8. Neural Networks

Over the years, linear regression models have attempted to characterise the 0.2% proof stress and ultimate tensile strength of austenitic stainless steels. Pickering and Irvine [2] carried out a sequence of tests using 18Cr-8Ni, 18Cr-12Ni and 16Cr-25Ni wt% base steels, varying carbon (0.006-0.11 wt%), chromium (16-25 wt%), nickel (8-25 wt%), molybdenum (0-4 wt%), vanadium (0-4 wt%), tungsten (0-4 wt%), aluminium (0-2 wt%), silicon (0.3-4 wt%), manganese (1-8 wt%), copper (0-2 wt%) and cobalt (0-8 wt%). From this study, they developed some empirical relations for these properties, mainly as a function of chemical composition:

$$\begin{aligned} \text{0.2\% Proof Stress / MPa} = {} & 15.4\,\{4.4 + 23C_C + 1.3C_{Si} + 0.24C_{Cr} \\ & + 0.94C_{Mo} + 1.2C_V + 0.29C_W + 2.6C_{Nb} + 1.7C_{Ti} \\ & + 0.82C_{Al} + 32C_N + 0.16(\delta\text{-ferrite}) + 0.46d^{-1/2}\} \end{aligned} \tag{8.1}$$

$$\begin{aligned} \text{UTS / MPa} = {} & 15.4\,\{29 + 35C_C + 55C_N + 2.4C_{Si} + 0.11C_{Ni} \\ & + 1.2C_{Mo} + 5.0C_{Nb} + 3.0C_{Ti} + 1.2C_{Al} \\ & + 0.14(\delta\text{-ferrite}) + 0.82t^{-1/2}\} \end{aligned} \tag{8.2}$$

where C_x is the wt% of an element x, t is the annealing twin spacing (mm) and δ-*ferrite* is the volume fraction of delta-ferrite. Carbon and nitrogen were identified as the most important variables.

However, there are inherent problems with this approach. Firstly, any extrapolations may be unsafe, as there is no account for the uncertainty of predictions made beyond a constant standard error. Secondly, the models are essentially linear. The inputs also appear to be treated independently, even though it is known that they may interact.

A more powerful alternative is the use of neural networks [40,42], a non-linear modelling technique, which allows more complex relationships. The fundamental principles are quite simple. Examples of input variables and their outputs are given to the neural network, so that it can "learn" to model the relationship between them. Bayesian probability theory is used to control model complexity and indicate prediction uncertainties.

In linear regression, the sum of each input x_i multiplied with a weight w_i and a constant bias value θ, indicating the influence on an output y:

$$y = \sum_i w_i x_i + \theta \tag{8.3}$$

The function in a neural network looks similar to equation 8.3:

$$y = \sum_i w_i h_i + \theta \tag{8.4}$$

However, flexibility is incorporated through the use of hyperbolic tangent functions:

$$h_i = \tanh\left(\sum_j w_{ij} x_j + \theta_i \right) \tag{8.5}$$

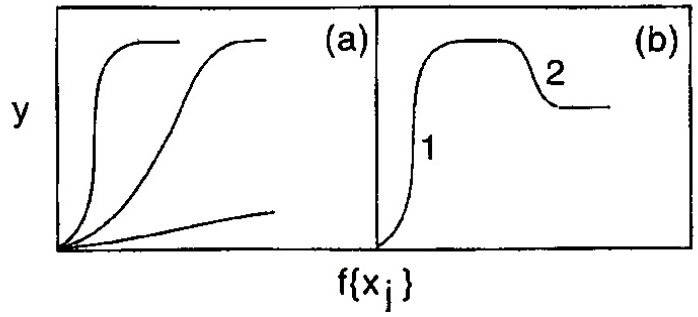


Figure 8.1(a) shows how flexible a hyperbolic tangent can be, as the weights are adjusted, whereas (b) shows two hyperbolic tangents can be added together to make a more complex function to address the non-linearity of the situation [40].

The integration of equation 8.5 into 8.4, together with the minimum and maximum values for the inputs, define the network. Non-linear interactions between the inputs and outputs can be encapsulated by adjusting the weights, as in figure 8.1(a).

Figure 8.2 shows a simple three layer neural network. It illustrates that inputs and outputs are connected through "hidden units". These directly relate to the number of hyperbolic tangents used and hence determine the neural network complexity (figure 8.1(b)).

As previously mentioned, ordinary linear regression does not account for prediction uncertainty. To solve this problem, neural networks can be trained within a Bayesian framework, developed by MacKay [42]. The method recognises that there are many functions that can fit plotted data, so a probability distribution of sets of weights is used. This means that predictions have large error bars if there are few data, as a multitude of solutions may be equally probable. Hence extrapolation can be much safer.

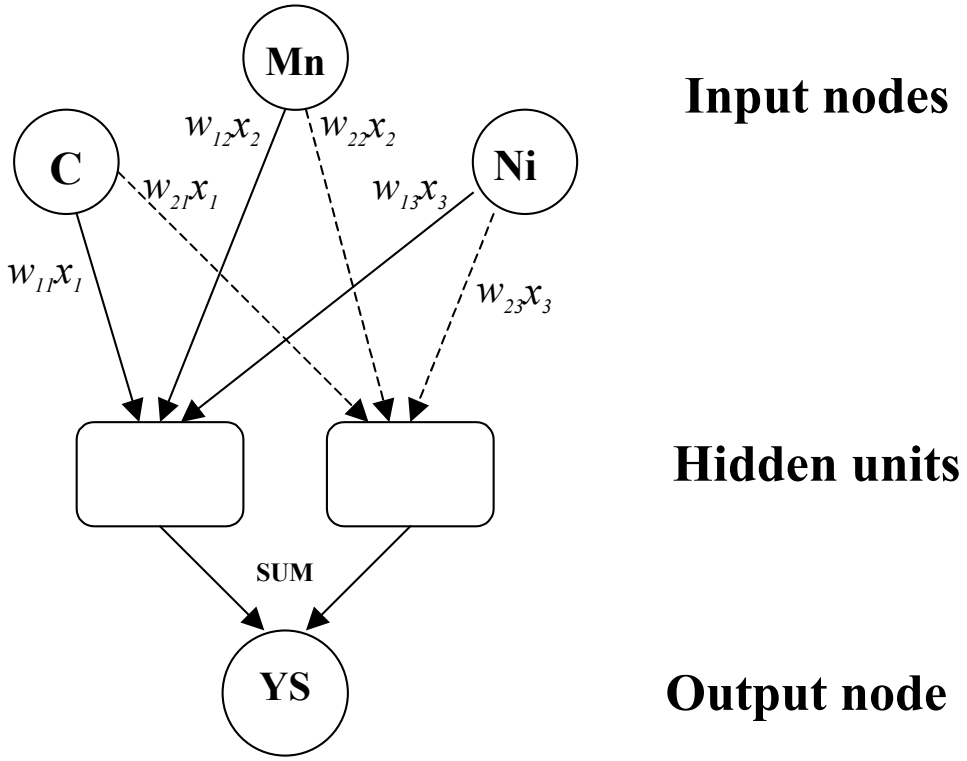


Figure 8.2 – An illustration of a three-layer neural network; consisting of inputs and outputs connected by hidden units that determine the complexity of the model. All the hyperbolic tangents are added to produce an output. The notation expressing the connection between each input and hidden unit is defined by equation 8.5 [40].

For a given model, it is possible to overfit the data by producing an overly complicated function, as shown in figure 8.3. To combat this problem, the data are separated into two random groups. The first set is used to train the neural network to produce models of varying complexity. Generally, as the model becomes more complicated, the associated error decreases. The second set of data, hidden from the initial training, are then used to assess how well the model generalises the unseen data. This analysis helps optimise the complexity of the neural network function.

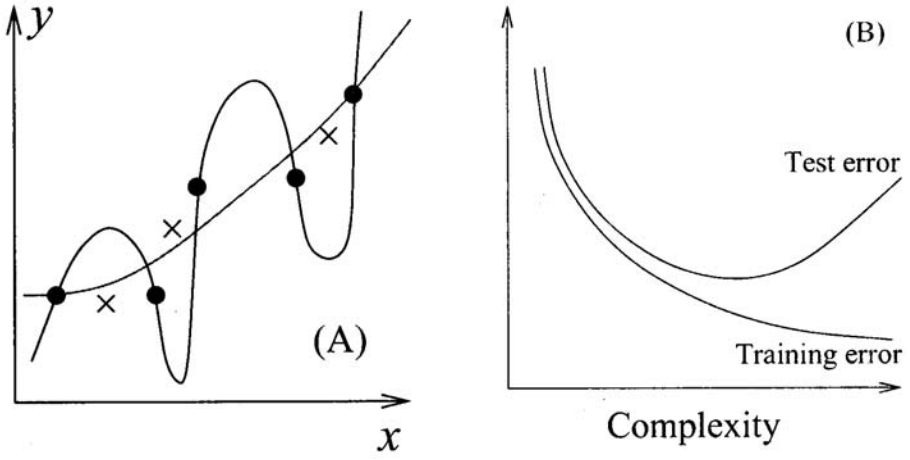


Figure 8.3: The first diagram shows plotted points on an x-y axis. The neural network has to decide whether a line of best fit is complex enough, or a line through all the lines is more appropriate. The second diagram shows that the training database will gradually set the function go through all the data points, hence reducing the overall error. However, the testing database can detect when the function is "overfitting", when it goes beyond the minimum on the test error curve [40].

Once a model has been trained, many models with varying complexity are produced. To calculate the error associated with each model, a regularisation function (M_w) is used:

$$M_w = \beta E_D + \alpha E_w \tag{8.6}$$

where α and β are parameters control model complexity. E_D is the test error; the difference between predicted and target values:

$$E_D = \frac{1}{2}\sum_i (t^i - y^i)^2 \tag{8.7}$$

where t is the target, y is the predicted value and i is the number of collective inputs and outputs involved. Also, to penalise heavily weighted models, the parameter E_w is used to limit overfitting:

$$E_w = \frac{1}{2}\sum_i w_i^{\,2} \tag{8.8}$$

The α and β parameters, as stated in equation 8.6, define the assumed weight variances (equation 8.9) and Gaussian noise (equation 8.10) respectively:

$$\sigma_\omega^{\,2} = \frac{1}{\alpha} \tag{8.9}$$

$$\sigma_v^{\,2} = \frac{1}{\beta} \tag{8.10}$$

where σ_v is the inferred noise as defined by the model.

α encourages the weights to decay in equation 8.6, so that simpler models are preferred to explain output variation, in accordance with Occam's Razor [43]. A low α value results in a large σ_w value, and is therefore a good measure of the significance of each input.

Models are usually ranked by minimum test error. Sometimes it may be more appropriate to evaluate the performance of a model by using the log predictive error (LPE). This error penalises large test errors, but compensates if the prediction has large error bars:

$$LPE = \sum_m \left[\frac{\frac{1}{2}\left(t^{(m)} - y^{(m)}\right)^2}{\left(\sigma_y^{\,(m)}\right)^2} + \log\left(\sqrt{2\pi}\sigma_y^{\,(m)}\right) \right] \tag{8.11}$$

where $\sigma_y^{(m)}$ is the fitting uncertainty for a given set of inputs $x^{(m)}$.

It may be reasonable to use a single model with the lowest test error. However, a group or "committee" of models used together could reduce the overall test error and allow more reliable predictions (figure 8.4) by minimising an error function. The average prediction of a committee p is:

$$p = \frac{1}{L}\sum_{n} y^{(n)} \tag{8.12}$$

where L is the number of models. The standard deviation error (σ) of p is as follows:

$$\sigma^2 = \frac{1}{L}\sum_{n} \sigma_y{}^{(n)2} + \frac{1}{L}\sum_{n} (y^{(n)} - p)^2 \tag{8.13}$$

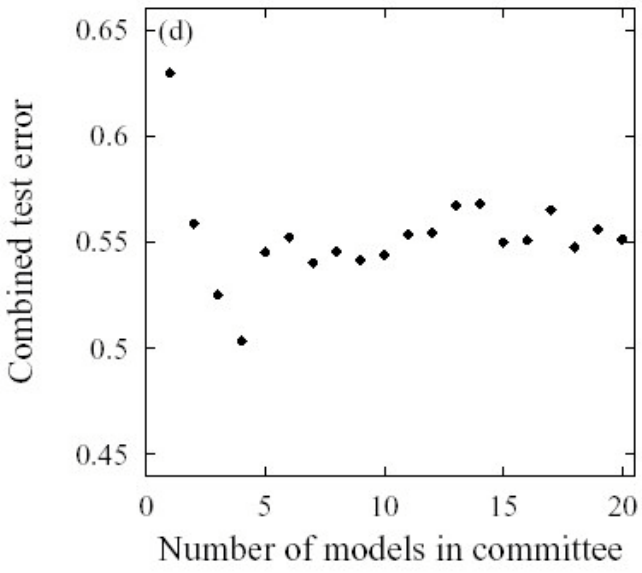


Figure 8.4: Shows how the combined test error goes through a minimum when models are grouped into a committee. In this case, four models is considered to be the optimum [44].

8.1 Database

The aim of this neural network analysis is to predict the yield strength and UTS for austenitic stainless steels as a function of multiple variables. This includes chemical composition, heat treatment temperature and test temperature. The dataset for both the yield strength and UTS are the same, except that the targets change for each case. The database was created based on published data, including the British Steelmakers Creep Committee (BSCC) [45], 16B, 33A and 38A datasheets from the National Research Institute for Metals (NRIM, Japan). These include the stainless steel types AISI 304, AISI 316, AISI 321 and AISI 347.

In total, 2011 experimental results were compiled to introduce as much diversity to neural network training as possible. A total of 184 results of data were omitted from the database due to missing variables. If there are missing variables, then the lines cannot "remain". For grain size, the majority of the database did not have a grain size measure, hence this input was removed. In 155 cases for sulphur and 171 cases for phosphorus, concentrations were found to be missing or recorded as zero. These elements are found in small quantities in virtually all metals, but their importance is hard to define in the context of a tensile test. Values of 0.021 wt% and 0.013wt% were therefore set for each case of phosphorus and sulphur respectively.

The range, mean and standard deviation of each input and output can be found in table 1. It lists each variable and indicates the range of data. However, the values do not define the range of applicability of the neural network model, as in linear regression analyses. Instead, the previously mentioned Bayesian framework is used to define the trained network applicability through the calculation of error bars. This is because each input may interact with others.

The stoichometric stabilisation ratio for titanium and niobium additions is:

$$Ratio = \frac{\left(\frac{C_{Ti}}{4}\right)+\left(\frac{C_{Nb}}{8}\right)}{C_C + C_N} \tag{8.14}$$

As the elements in equation 8.14 are involved in fine MX carbide and nitride precipitation, they may affect strength. However, each compositional element should still be included as an independent input in order to avoid bias in the analysis. Each input has its own effect on steel strength, but there may be a collective influence where an optimum value may offer superior mechanical properties.

All the variables used within the model were normalised to allow an easy comparison of variables:

$$x_n = \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) - 0.5 \qquad (8.15)$$

where x_n is the normalised value, x is the real value, and x_{min} and x_{max} are the minimum and maximum values of the dataset respectively. Through this operation, the input and output values are normalised between ±0.5. Nevertheless, any predictions made from the model can easily have values outside this region.

Figures 8.5(a) to (r) offer a visual picture of the spread of data for 0.2% proof stress, although it is very similar for UTS. Since the inputs interact with each other, the charts should not be scrutinised too heavily. Generally, inputs such as chromium, carbon, sulphur, phosphorus and test temperature were well studied by the model. However, certain gaps in knowledge were visible from the plots. Nickel concentrations between 15-30 wt% were not in the database. However, this range may not be useful for austenitic stainless steels. Molybdenum between 1-2 wt%, manganese at 0-0.8 wt%, niobium in the range of 0.2-0.6 wt%, including boron at much greater concentrations than the maximum of 0.02 wt% should also be studied.

Variable	Minimum	Maximum	Average	Std. Deviation
Cr	15.90	21.06	17.81	0.99
Ni	8.40	34.45	12.57	5.14
Mo	0.00	2.91	1.02	1.17
Mn	0.61	1.82	1.46	0.23
Si	0.00	1.15	0.50	0.14
Nb	0.00	0.95	0.10	0.26
Ti	0.00	0.56	0.15	0.20
V	0.00	0.06	0.0022	0.01
Cu	0.00	0.35	0.0424	0.0812
N	0.00	0.08	0.01	0.01
C	0.01	0.12	0.06	0.01
B	0.00	0.02	0.0004	0.0012
P	0.00	0.04	0.02	0.01
S	0.00	0.05	0.01	0.01
Co	0.00	0.54	0.04	0.1
Al	0.00	0.52	0.03	0.1
Addition ratio for Nb and Ti	0.00	3.0625	0.7737	0.9268
Heat Treatment Temperature / K	1279	1473	1360.83	35.77
Test Temperature / K	293	1273	674.42	232.98
0.2% Proof Stress / MPa	35	341.27	158.49	46.5
Ultimate Tensile Strength / MPa	47	713.92	433.39	93.63

Table 8.1: Analysis of the dataset

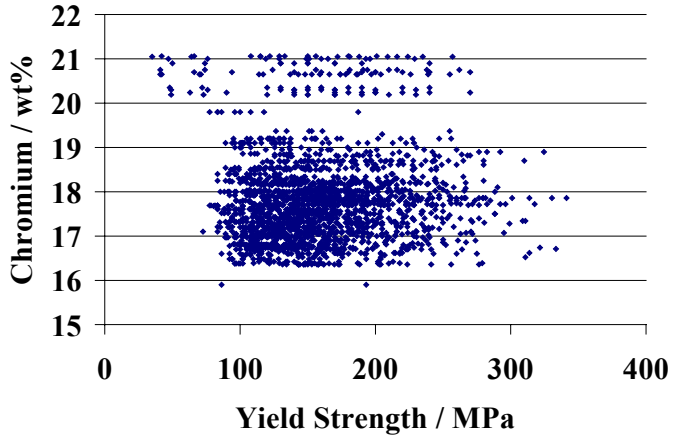


Figure 8.5a: Distribution of chromium data

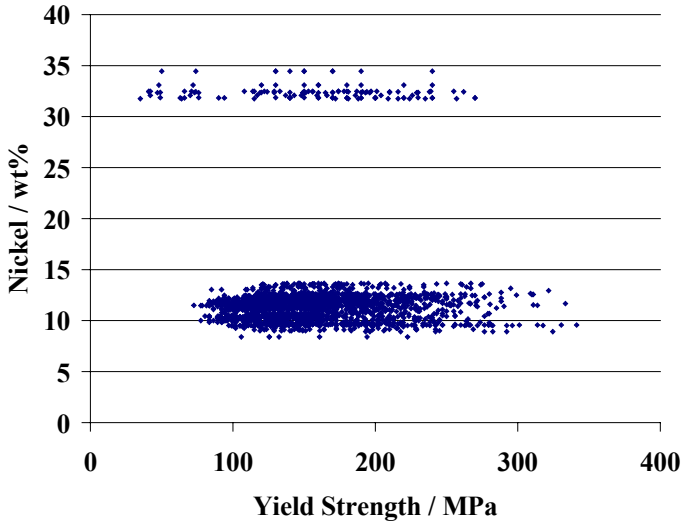


Figure 8.5b: Distribution of nickel data

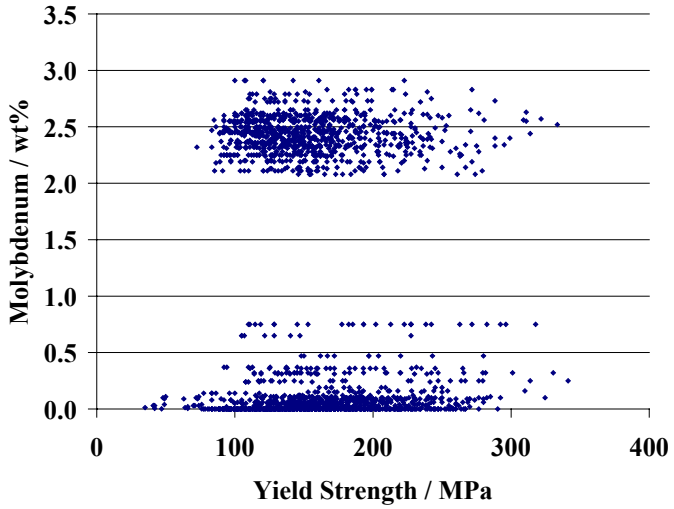


Figure 8.5c: Distribution of molybdenum data

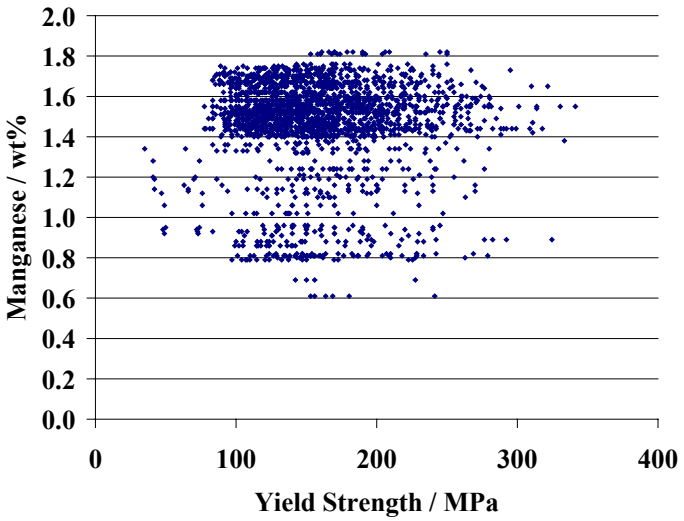


Figure 8.5d: Distribution of manganese data

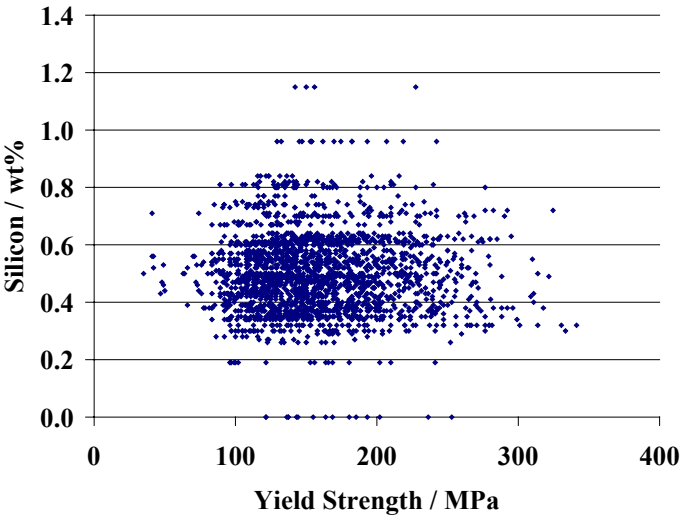


Figure 8.5e: Distribution of silicon data

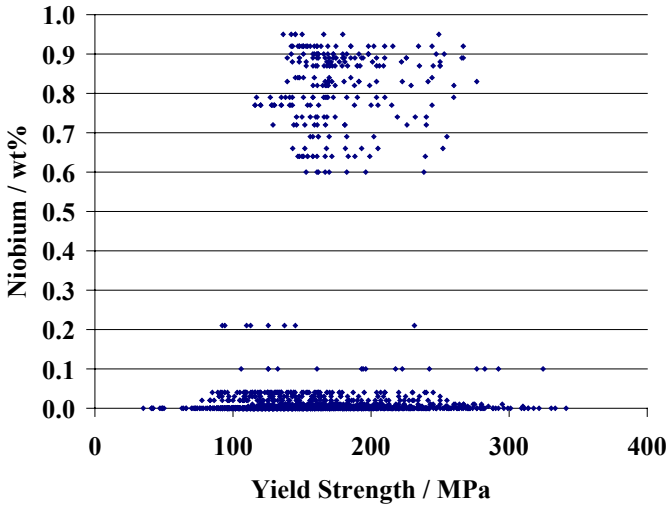


Figure 8.5f: Distribution of niobium data

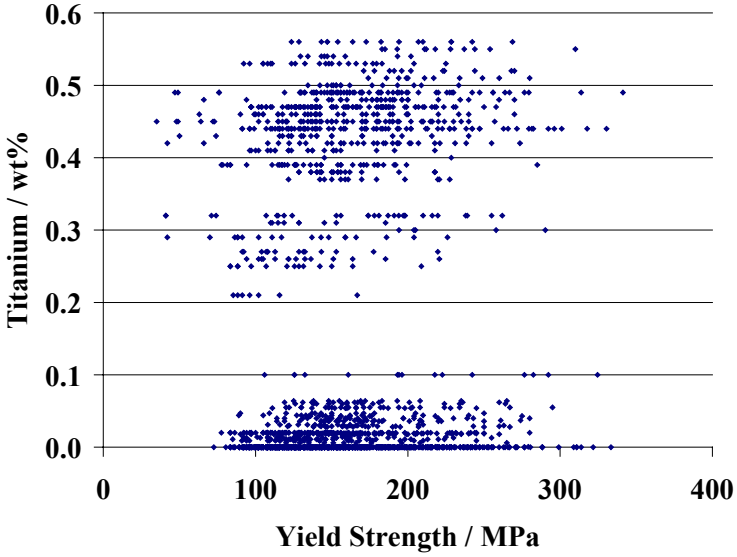


Figure 8.5g: Distribution of titanium data

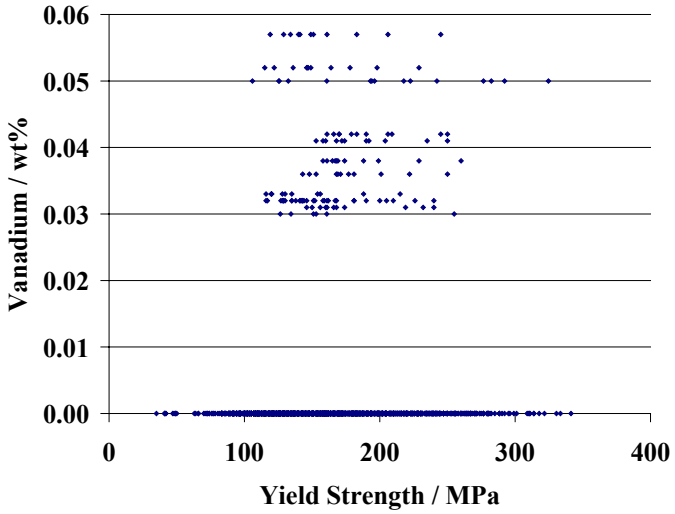


Figure 8.5h: Distribution of vanadium data

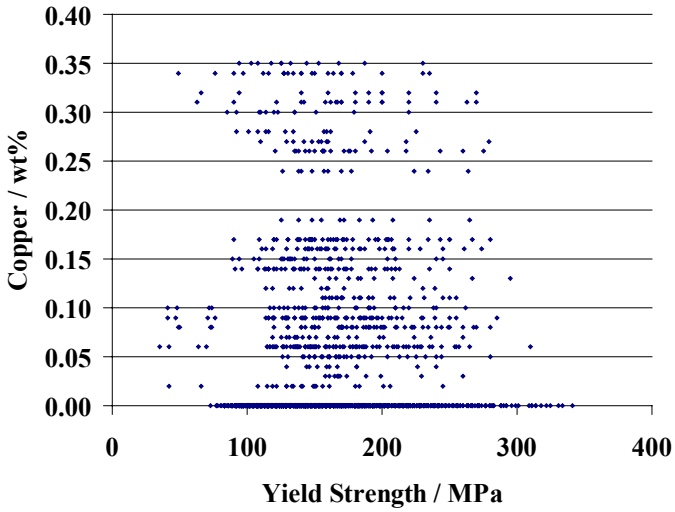


Figure 8.5i: Distribution of copper data

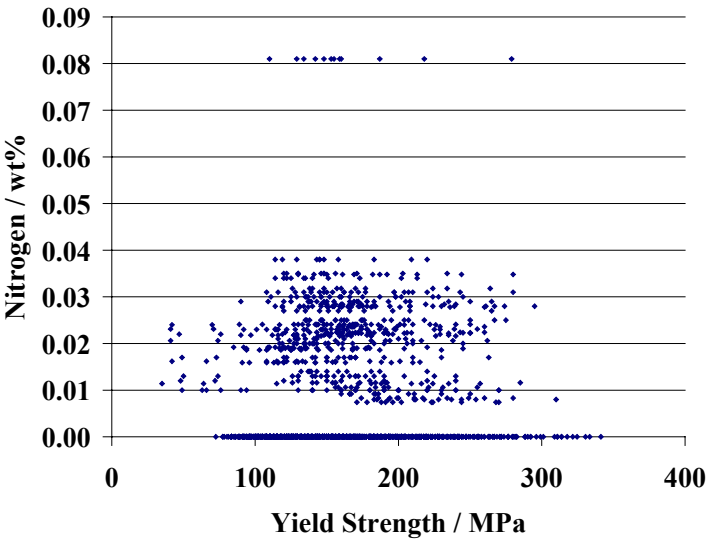


Figure 8.5j: Distribution of nitrogen data

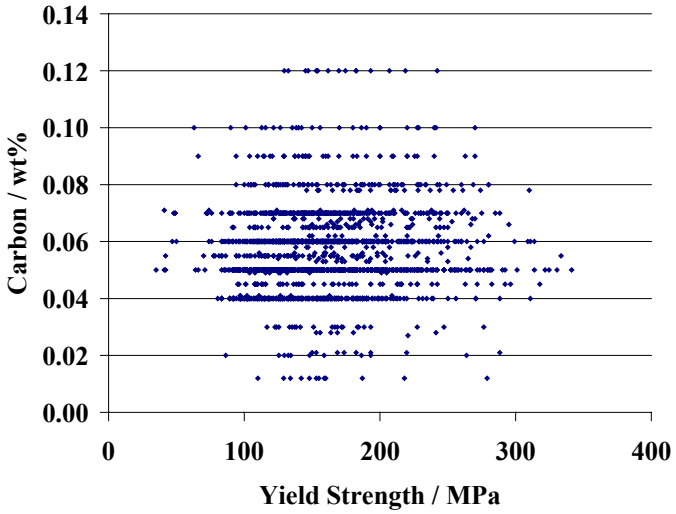


Figure 8.5k: Distribution of carbon data

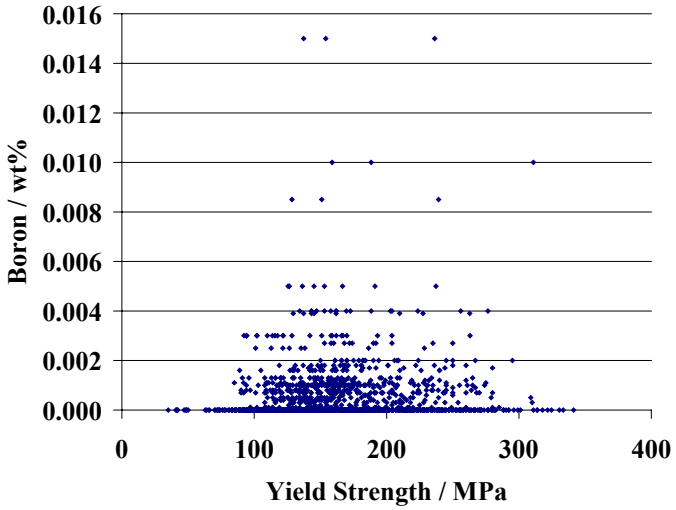


Figure 8.5l: Distribution of boron data

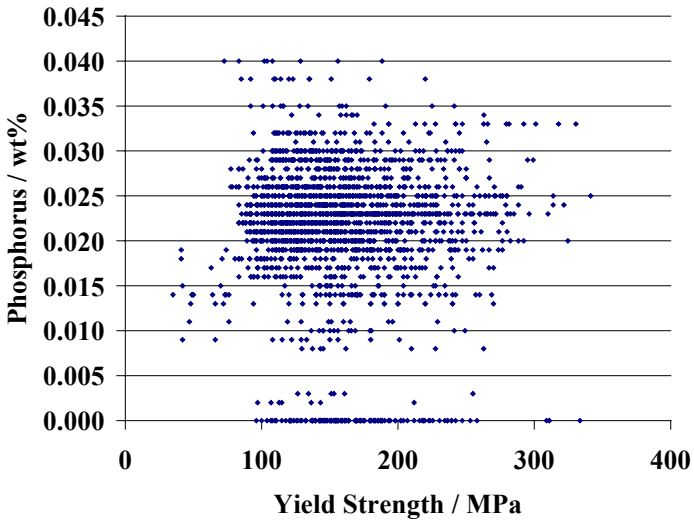


Figure 5m: Distribution of phosphorus data

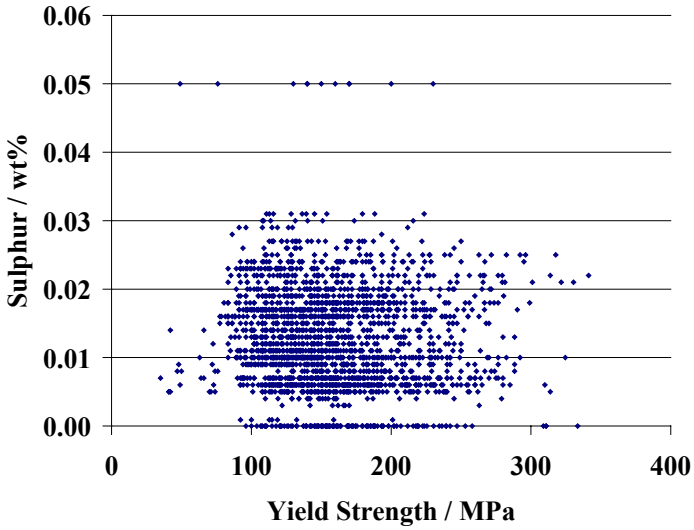


Figure 5n: Distribution of sulphur data

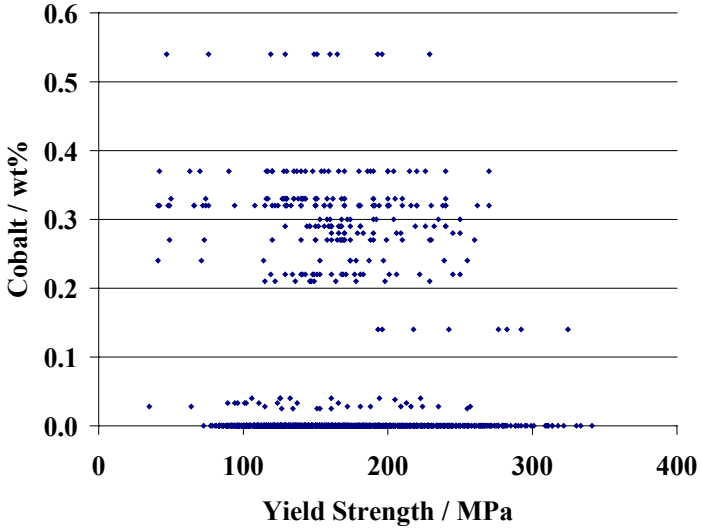


Figure 5o: Distribution of cobalt data

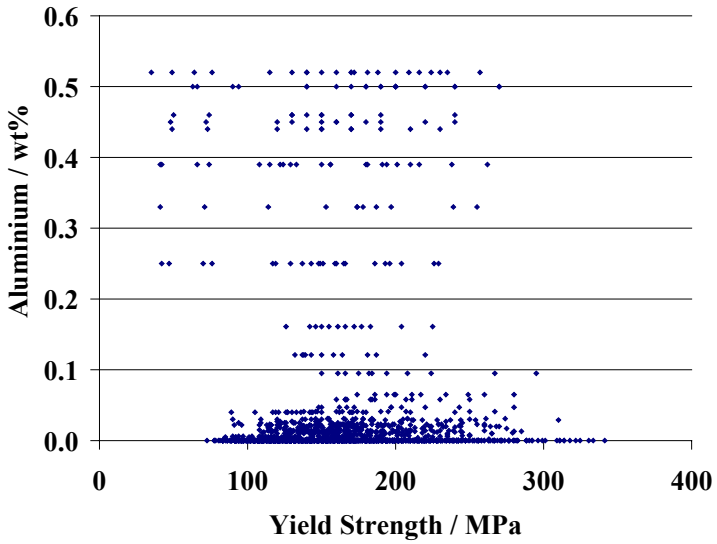


Figure 8.5p: Distribution of aluminium data

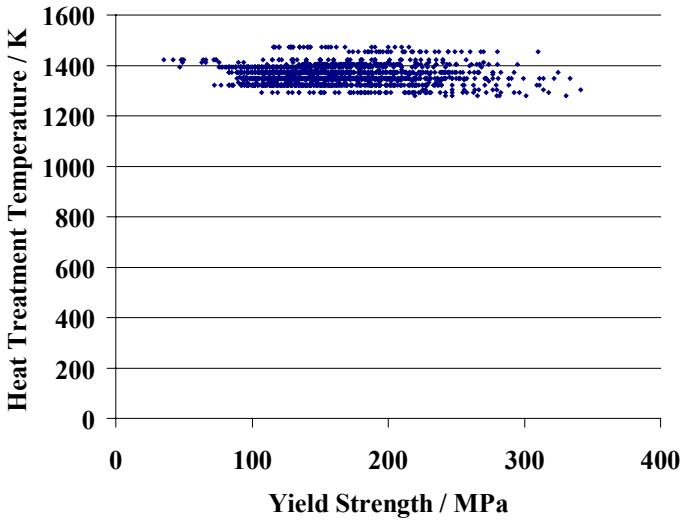


Figure 8.5q: Distribution of heat treatment temperature data

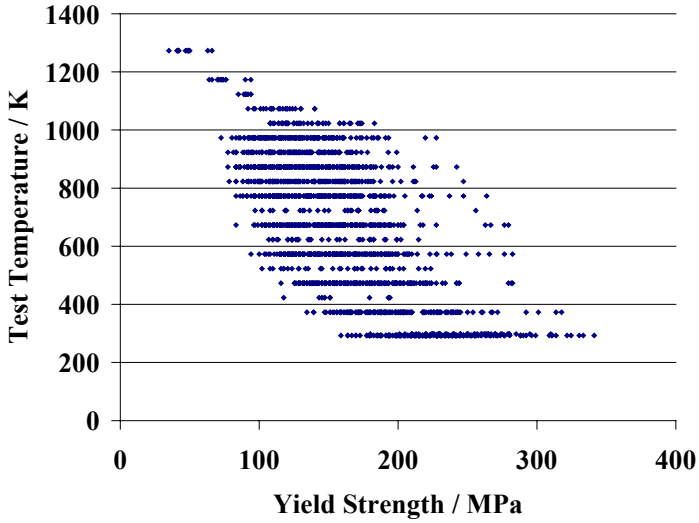


Figure 8.5r: Distribution of test temperature data

8.2 Randomising the dataset

The conventional method for randomisation is to take out every other line from the training set, and place it in the test set. However, each composition had many test temperature results that were grouped together. This meant that the dataset may be unfairly split using this method.

Because of this, a different method was used whereby a set of four compositions were distributed at a time. In each group, every third line would be moved to the test set. This would continue until the fourth composition group was reached, which was totally transferred into the test set.

8.3 Yield Strength Model

Around 200 neural network models were trained using a randomly selected training dataset, which accounted for half the total database. The remaining data later became part of the testing dataset to see how the model generalised the unseen data. Each generated model had a different number of hidden units or random starting-seed, but the same source of input variables, as in table 8.1. The results for this can be found in figure 8.6, which include the sigma noise, test error and LPE.

Figure 8.6a shows that the perceived noise level of the output decreases as the number of hidden units involved increase. This is consistent with expectation. The test error was smallest when it involved 6 hidden units (fig. 8.6b), or a maximum of 4 hidden units with the LPE (fig. 8.6c).

Figures 8.6e and 8.6f show how well the best model generalised the test set. In the latter figure, there are a number of outliers in the test dataset plot for predicted vs. measured yield strength. However, these were seen to be unique points not found within the training set.

Note that any error bars plotted in these figures include $\pm 1\sigma$ and fitting uncertainty, as calculated by the Bayesian framework.

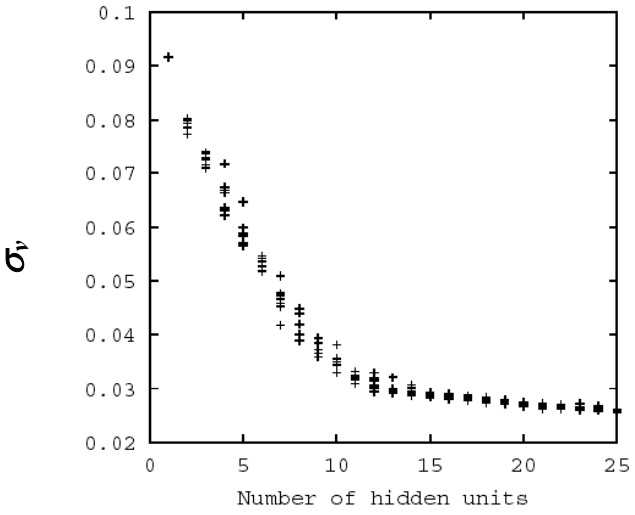


Figure 8.6a: The perceived level of noise σ_ν vs hidden units

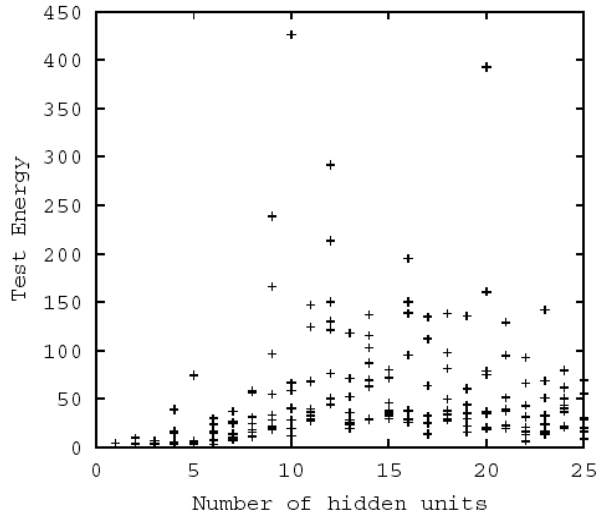


Figure 8.6b: Test error vs hidden units

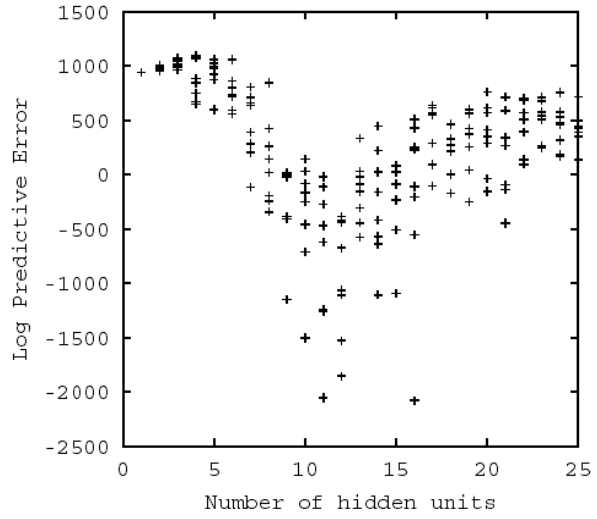


Figure 8.6c: Log predictive error vs hidden units

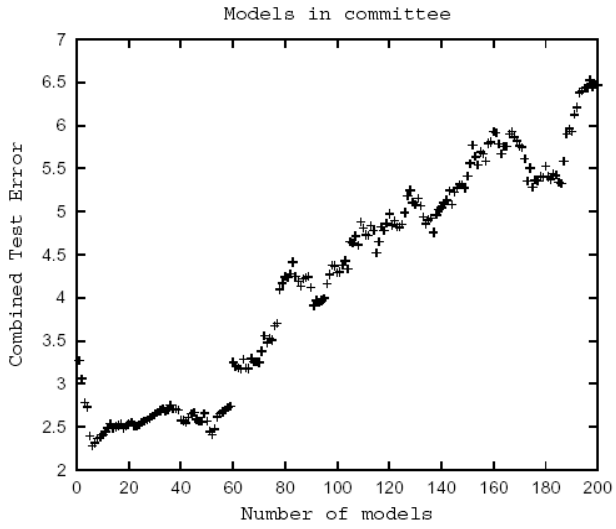


Figure 8.6d: Test error vs models in committee

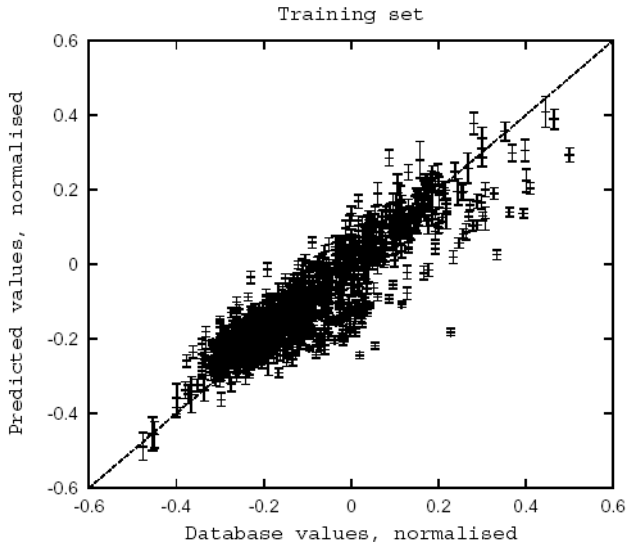


Figure 8.6e: Predicted vs measured yield strength (training set)

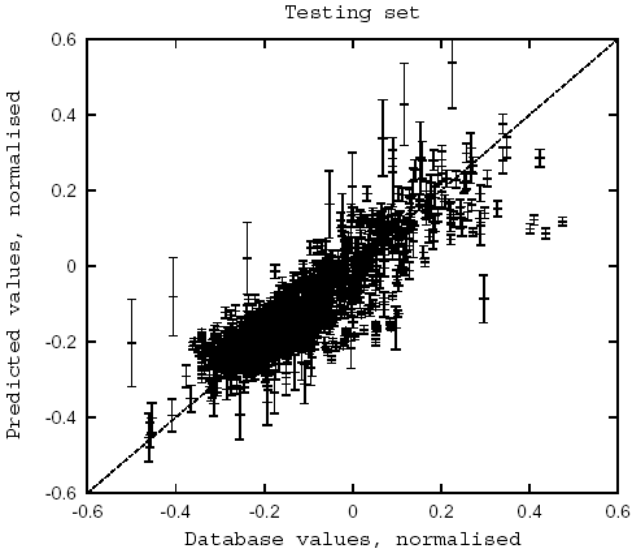


Figure 8.6f: Predicted vs measured yield strength (test set)

As mentioned previously, a committee of models may allow predictions to be more reliable by lowering the test error. This is certainly the case here, where the error of the best single model is 3.236 with four hidden units, but the committee is 2.277 (figure 8.6d). This combined test error was optimised by combining 6 models together. After this selection process, the model committee was fixed and subsequently re-trained on the whole database. A final comparison of predicted and measured yield strength is shown in figure 8.10a. It can be seen that the plot has fewer outliers.

The significance, σ_w, is a measure of how much the inputs influence the output. Each variable was investigated for the best 4 models, and presented in figure 8.7. The test temperature is seen to have a large influence on the output, which is consistent with metallurgical theory. However, each input was seen to offer at least a moderate contribution to the output. This therefore confirmed that they were a good choice of inputs.

In summary, it was important to find out which variables are considered to be most significant, or those that contributed very little to the output. However, it is well understood that many of the variables have some bearing on steel strength. Overall, the aim was to obtain meaningful inputs that allow optimisation of mechanical properties within a predictive framework.

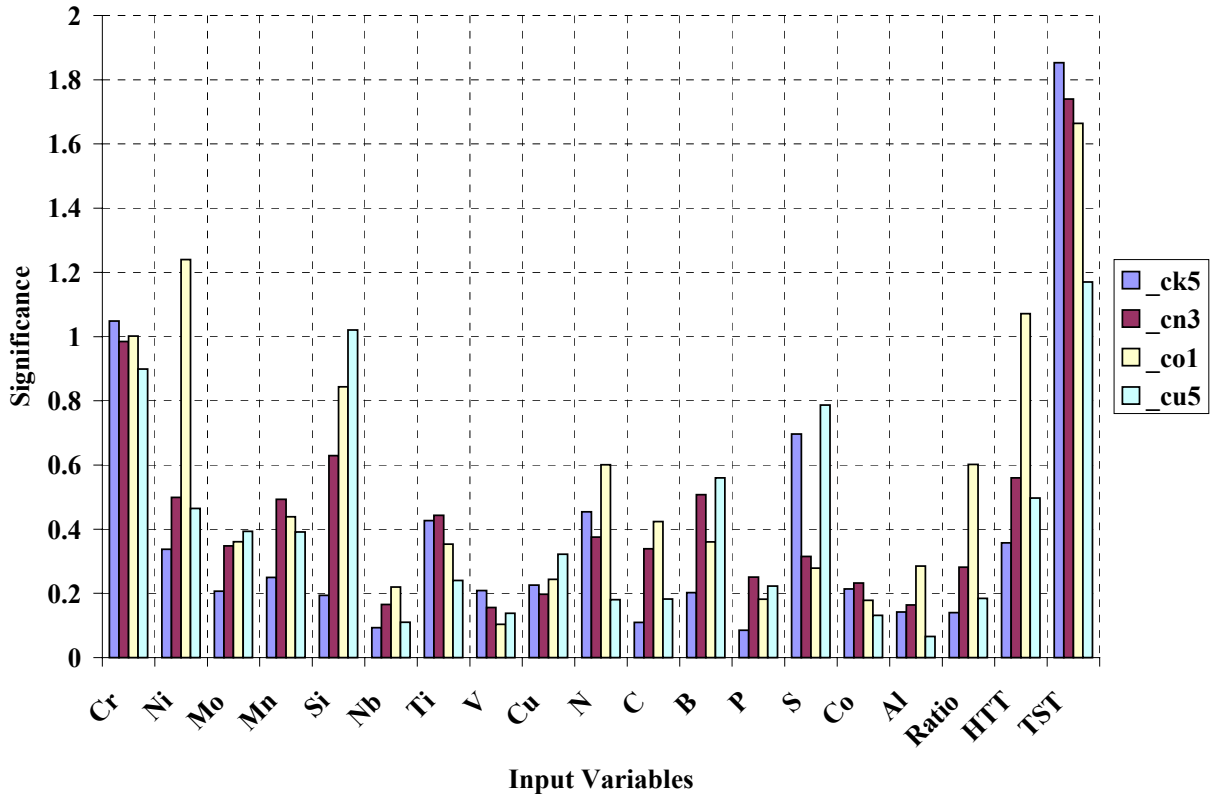


Figure 8.7 – A chart illustrating the perceived significance σ_w for the four best performing models, based on highest LPE. The inputs compromise of elements, ratios and temperature treatments. "Ratio" refers to the stoichometric addition of niobium and titanium, whereas "HTT" and "TST" are heat treatment temperature and test temperature respectively.

8.4 Ultimate Tensile Strength Model

Around 125 neural network models were trained identically to the yield strength model, with the database split in the same manner. The procedures followed for yield strength were also identical for this model, with figure 8.8 showing the results. Figure 8.10b highlights the performance of the final optimised committee of 5 members. The significance σ_w of the best 4 models (fig. 8.9) shows that test temperature is again the most influential variable. Again, the plotted error bars include $\pm 1\sigma$ and the fitting uncertainty.

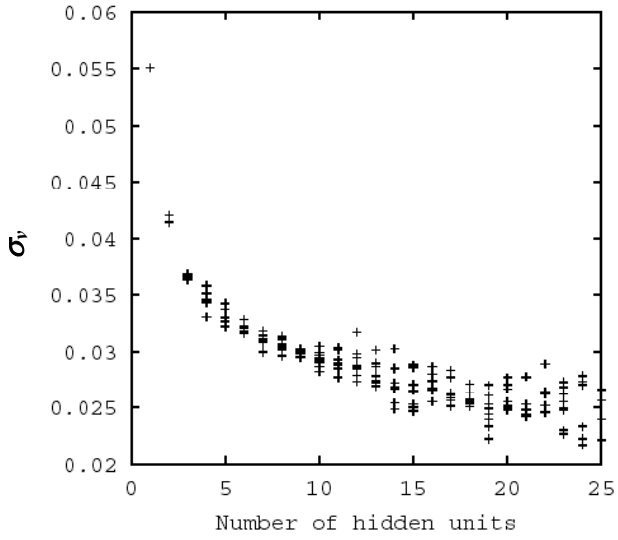


Figure 8.8a: The perceived level of noise σ_ν vs hidden units

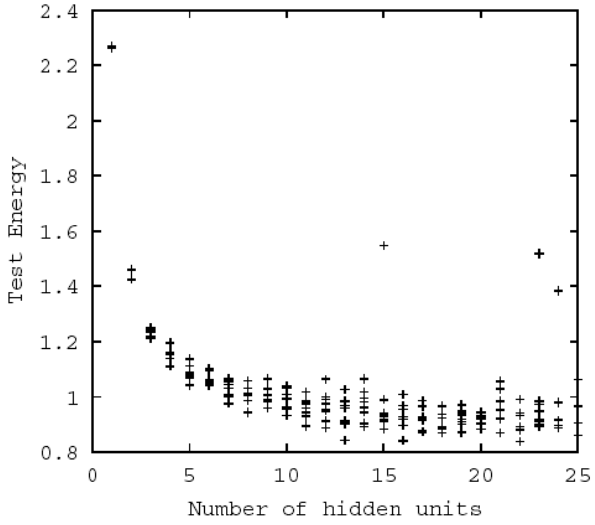


Figure 8.8b: Test error vs hidden units

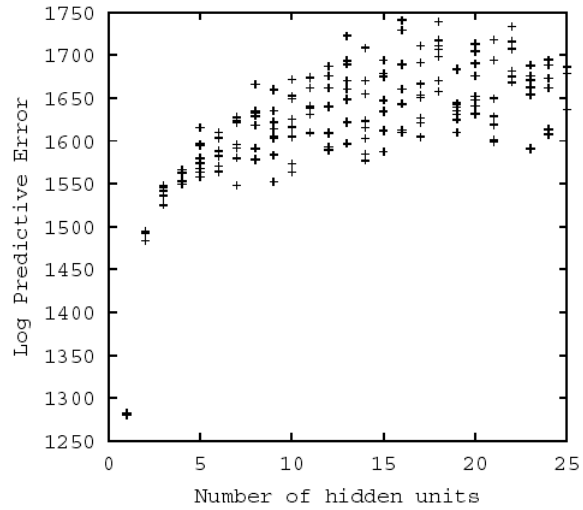


Figure 8.8c: Log predictive error vs hidden units

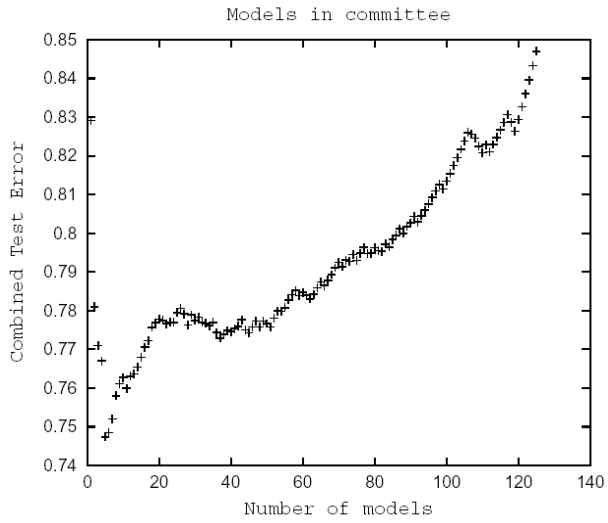


Figure 8.8d: Test error vs models in committee

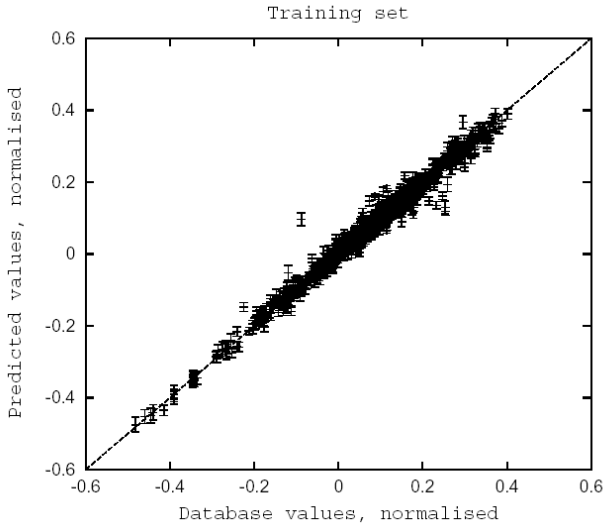


Figure 8.8e: Predicted vs measured UTS (training set)

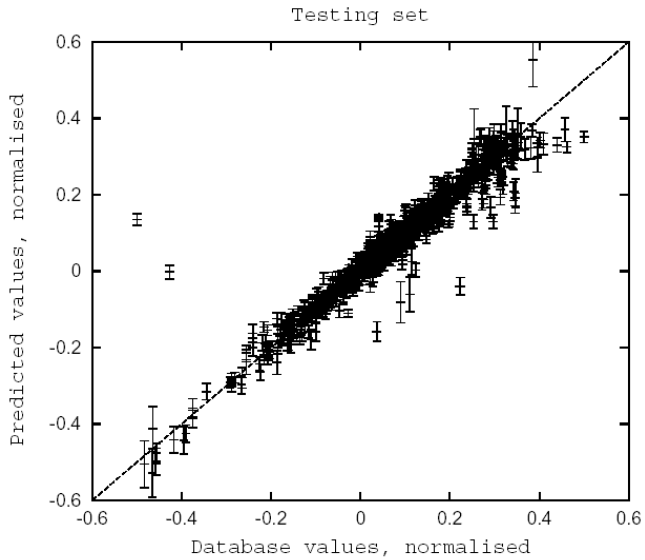


Figure 8.8f: Predicted vs measured UTS (test set)

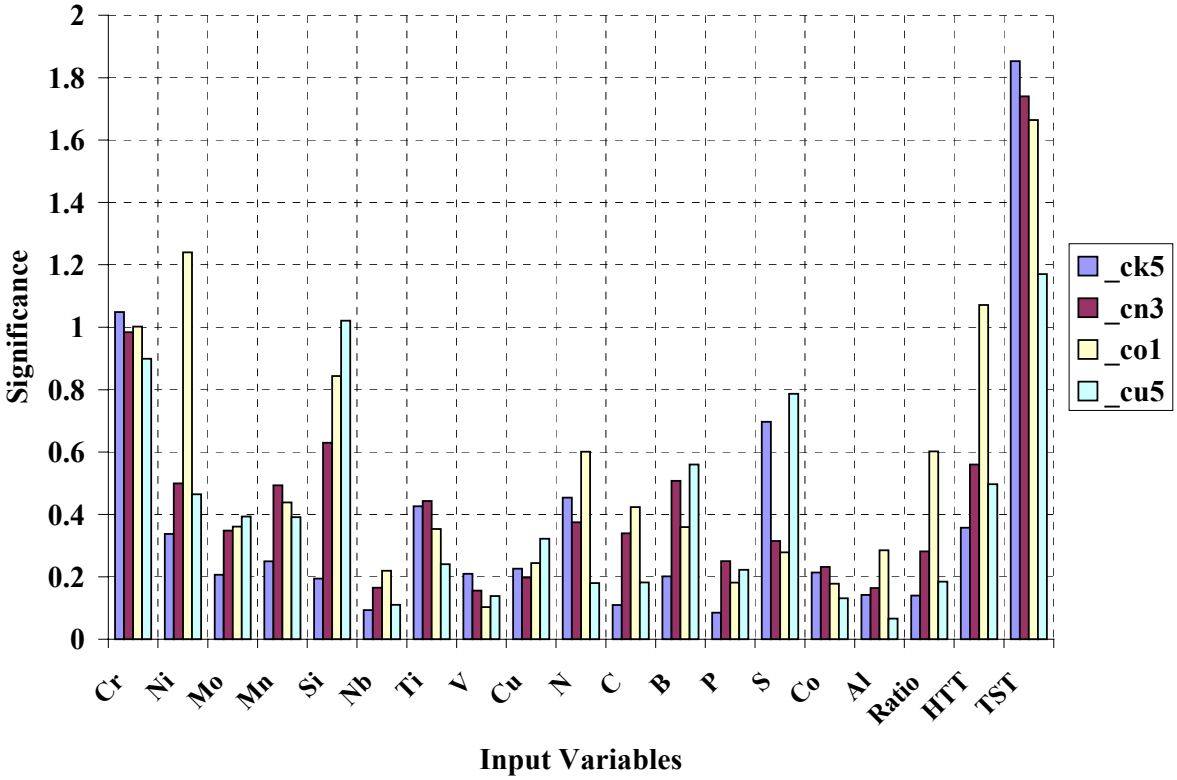


Figure 8.9 – same as figure 8.7, but shows the perceived significance σ_w for the four best performing models for UTS, based on highest LPE.

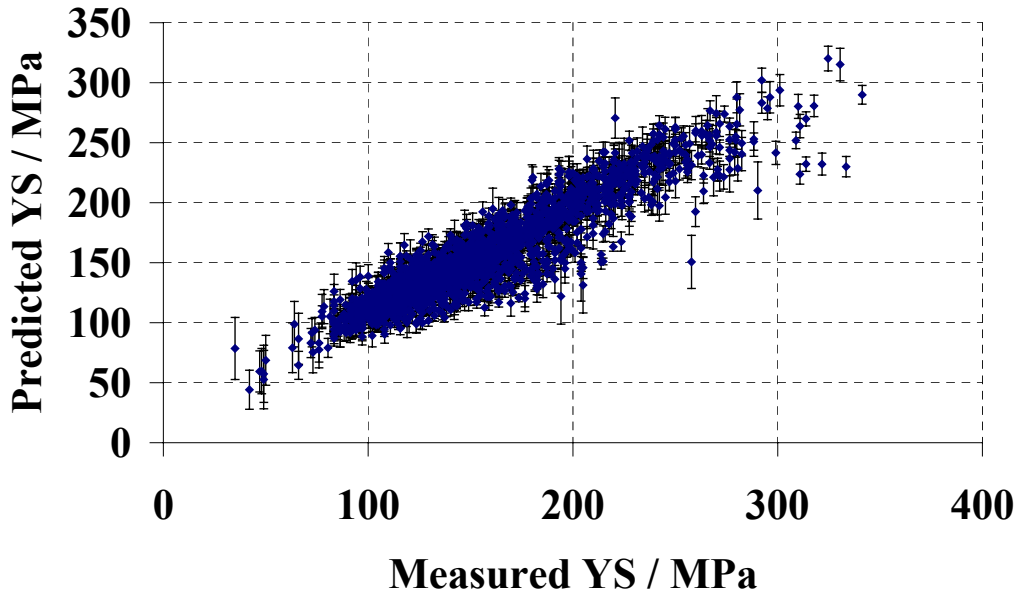


Figure 8.10a – Final committee model of predicted vs. measured yield strength.

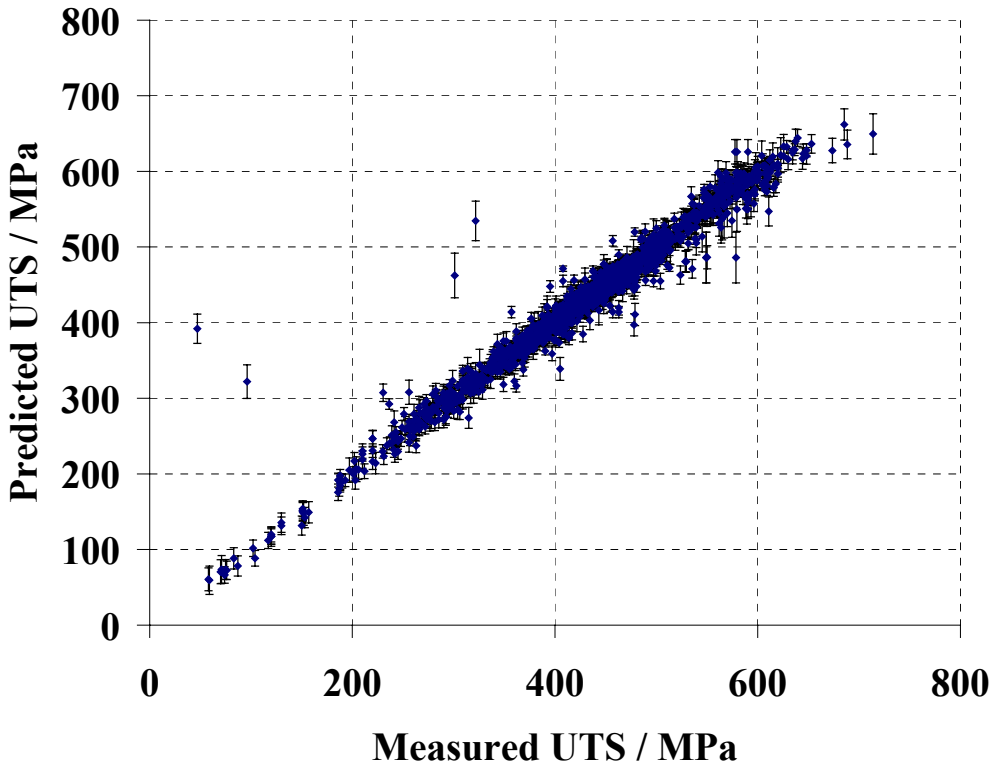


Figure 8.10b – Final committee model of predicted vs. measured UTS.

8.5 Applications

Once the models were developed, their behaviour was compared to findings in the literature using compositions in table 2. The aim was to show that they agreed with existing knowledge of stainless steels had enough complexity to describe different relationships. All error bars plotted in this section include $\pm 1\sigma$ and the fitting uncertainty.

8.5.1 Chromium

Predictions were made using composition A, shown in figure 8.11. For a typical type 316 stainless steel, the variations are well within uncertainty. The model indicates a negligible effect of Cr, as shown by Sourmail in another context [44].

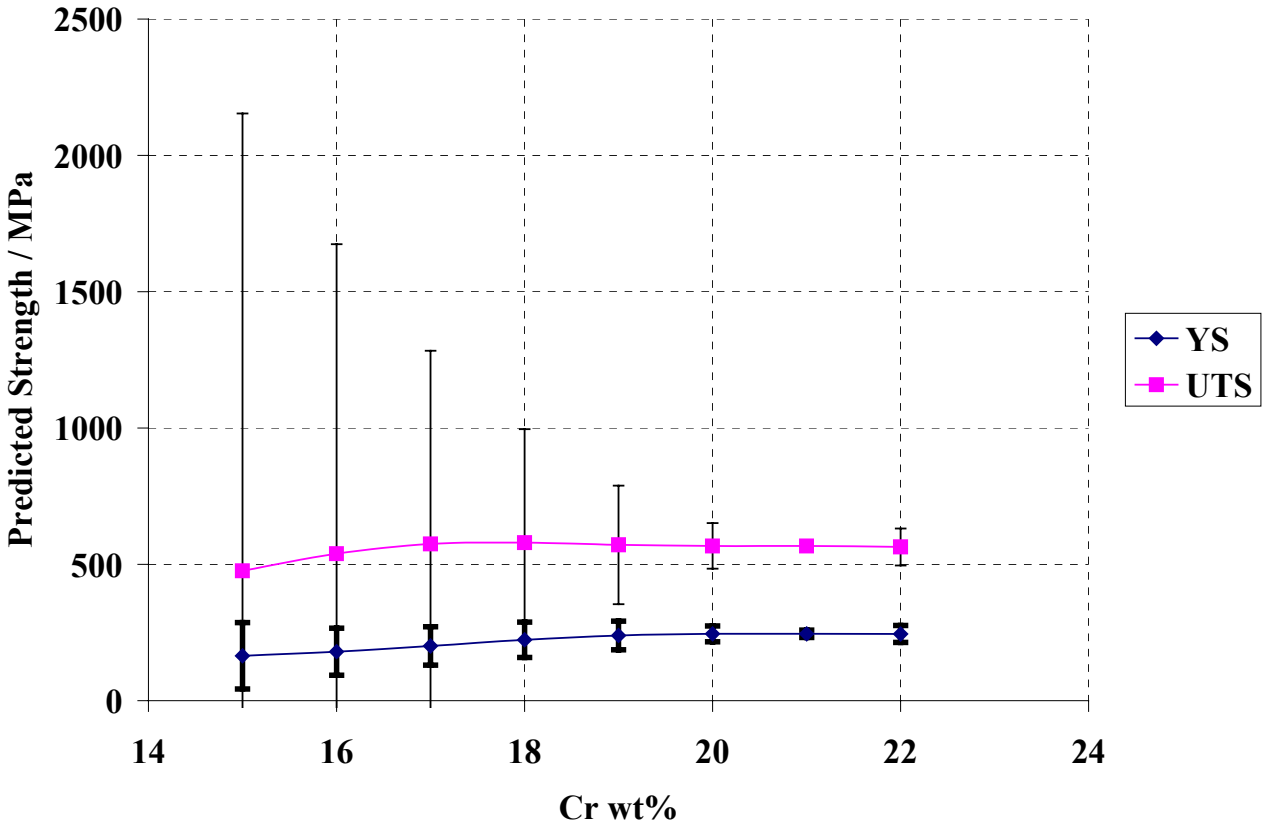


Figure 8.11 – The effect of chromium on the YS and UTS using composition A.

The dataset contained chromium values mainly between 17-19 wt%. This was the region where the most reliable predictions are made, but the large error bars outside this region indicate that more knowledge and data are required. Nevertheless, consistent with literature [2], chromium is found to have little effect on tensile properties.

	Cr / wt%	Ni / wt%	Mo / wt%	Mn / wt%	Si / wt%	Nb / wt%	Ti / wt%	V / wt%	Cu / wt%	N / wt%	C / wt%	B / wt%	P / wt%	S / wt%	Co / wt%	Al / wt%	Ratio	HTT / K	Temp / K
Composition A	-	10.26	0.31	1.58	0.6	0.04	0.04	0	0.12	0.028	0.07	0	0.022	0.012	0	0.014	0.153	1343	298
Composition B	17	12.2	0.5	0.81	-	0	0	0	0	0	0.06	0.0039	0.026	0.017	0	0	0	1373	298
Composition C	17.05	12.6	2.24	1.1	0.7	0.001	0.03	0	0.31	0.017	0.05	0.03	0.034	0.013	0	0.002	0.114	1373	-
Composition D	17.34	12.66	-	1.44	0.42	0	0	0	0	0	0.06	0.005	0.025	0.013	0	0	0	1373	298
Composition E	18.5	10.3	0	2.0	1.0	0	0	0	0	0	0.03	-	0.025	0.013	0	0	0	1373	298
Composition F	18.05	12.13	0.06	1.76	0.49	0.88	0.036	0.08	0	0.0241	0.065	0.0018	0.027	0.019	0.22	0.007	1.336	-	298

Table 8.2: Composition information used for making predictions for both YS and UTS.

8.5.2 Silicon

In the database, there were more values below 0.6 wt% than above (fig. 8.5e). This is reflected in figure 8.12, where the error bars get larger at high silicon concentration. Much more data are needed in this region, where effects such as increased solid solution hardening may occur [46].

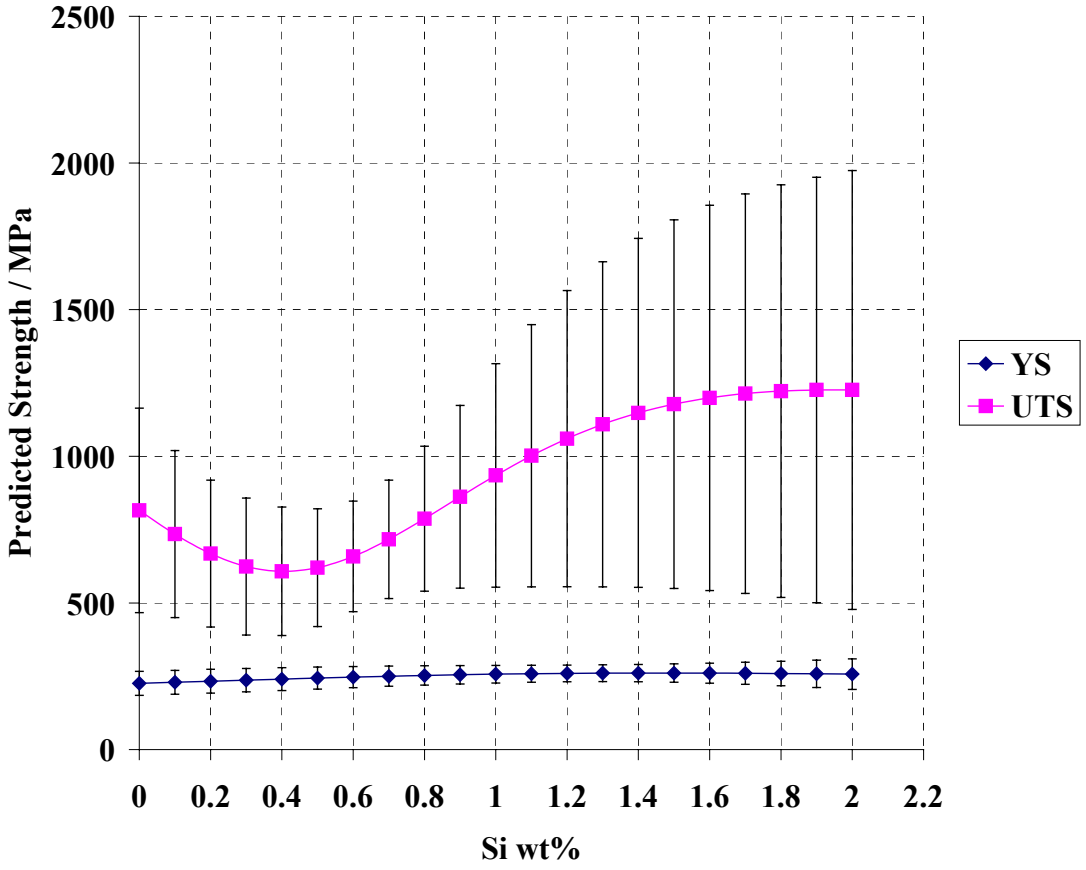


Figure 8.12 – The effect of silicon on the YS and UTS using composition B.

In 1991, Kivineva and Karjalanen [46] conducted work on high silicon austenitic stainless steels. The neural network model had not seen silicon above 1.15 wt%. Nevertheless, a prediction was made with a quoted steel composition to see how the model would perform.

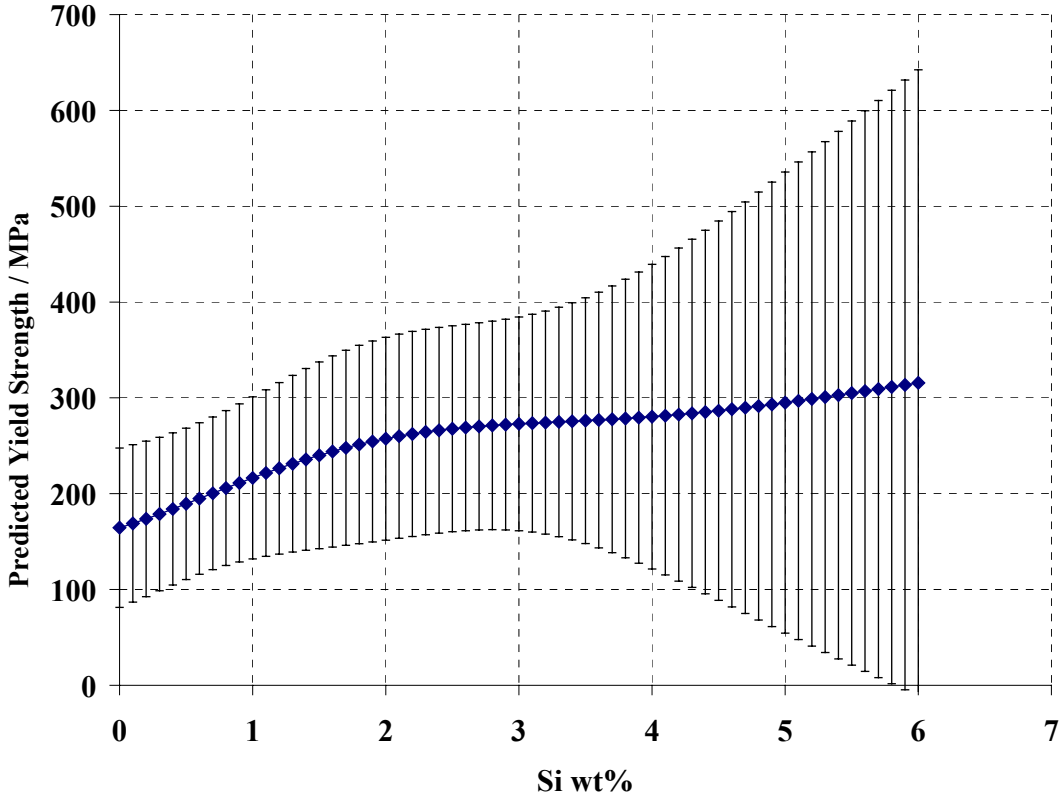


Figure 8.13a – Predicted effect of high silicon concentration on the YS using Fe-17.5Cr-8.9Ni-0.13Mo-1.38Mn-xSi-0.017C-0.024N-0.03P-0.004S-0.19Cu-0.06Al wt% steel.

From both figures 8.13(a) and (b), large error bars can be seen as expected. However, it is interesting to note that the measured value of yield strength at 3 wt% is 240 MPa, compared to 272 MPa by the network. The UTS is more unsure, but predictions in the mid-range of the plot are quite good. For example, at 4.5 wt%, the measured value from [46] is 654 MPa, whereas the predicted value is 676 MPa.

This shows that good predictions are being made, even though large error bars signify great uncertainty.

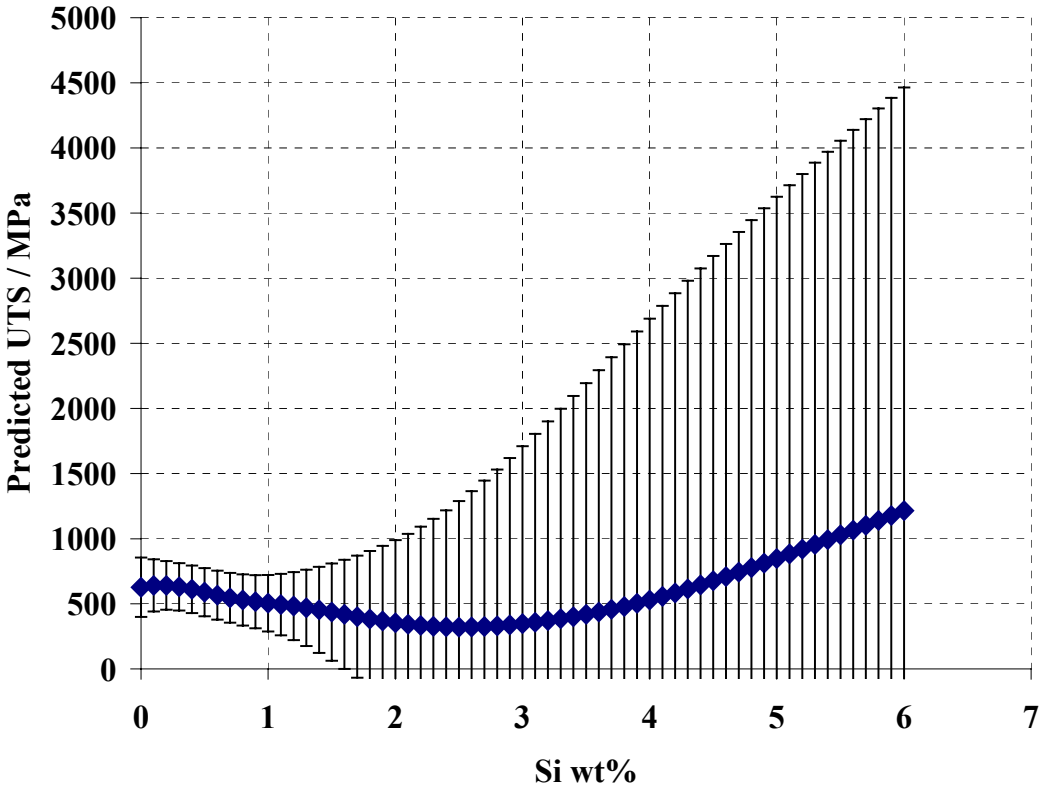


Figure 8.13b – Predicted effect of high silicon concentration on the UTS using Fe-17.5Cr-8.9Ni-0.13Mo-1.38Mn-xSi-0.017C-0.024N-0.03P-0.004S-0.19Cu-0.06Al wt% steel.

8.5.3 Test temperature

It has been shown in the literature that high temperature conditions have many consequences. This includes easier dislocation motion and lower SFE values. Using composition C, figure 8.14 correlates well with this reasoning, showing a strong decrease with increasing test temperature.

Notice that the difference between yield strength (YS) and UTS decreases at high temperatures, which indicates that work hardening occurs less readily, as thermal activation allows barriers to be overcome.

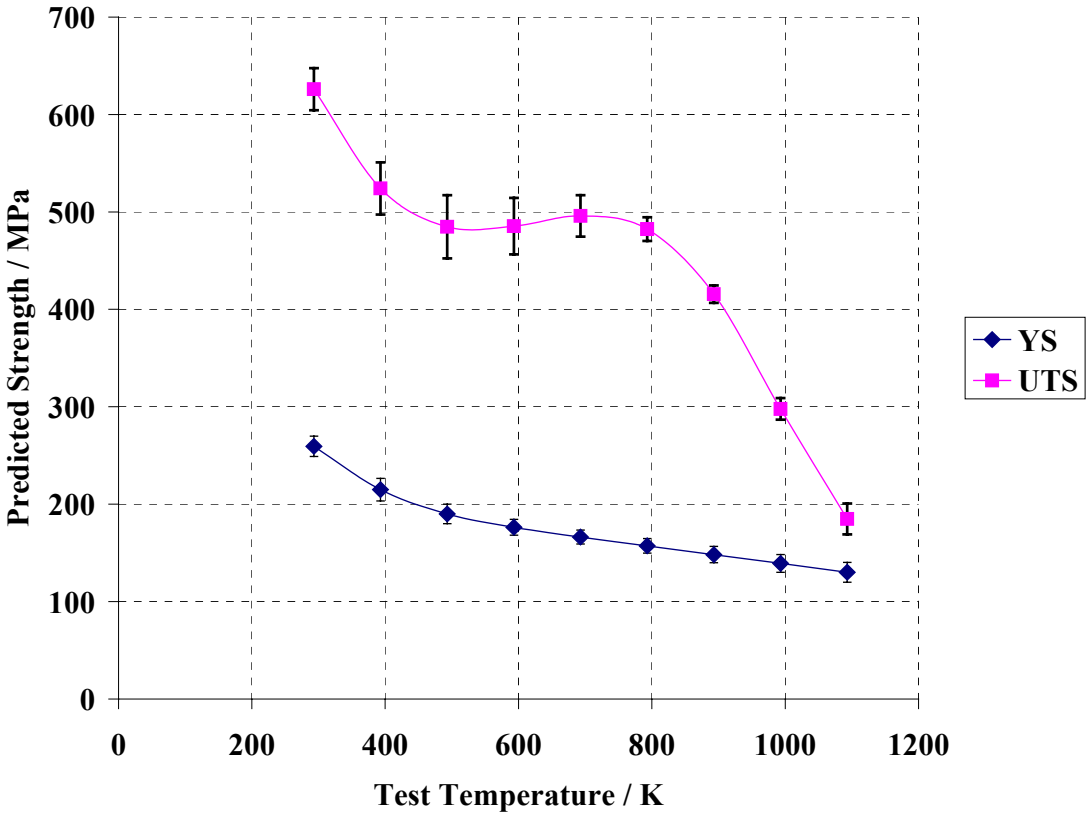


Figure 8.14 – The effect of test temperature on the YS and UTS using composition C.

To show this, compositions A and C were tested, where the former was considered to have a lower SFE due to higher concentrations of manganese and nitrogen (as described in section 3.2). From figure 8.15, it is shown that the (UTS - YS) difference is larger for the low SFE steel. This indicates that low SFE elements included in the models may predict increased work hardening at high temperatures.

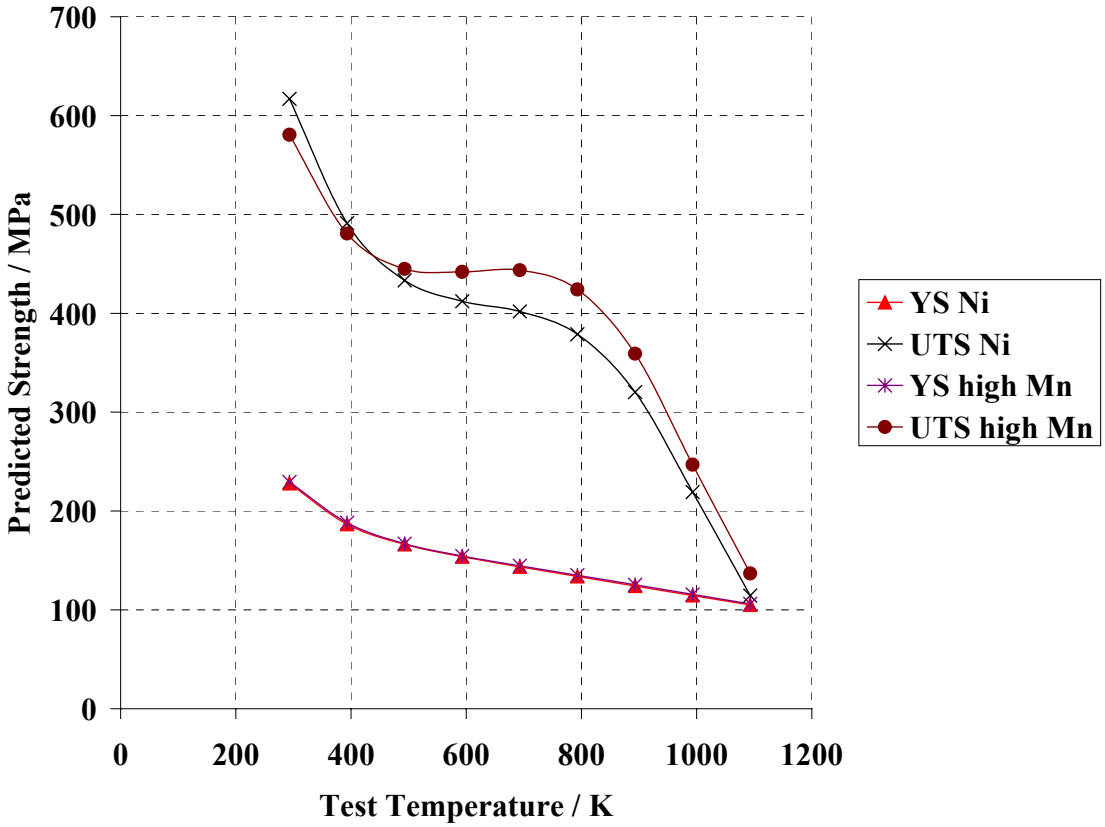


Figure 8.15 – Investigating the effect of SFE at higher temperatures using compositions A and C (error bars have been omitted for clarity).

8.5.4 Heat treatment temperature

This property was also considered, as it was the second most important variable in the model. The aim of heat treatment is to dissolve precipitates and allow *MX* carbide formation upon cooling to ambient temperature. Hence niobium and titanium were included as inputs for predictions, composition F. The UTS in figure 8.16 is predicted to behave differently at various temperatures. However, large error bars at either ends of the curve indicate large uncertainty. The main area of interest lies around 1300-1400 K, where a steady decline is shown. This confirms the belief that increased temperatures cause more carbides to be drawn into solid solution.

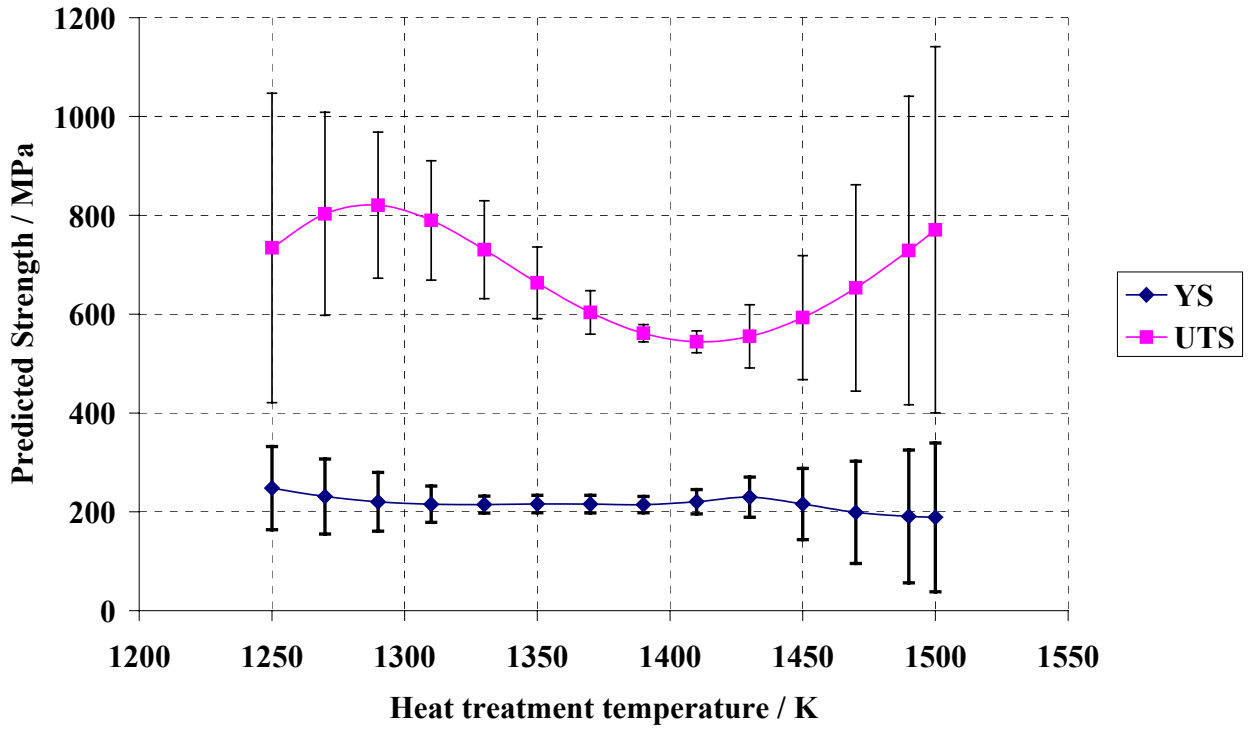


Figure 8.16 - The effect of heat treatment temperature on the YS and UTS using composition F.

8.5.5 Boron

Wilkinson and King [47] reported positive correlation between boron and yield strength with values up to 2 wt%. However, the neural network database was limited by a maximum boron concentration of 0.015 wt%, as well as small amounts of data. Nevertheless, the yield strength was still to found to increase with boron addition using the same composition in [47]; composition E, figure 8.17. The UTS also predicted very well in higher concentrations, even though it was accompanied with expected large error bars. The results show that the model is making good predictions, but the large error bars underline the need for more information to improve model predictability.

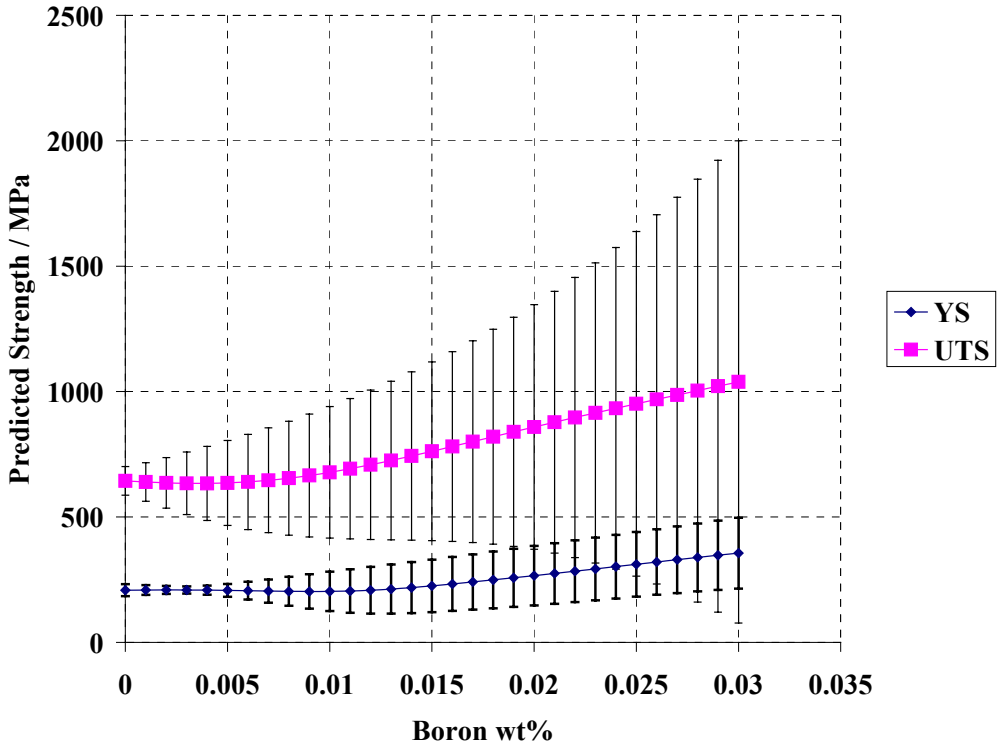
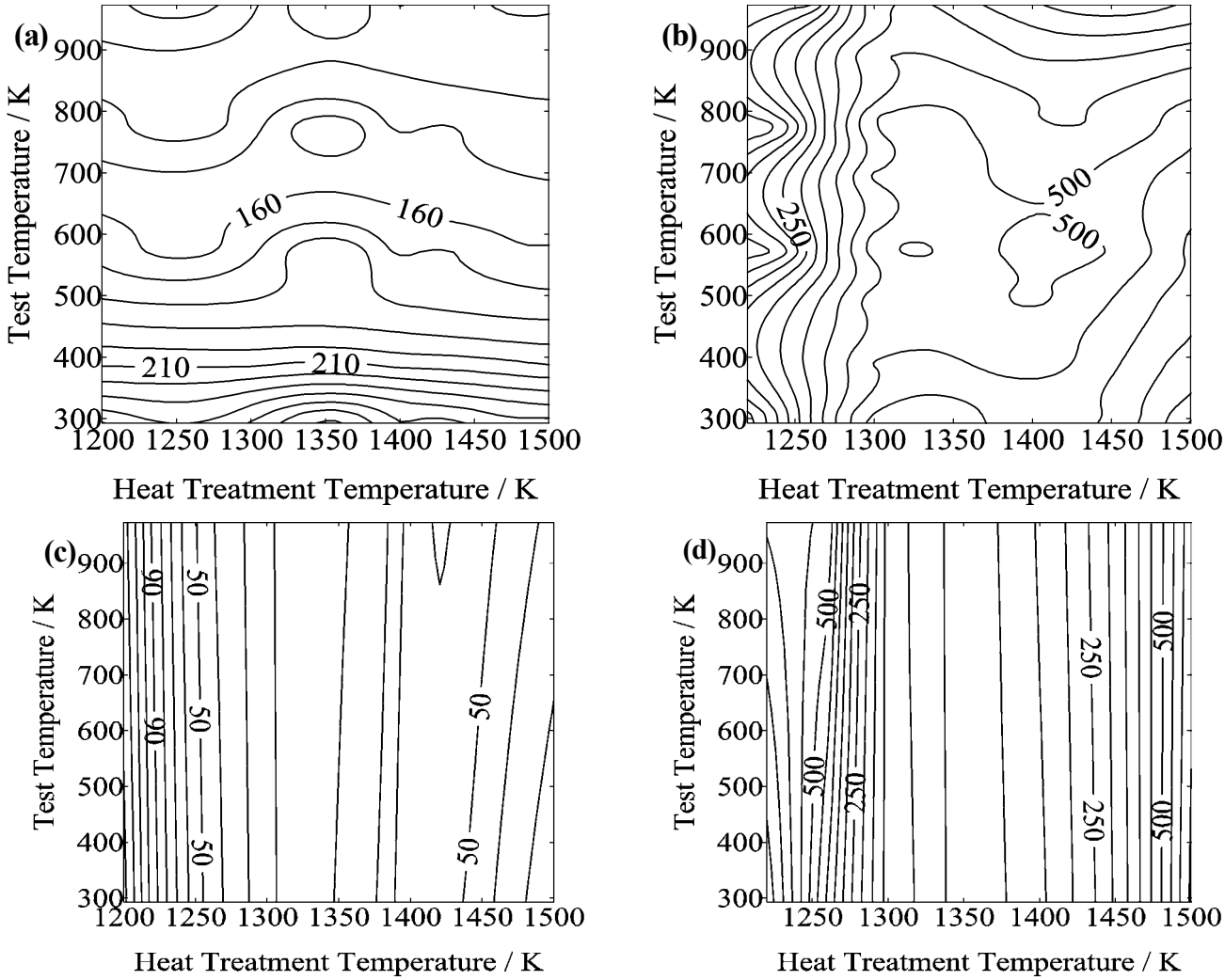


Figure 8.17 – The effect of boron on the YS and UTS using composition E.

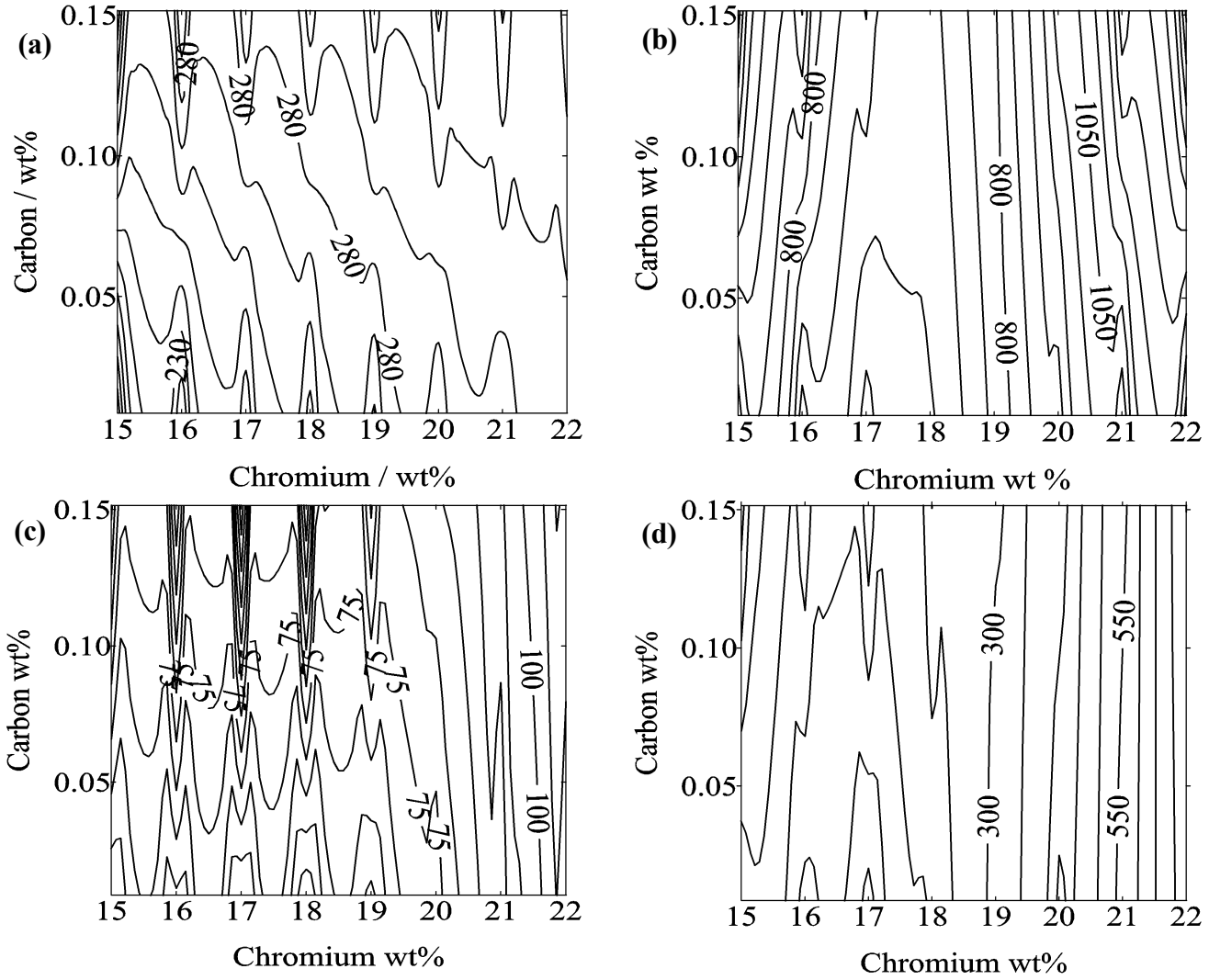
8.5.6 Non-linear effects

It would be interesting to find any non-linear relationships that are not captured by linear regression methods, as with Pickering in equations 8.1 and 8.2. This would show that the flexibility of neural networks may be more suitable for making predictions. To illustrate this, contour plots of YS and UTS were made for heat treatment temperature against test temperature (fig. 8.18). As curves are found for these plots, it is shown that neural networks can capture more complex interactions, as linear regression would simply be a set of straight lines. Chromium vs. carbon was also plotted for YS and UTS. As carbon is principally a solid solution strengthener, it is thought to have more effect on YS than on work hardening, hence the presence of linear lines in figure 8.19b.



Figures 8.18 (a) and (b): Contour plots of YS and UTS respectively for heat treatment temperature vs. test temperature using composition D; (c) and (d) are plots of YS and UTS error bars respectively.

Figures 8.18(c) and (d), along with 8.19(c) and (d), indicate the error for each prediction, which includes $\pm 1\sigma$ and fitting uncertainty. Note that features found for the prediction contour plots may not be realistic if the corresponding error bars are large enough to cause doubt.



Figures 8.19 (a) and (b): Contour plots of YS and UTS respectively for chromium vs. carbon using composition D; (c) and (d) are plots of YS and UTS error bars respectively.

8.5.7 Software

There were a number of trends examined, however there are too many to report upon, due to the countless permutations possible. Despite this, the results appeared reasonable from metallurgical understanding. It has been shown that the neural network has a good understanding of the complexity of modelling steel strength. As mentioned previously, the advantage with this technique is the safer extrapolation, with error bars associated with all predictions. The software for calculating these tensile properties can be obtained from the web:

http://www.msm.cam.ac.uk/map/map.html

8.6 Summary

The yield strength and UTS have been analysed using neural networks with a Bayesian framework. A variety of sources were used to form the database of composition and temperature information. Good agreement was found when the models were tested against the literature. This includes the effects of varying boron and silicon concentrations, and work hardening behaviour at high temperatures. Their ability to perceive non-linear interactions between different inputs was also indicated by the curves from contour plots.

Now that a feasible predictive model exists, it would be useful to find new compositions for austenitic stainless steels. Each input variable has some influence on the final steel strength. So to aid this investigation, a useful approach would be the use of “genetic algorithms”. By using this technique in conjunction with neural networks, it is hoped that an efficient search of the envelope of all inputs can lead to the same or improved strength.