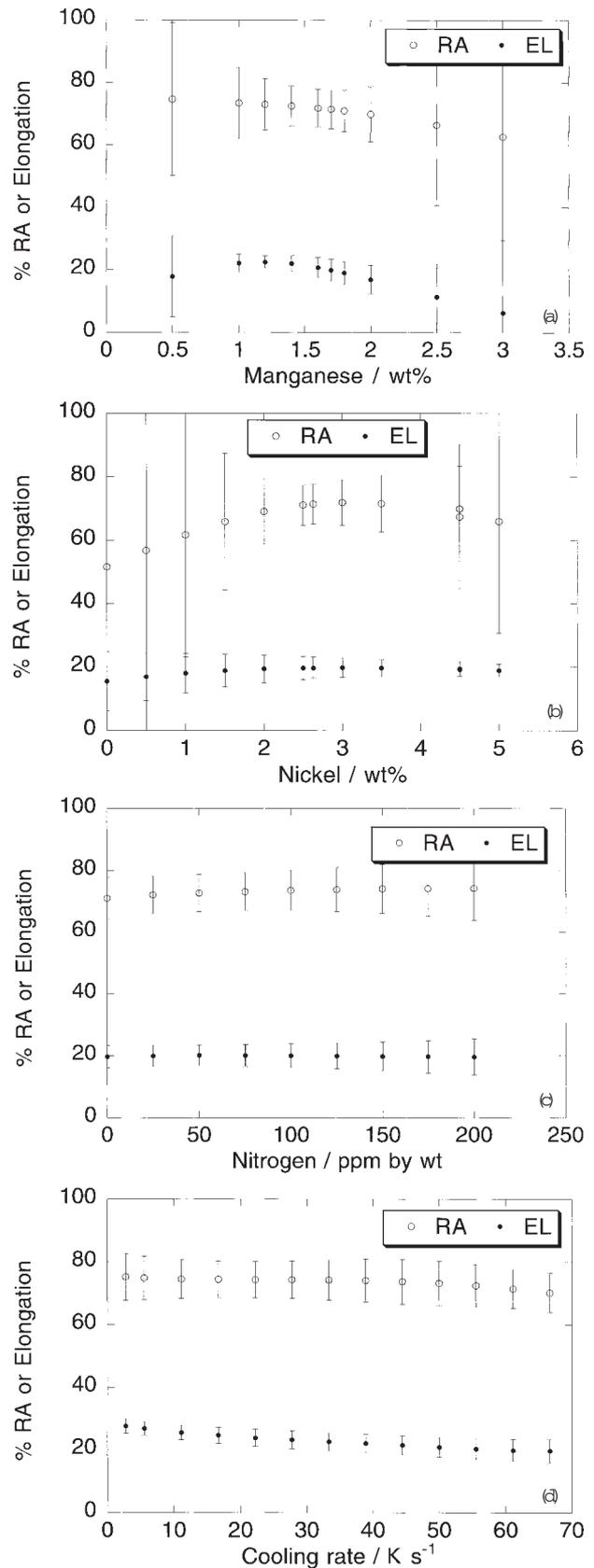
Predictions

As in the 'Predictions' subsection for strength above, the set of baseline inputs used to study trends is given in Table 3, and the cooling rate was set at 61.2 K s^{-1} . As pointed out above, an alloy such as this should have a mixed microstructure of bainite and low carbon martensite, given the combination of high manganese and nickel contents in the presence of molybdenum.

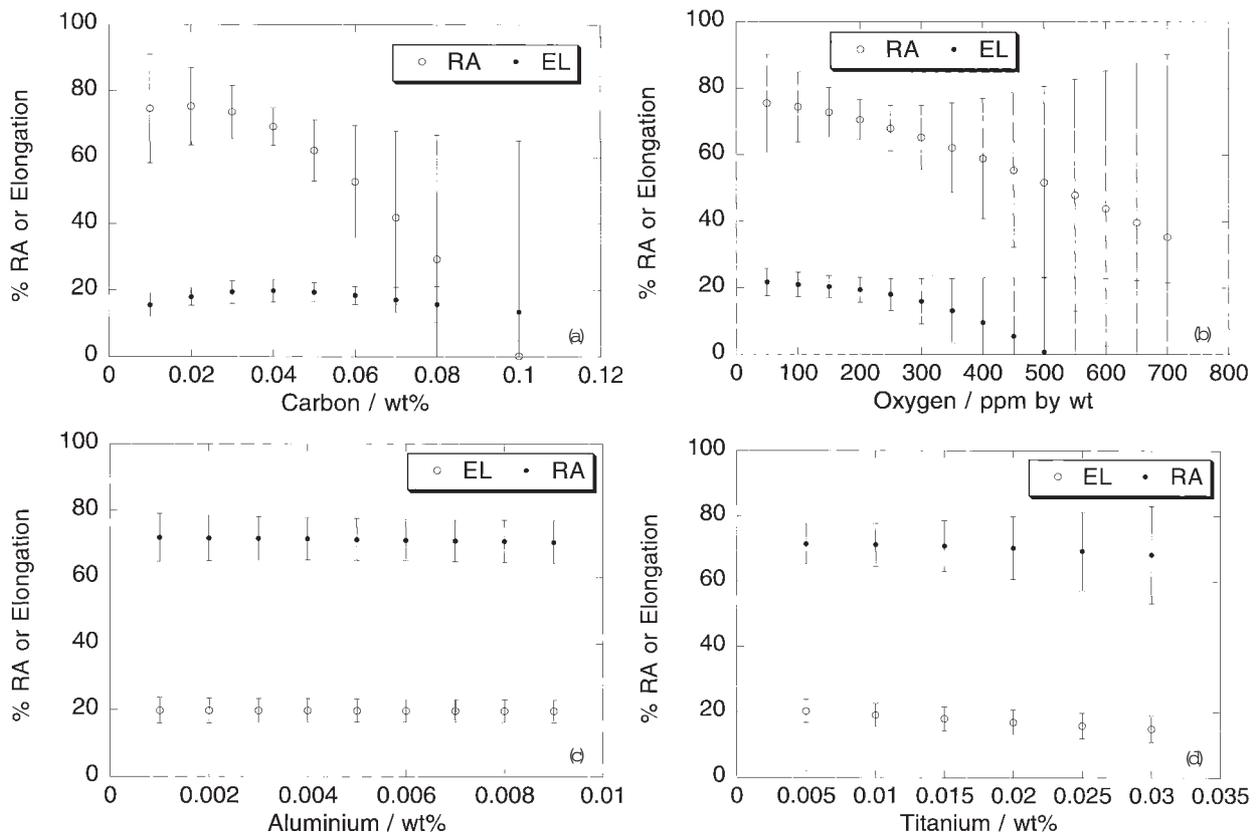
Figure 11 shows plots of the effects of varying the manganese, nickel, and nitrogen concentrations and the cooling rate on the elongation and RA. It is particularly interesting that all of these cause a substantial increase in strength but have only a small effect on ductility. Elements such as chromium and molybdenum show similar variations but their plots are not included for the sake of brevity. All the plots have the same vertical scale.

By contrast, those elements that form oxides or carbides show a significant deterioration in ductility as their concentrations are increased (Fig. 12). This is fully expected given that inclusions nucleate voids and hence play an important role in determining ultimate fracture. This interpretation must naturally be modified if alloying elements cause a great change in the work hardening behaviour, which has an effect on the extent of uniform strain before necking.

It is noteworthy that in Fig. 12c and d, the effects of aluminium and titanium can be seen to be rather small, and yet in Fig. 10 their roles are perceived to be significant. This is because although the elements are indeed significant in determining ductility, their quantitative effect is already included in the oxygen plot.



11 11 Variations in elongation and RA, predicted using committee models, as function of *a* manganese, *b* nickel, and *c* nitrogen concentrations, and *d* cooling rate: error bars represent $\pm 1\sigma$ values



12 Variations in elongation and RA, predicted using committee models, as function of *a* carbon, *b* oxygen, *c* aluminium, and *d* titanium concentrations: error bars represent $\pm 1\sigma$ values

CONCLUSIONS

A neural network method within a Bayesian framework has been used to model experimental data for the tensile properties of ferritic steel weld metals appropriate for the welding of high strength low alloy steels. The inputs for the neural network models consist of the detailed chemical composition and the weld cooling rate at 538°C. The models are found to be well behaved with a demonstrated ability to reproduce known metallurgical trends.

ACKNOWLEDGEMENTS

Funding support from the Office of Naval Research, Arlington, VA, USA, and from the Cambridge Commonwealth Trust, is gratefully acknowledged.

REFERENCES

- J. J. DELOACH, JR: in 'Welding and weld automation in shipbuilding', (ed. R. DeNale), 85–104; 1996, Warrendale, PA, TMS.
- D. J. C. MACKAY: *Neural Comput.*, 1992, **4**, 415.
- D. J. C. MACKAY: *Neural Comput.*, 1992, **4**, 448.
- D. J. C. MACKAY: *Neural Comput.*, 1992, **4**, 698.
- D. J. C. MACKAY: in 'Mathematical modelling of weld phenomena 3', (ed. H. Cerjak), 359–389; 1997, London, The Institute of Materials.
- H. K. D. H. BHADSHIA: *ISIJ Int.*, 1999, **39**, 966–979.
- L. P. CONNOR (ed.): 'Welding handbook', 8th edn, Vol. 1, 75–79; 1987, Miami, FL, American Welding Society.
- J. J. DELOACH, JR: personal communication to E. A. Metzbower, US Navy Surface Warfare Center, West Bethesda, MD, 2000.
- W. C. LESLIE: in 'The physical metallurgy of steels', Chap. 6; 1981, New York, Hemisphere Publications.
- H. K. D. H. BHADSHIA, D. J. C. MACKAY, and L.-E. SVENSSON: *Mater. Sci. Technol.*, 1995, **11**, (10), 1046–1051.
- T. COOL, H. K. D. H. BHADSHIA, and D. J. C. MACKAY: *Mater. Sci. Eng. A*, 1997, **A223**, 186–200.
- H. D. K. H. BHADSHIA, L.-E. SVENSSON, and B. GRETOFT: *Acta Metall.*, 1985, **33**, 1271–1283.