Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Nuclear Materials 348 (2006) 311–328
journal of nuclear materials
www.elsevier.com/locate/jnucmat

Neural-network analysis of irradiation hardening in low-activation steels

R. Kemp [a,*], G.A. Cottrell [b], H.K.D.H. Bhadeshia [a], G.R. Odette [c], T. Yamamoto [c], H. Kishimoto [c]

[a] *Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK*
[b] *EURATOM/UKAEA Fusion Association, Culham Science Centre, Abingdon, Oxon OX14 3DB, UK*
[c] *Department of Mechanical and Environmental Engineering and Department of Materials, University of California Santa Barbara, Santa Barbara, CA 93106, USA*

Received 26 January 2005; accepted 30 September 2005

Abstract

An artificial neural network has been used to model the irradiation hardening of low-activation ferritic/martensitic steels. The data used to create the model span a range of displacement damage of 0–90 dpa, within a temperature range of 273–973 K and contain 1800 points. The trained model has been able to capture the non-linear dependence of yield strength on the chemical composition and irradiation parameters. The ability of the model to generalise on unseen data has been tested and regions within the input domain that are sparsely populated have been identified. These are the regions where future experiments could be focused. It is shown that this method of analysis, because of its ability to capture complex relationships between the many variables, could help in the design of maximally informative experiments on materials in future irradiation test facilities. This will accelerate the acquisition of the key missing knowledge to assist the materials choices in a future fusion power plant.
© 2005 Elsevier B.V. All rights reserved.

PACS: 61.82 Bg

1. Introduction

Future magnetic confinement fusion power plants will need to be fabricated from components having service lifetimes, in the high neutron-flux region of the first wall, of typically five years. Structural first-wall materials such as steels will have to be developed that can retain good mechanical properties in the presence of neutron irradiation damage of up to 200 atomic lattice displacements per atom (dpa) with transmutation helium gas at levels up to 2000 atomic parts per million (appm) over the design life [1]. The materials also need to minimise activation through the formation of long-lived radionuclides from alloying and impurity elements.

A reliable fusion power-plant design should ideally be based on a comprehensive, multi-variable

* Corresponding author. Tel.: +44 (0) 1223 334 336; fax: +44 (0) 1223 334 567.
E-mail address: rk237@cam.ac.uk (R. Kemp).

0022-3115/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jnucmat.2005.09.022

materials database well-populated with the measured mechanical properties of candidate fusion materials, covering the ranges of physical conditions expected. However, at present, few data exist on candidate alloys at high dpa and there are no high-dose data at fusion-relevant He/dpa ratios. Proposed materials testing facilities, such as IFMIF [2], are designed to provide relevant data but owing to limitations on the number of samples that can be irradiated in such a facility, it will be vital to prioritise the most effective experiments to fill the gaps in our knowledge.

In the absence of a comprehensive experimental database, properties can be estimated based on models that have been fitted to the available experimental data. It is important therefore to establish techniques that make the best use of complex, multi-variate data that are not uniformly distributed in the variable space. The modelling technique we describe below is not unique to materials science. Previous complex materials problems where it has been successfully applied include the modelling and optimisation of the Charpy toughness and strength of steel weld metals, the yield and ultimate tensile strength of nickel-base superalloys, the behaviour of high-temperature, low-creep steels, and properties of polymeric and inorganic compounds and ceramics. A review of these applications has been given by Bhadeshia [3]. However, as far as we are aware, the modelling method has so far not been applied to the mechanical properties of neutron-irradiated materials, such as those that are being considered for future fusion power plants.

In this paper we describe a technique to represent and analyse a relatively small and sparse database of irradiated materials properties for irradiation hardening of reduced-activation ferritic/martensitic (RAFM) steels. Irradiation hardening is widely believed to be caused by microstructural changes such as the formation of neutron-induced precipitates, dislocation loops, voids and helium-containing bubbles. Such defects contribute to dispersed-barrier hardening and embrittlement, and in some cases, swelling. However, the details of the hardening process depend on many variables including the chemical composition of the steel and irradiation variables such as the dose, dose rate and temperature. There is currently no comprehensive physical model that can predict the hardening expected as a function of all these inputs.

An artificial neural network (ANN) is therefore used to model the experimental irradiation hardening data (the term 'artificial' is used to indicate that these networks are computer programs, and therefore to distinguish them from 'real' neural networks, such as the human brain). Such a network is able to capture complex non-linear patterns within the data. The specific method used is based on a Bayesian statistical approach, whereby the uncertainty associated with any prediction takes account both of experimental noise and the modelling uncertainty. It is possible therefore to assign confidence limits that depend on the position in the input space where the prediction is attempted. This not only warns of excessive extrapolation beyond the knowledge base, but also indicates domains where the dataset may be sparsely populated, resulting in uncertain predictions. These methods have previously been extensively used in modelling the properties of materials which have not been subjected to irradiation [3].

2. Analysis method

An ANN is used to perform multi-dimensional non-linear regression on a dataset. The yield strength of unirradiated and irradiated steels is expressed as a non-linear function, f, of a number of experimental variables in the database,

$$\sigma_y = f(\{c_i\}, c_{He}, \mathrm{CWT}, K, t, T_{irr}, T_{test}, \ldots), \quad (1)$$

where $\{c_i\}$ is the chemical composition of the alloy, c_{He} the concentration of any transmutation helium produced, CWT a parameter to describe any cold-working treatment prior to irradiation, K the irradiation rate (in dpa s^{-1}), t the duration time of the irradiation – giving the overall irradiation damage as Kt – and T_{irr} and T_{test} are respectively the irradiation and test temperatures of the sample.

Because assumptions are not made about the form of this function (other than that it should be differentiable), the neural network used is able to respond flexibly to the demands made by the data, capturing non-linear interactions between the parameters. The form of the function f is inferred through a training process during which the network is exposed to a training dataset. If, on the basis of what is known about irradiation effects, specific combinations of inputs are justified then they are included as additional inputs; this is discussed in Section 4. The aim is to deduce Eq. (1) and to provide a measure of the uncertainty associated with the predictions made from it. To control the complexity of the fitting function, and thus any

tendency to overfit the data, the training of the model involves minimising an objective function, which penalises over-complex functions; the overfitting is also managed by testing the generalisation of the model on unseen test data.

A brief description of the general procedure is as follows. The arguments of f, or *input variables*, are gathered in a set of what are termed *hidden units*. A hidden unit is a non-linear operator – a 'neuron' – that is activated by a weighted sum of the input variables. Each different neuron (i) possesses a weight (w_{ij}) for each of the j input variables, as well as a constant *bias term* θ_i. As noted previously, the input variables can be specified functions of the actual independent irradiation and material variables, or combinations of variables. The linear weighted sum of the input variables and the bias term form the arguments for the neuron activation function, h_i – in our case tanh(). The predicted network output, y, is then computed as the linear weighted sum of all h_i. An untrained model is thus defined by the choice of input variables, the number of hidden units, and the set of initial guesses for w_{ij} and θ_i, the *seed*. A series of models is thus trained with different numbers of hidden units and seeds, so that the optimal model or models can be located. A more accurate prediction can often be produced by creating a *committee* of more than one model, and combining their predictions. The overall predicted output is then the average value of y from each member of the committee.

The ANN training algorithm searches for an optimal set of weights by a constrained minimization of the variance between the predicted and measured values on a set of *training data*. The model is then evaluated against a roughly equally sized set of independent *test data* that was not used in the training itself. Over-correlation of the training data is avoided by minimising this *test error*. The procedure used here (due to MacKay [4]) provides uncertainty estimates for the predictions; these uncertainties are large in regions of sparsely populated data space and for significant extrapolations. This procedure is described in more detail in Appendix A.

3. The database

A database of the tensile properties of a set of RAFM steels was compiled, primarily by Yamamoto and co-workers, from the published literature [5–49]. The database includes leading candidate RAFM steels such as F82H (8Cr) and Eurofer 97 (9Cr) as well as conventional 9Cr steels such as T91 (9Cr–1Mo) and EM10 (9Cr–1Mo). A set of 2.5Cr bainitic steel data is also included in the database, which contains a total of about 1800

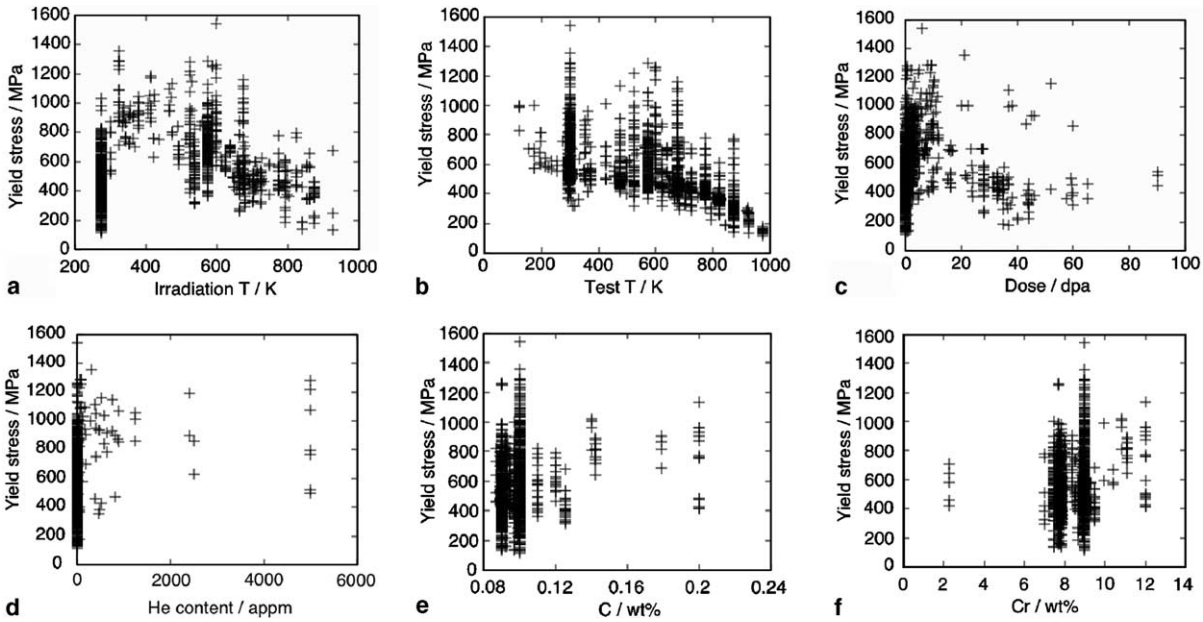

Fig. 1. Distribution of a selection of different inputs, to illustrate the spread of data.

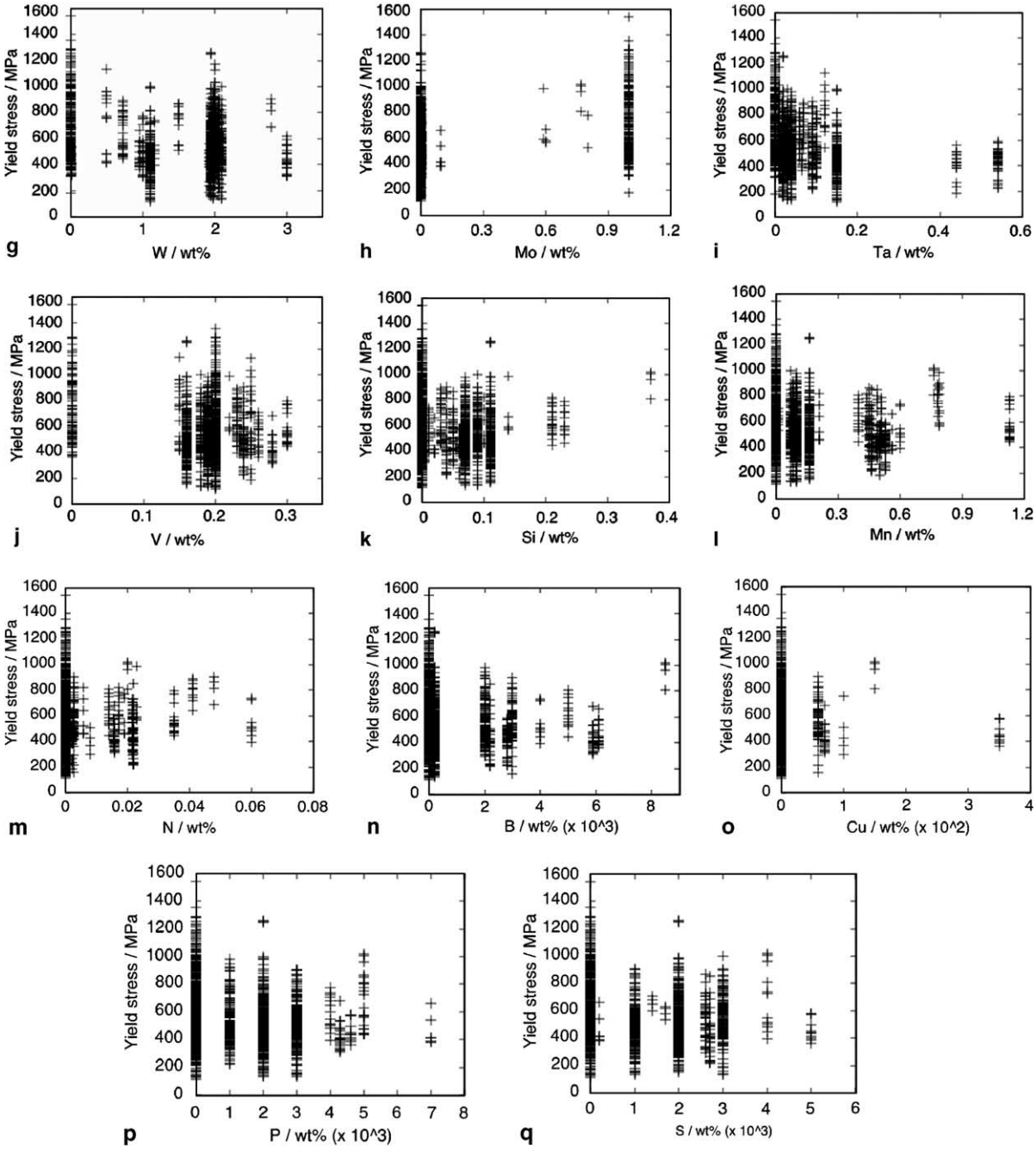

Fig. 1 (*continued*)

datapoints. The data were randomly divided into two equal groups designated the training and test datasets. Fig. 1 illustrates the range of many of these variables plotted against the yield stress, σ_y. These graphs do not describe functional dependences, but merely demonstrate that σ_y varies with many inputs. It is evident that the data are not uniformly distributed and further heterogeneities may exist when correlations between the input variables are considered. However, as will be seen later, the present method is able to highlight regions of the input space that lack information by associating larger modelling uncertainties with predictions made in such regions.

Table 1
Various inputs in the data set

Input variable	Min.	Max.	Mean.	Standard deviation
Irradiation temperature, T_{irr}, K	273	925	401.2	179.3
Test temperature, T_{test}, K	123	973	549.5	209.4
Dose, dpa	0	90	3.47	10.04
He content, appm	0	5000	38.4	359.8
Cold working, %	0	10	0.09	0.94
Composition, wt%				
C	0.087	0.2	0.097	0.013
Cr	2.25	12	8.325	1.027
W	0	3	1.485	0.778
Mo	0	1	0.16	0.363
Ta	0	0.54	0.064	0.102
V	0	0.3	0.182	0.054
Si	0	0.37	0.055	0.052
Mn	0	1.13	0.145	0.204
N	0	0.06	0.0025	0.081
Al	0	0.054	0.0008	0.037
As	0	0.005	0	0.0003
B	0	0.0085	0.0007	0.0013
Bi	0	0.005	0	0.0003
Ce	0	0.036	0.0001	0.0022
Co	0	0.01	0.0002	0.0009
Cu	0	0.035	0.0006	0.0032
Ge	0	1.2	0.0139	0.128
Mg	0	0.01	0	0.0006
Nb	0	0.16	0.0017	0.011
Ni	0	2	0.0576	0.31
O	0	0.009	0.0002	0.0011
P	0	0.007	0.0013	0.0014
Pb	0	0.005	0	0.0003
S	0	0.005	0.0012	0.0011
Sb	0	0.003	0	0.0002
Se	0	0.003	0	0.0002
Sn	0	0.003	0	0.0002
Te	0	0.005	0	0.0003
Ti	0	0.25	0.0099	0.046
Zn	0	0.005	0	0.0003
Zr	0	0.059	0.0003	0.0036

The input variables are listed in Table 1 and include most of the parameters thought to influence the strength of irradiated steel: chemical composition and degree of cold working, irradiation temperature, test temperature, total irradiation displacement damage, and helium content. Because the algorithm used is able to deduce the relevance of each individual input variable, it is appropriate to include all the available variables in the analysis. Those variables which are not found to greatly influence the strength are associated by the model with minimal weights [50]. In some cases, full compositional data were not available and in these instances the concentrations of deliberately added chemical elements were set to zero, while the concentrations of impurities, which are inevitably present, were set to be equal to the average of the available data. The data contained no information on the dose rate, K, or irradiation time [51,52]. Pre-irradiation heat-treatment information was also missing for most of the data, and hence could not be included in the training process. By highlighting these difficulties it is hoped that future experiments are better defined with a full reporting of the variables deemed here to be useful.

4. Yield stress model

The yield stress was modelled directly, rather than the change caused by irradiation, as this is well

characterised in unirradiated materials and so provides a good baseline for the algorithm; $\Delta\sigma_y$ can then be calculated from the model outputs. The choice of input variables is clearly important in the training of a network. It is known, for example, that the yield stress is affected by many factors, including grain size, dislocation and precipitate densities, and the presence and concentration of such alloying elements as Ta and W [53,54]. However, although the principles of these hardening mechanisms are understood, it is not possible to make quantitative predictions directly from theory owing to the complexity of the phenomena involved. An attempt can nevertheless be made to include aspects of known physics by using appropriate combinations of input variables. It should be noted that, if these functional forms are not useful in explaining the data, they will be effectively ignored by the network. Thus, in addition to the raw inputs (Table 1), terms were included as follows:

(a) Two Arrhenius-type relations for the temperatures, T_{irr} and T_{test}, to take account of any thermally activated processes. Although the processes involved cannot be simply described by Arrhenius rate expressions, including this term may help the model to capture interactions that involve it. These have the standard form

$$\exp\left(\frac{-Q}{kT}\right), \tag{2}$$

in which Q is an activation energy and k is Boltzmann's constant.

(b) The He concentration-to-dpa ratio was included, to complement the c_{He} input to allow for possible interactions between gas and displacement damage.

(c) A term defining saturation in the hardening as the microstructure reaches a steady state in terms of defect formation and annihilation as is thought to occur under some irradiation conditions [55,56].

$$1 - \exp\left(-\frac{Kt}{\alpha}\right). \tag{3}$$

Here, α is a fitting parameter dependent on the material and irradiation conditions – particularly T_{irr}. Owing to the functional form of Eq. (3), it is not possible to define the problem in such a way that α can be inferred by the network – an unfortunate limitation of the methodology. Therefore, multiple values of α were included as separate inputs leaving the network to adopt the most suitable value(s). In our case, a set of initial models was trained for 11 values of α between 0.1 and 10 dpa. The significances (equivalent to partial correlation

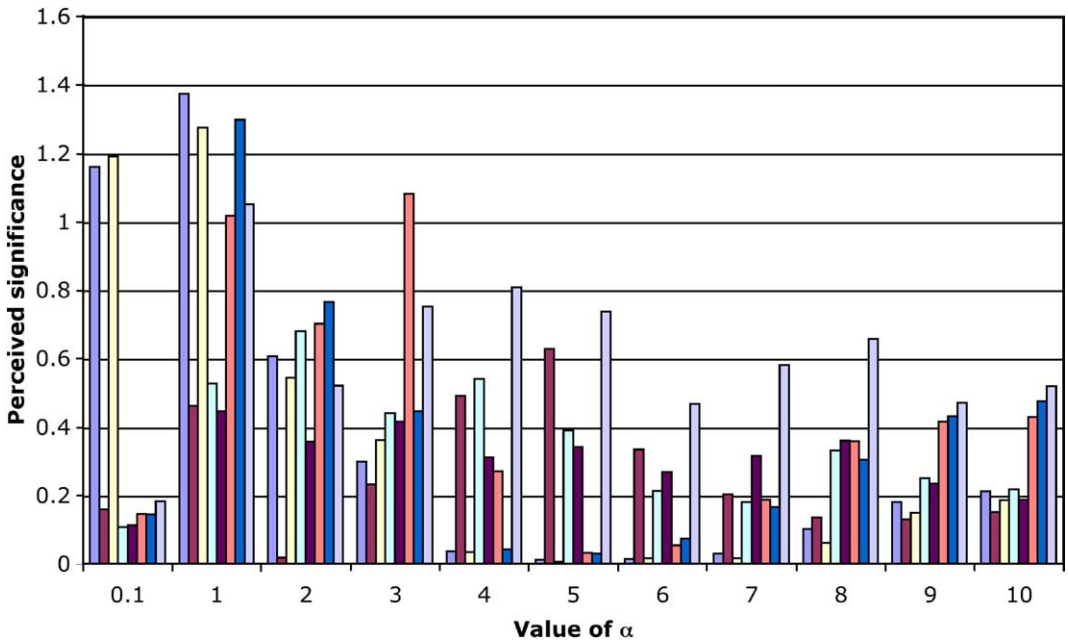

Fig. 2. Perceived significances for inputs with different values of α. Each bar represents a different model in the committee – there were eight models in the committee in total.

coefficients) inferred by the network for each of these values were then examined (Fig. 2) – each bar represents one of the eight models in the committee used. The significance is maximum when an input is able to explain a large amount of the variation in the output. It is striking that low values ($\alpha \sim 1$–4) are generally preferred. This is consistent with there being a strong low-dose/short time scale influence, and such values are in agreement with those observed by Yamamoto et al. [56]. Although all values could have been included in the final model, we chose to set $\alpha = 1$ for final model training. It would have been desirable to also incorporate at least one higher value, but models trained with multiple values of α were more complex and suffered an accompanying increase in uncertainty. This frequently occurs when there are a large number of similar inputs and limited data to allow them to be distinguished.

To avoid biasing the model towards any particular combinations of parameters such as those described above, the individual variables making up these inputs were also included, so that their direct influence, if any, could also be detected. The model was trained on the natural logarithm of the yield stress, rather than the yield stress itself. This prevented the model from producing non-physical negative predictions, and allowed the simplification of some derived inputs – for example, the Arrhenius relations could be included as an input of the form $(-1/T)$ rather than the exponent (with the Q/k multiplier inferred by the network). This procedure also allowed the model to easily capture any potential power law dependencies, such as the frequently used $\sqrt{Kt}$, which is related to hardening due to microstructural evolution under irradiation [57]. Modelling $\ln(\sigma_y)$ results in a slight decrease in sensitivity over modelling σ_y directly, but the benefits described above more than compensate.

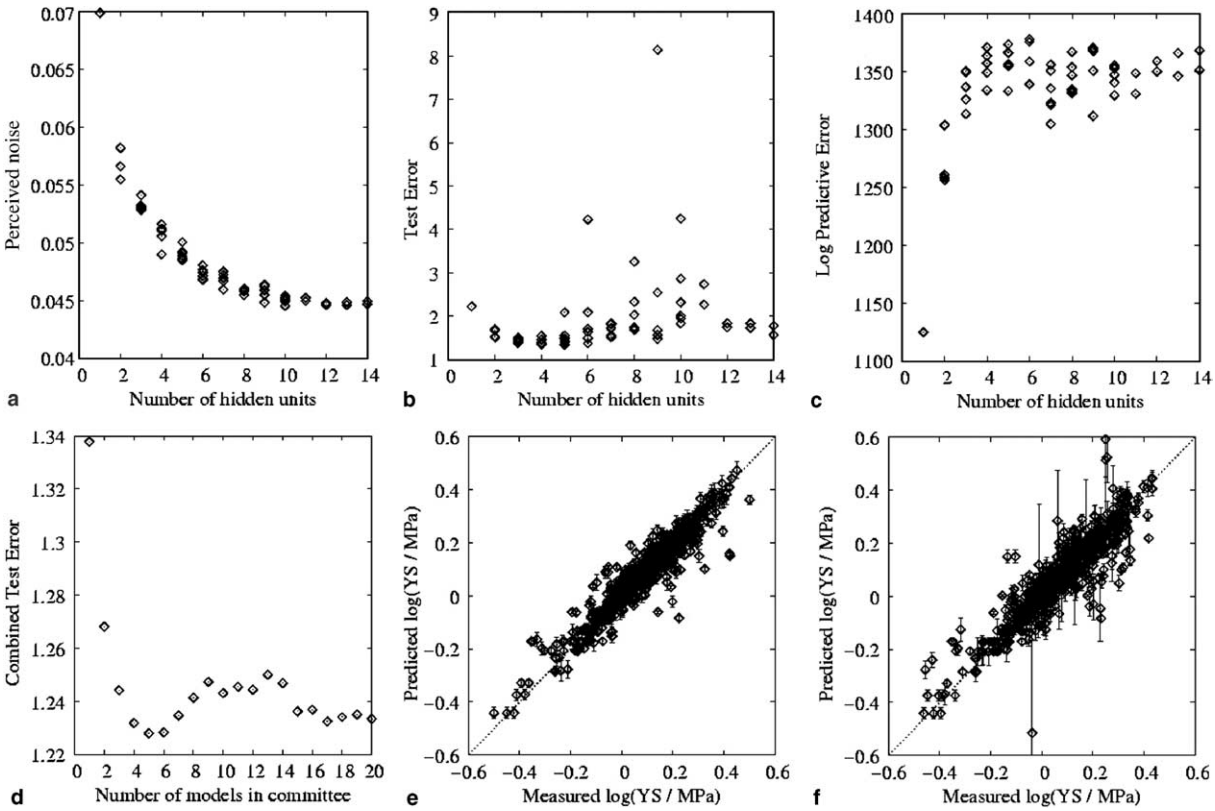

Fig. 3. Training the model. (a) Perceived level of noise for training data, (b) test error, (c) log predictive error for increasing model complexity, (d) combined test error for different sizes of committee, (e) best single model performance on training data and (f) test data – in this case, the data are normalised.

A total of 120 networks were trained with from 1 to 20 hidden units, and from six different seeds. As the networks became more complex (that is, they possess greater numbers of hidden units and are hence more flexible) the perceived noise on the training data decreased as expected (Fig. 3(a)). This is a measure of the 'noise' – i.e. the experimental error – required for the model to be correct given the data – that is, the *likelihood* of the data, assuming the model is correct. However, as shown in Fig. 3(b) the ability of the models to predict σ_y for the test data does not decrease monotonically and is minimum in the range of ≈4 hidden units. Fig. 3(c) shows that the error metric that is used to consider uncertainty and avoid over-correlation – the log predictive error (LPE) (described in Appendix A) – also peaks at about 4–6 hidden units. Also, as anticipated, more accurate results were obtained by building a committee. In this case, the optimum committee was found to have five members (Fig. 3(d)).

As mentioned above, the model provides a measure of the network-perceived significance, σ_w, for

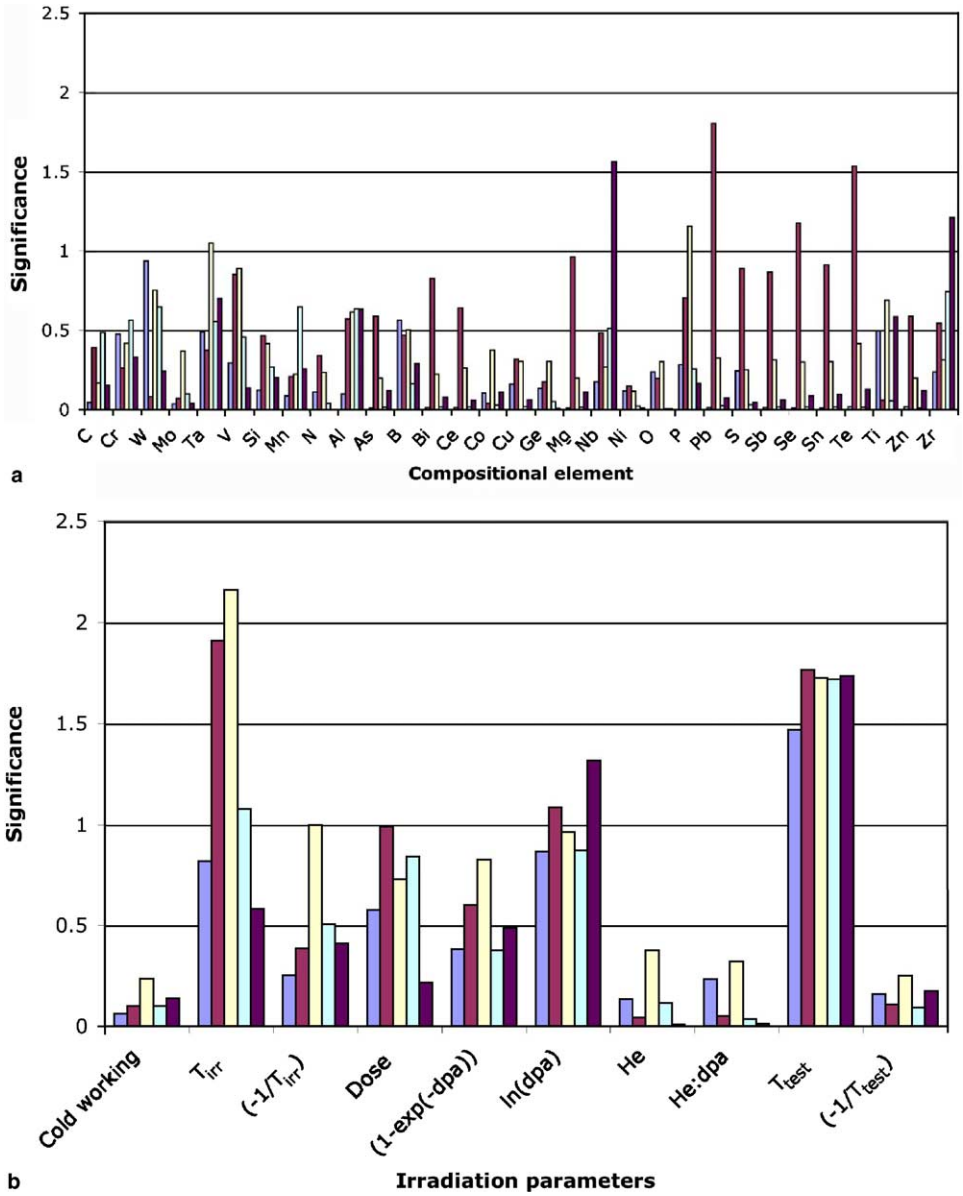

Fig. 4. (a) Perceived significances σ_w for the committee members of the model for alloying elements and (b) perceived significances σ_w for the committee members of the model for irradiation and tensile test conditions.

each input. This is the extent to which each input contributes to the output, and is akin to a partial correlation coefficient in a multiple linear regression analysis. These significances are shown for each sub-model in the committee in Fig. 4(a) (for composition) and Fig. 4(b) (for irradiation and test parameters).

As shown in Fig. 4(a), most of the compositional variables were found to be insignificant. However, this is not at all surprising, even for alloying elements that are known to potentially influence σ_y but that in the database: (a) have a narrow range of values (e.g. C and V); (b) have only one or a few values, some of which may be associated with potentially confounding variants such as different heat treatments (e.g. Cr, Si, Mn, B, Mo, W, Ni, P and Nb); or (c) are alloying or trace impurity elements with missing values set to zero (these are not included in the model). The same can be said for cold work. However, as discussed later, the various alloys did have different σ_y as well as $\Delta\sigma_y$ dependencies on both dpa and T_i. For example, the alloy doped with 2 wt% Ni has a higher σ_y after irradiation. T_i, T_t, and the three dose-related variables were significant, particularly in combination. It is well known that test temperature has a strong effect on tensile tests, and it is to be expected that irradiation temperature will also have a powerful influence as a heat treatment in its own right – this is why the time elapsed during irradiation is potentially an important parameter. Notably neither He nor He/dpa were found to play a significant role in determining σ_y. As an example, Fig. 5 shows that the procedure describes well the non-linear $\sigma_y(T_t)$ of Eurofer 97.

Predictions of the final trained committee model on the input dataset are shown in Fig. 6. Given the sparseness of the dataset and assumptions made for the inputs, there are relatively few outlying points, and those that belong to that class are accompanied by a large uncertainty. There is a clear improvement over the best single model, shown in Fig. 3(e) and (f). For the final committee model, the standard deviation of the predicted versus measured σ_y was 95.17 MPa. There were a total of 42 (from 1814 data points) outliers (more than three standard deviations away).

5. Properties of the trained network

In this section we discuss the properties of the trained neural network, comparing its predictions

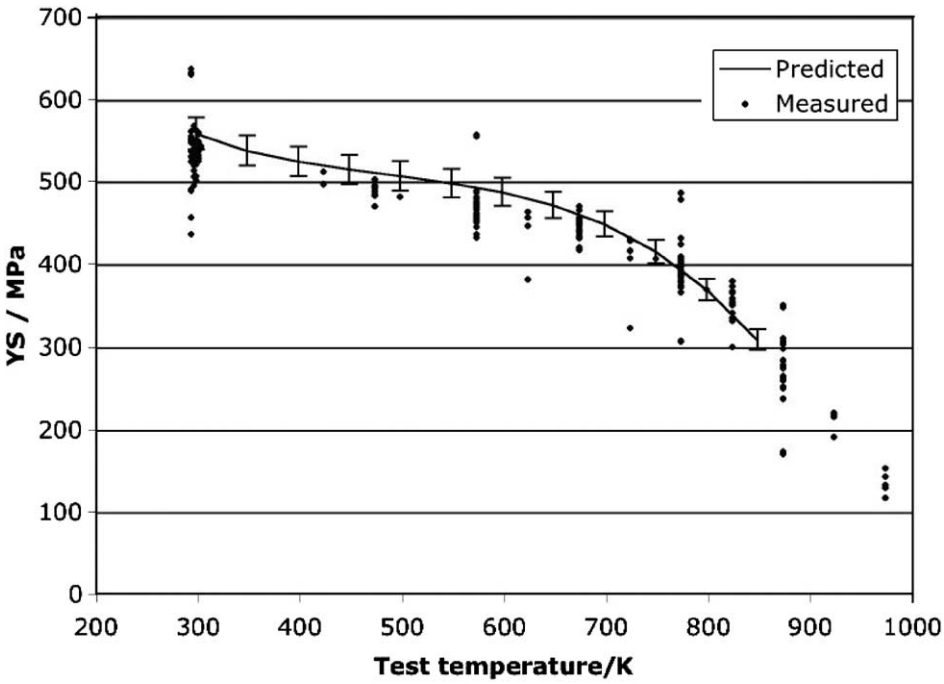

Fig. 5. Model prediction for the tensile test behaviour of unirradiated Eurofer'97 at different test temperatures plotted with measured data. The bars represent modelling uncertainty.

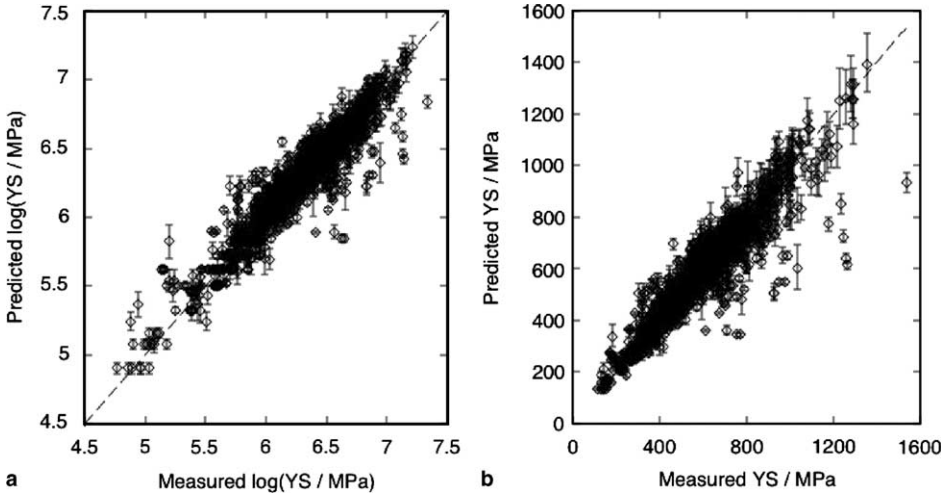

Fig. 6. Performances of the final committee model on the whole database. (a) the direct output $\ln(\sigma_y)$ and (b) converted to σ_y.

Table 2
Chemical composition inputs (in wt%) for model predictions

	Alloy					
Element	F82H	Eurofer'97	T91	9Cr–1WVTa	9Cr–1MoVNb–2W	2.5Cr–1.4WV
C	0.09	0.1	0.1	0.1	0.1	0.1
Cr	7.7	9	9	9	9	2.5
W	1.94	1.1	0	1	0	1.4
Mo	0	0	1	0	1	0
Ta	0.02	0.15	0	0.1	0	0
V	0.16	0.2	0.2	0.2	0.2	0.2
Si	0.11	0	0	0	0	0
Mn	0.16	0	0	0	0	0
N	0	0	0	0	0	0
Al	0	0	0	0	0	0
As	0	0	0	0	0	0
B	0.0002	0	0	0	0	0
Bi	0	0	0	0	0	0
Ce	0	0	0	0	0	0
Co	0	0	0	0	0	0
Cu	0	0	0	0	0	0
Ge	0	0	0	0	0	0
Mg	0	0	0	0	0	0
Nb	0	0	0.005	0	0.006	0
Ni	0	0	0	0	2	0
O	0	0	0	0	0	0
P	0.002	0	0	0	0	0
Pb	0	0	0	0	0	0
S	0.002	0	0	0	0	0
Sb	0	0	0	0	0	0
Se	0	0	0	0	0	0
Sn	0	0	0	0	0	0
Te	0	0	0	0	0	0
Ti	0	0	0	0	0	0
Zn	0	0	0	0	0	0
Zr	0	0	0	0	0	0

with previous conventional analyses of the same database and assessing its ability to predict new, unseen observations. Due to the complexities of a neural-network model, the best way to uncover its

properties is through 'virtual experiments' – using it to make predictions – rather than examination of the weights or structure. The input compositions for the predictions are shown in Table 2. Predictions were not made for all alloys included in the database.

5.1. Test and irradiation temperature

The dashed line in Fig. 7 shows that the model predicts a peak in the T_i ($=T_t$) dependence of $\Delta\sigma_y$ for Eurofer 97. The corresponding predictions based on direct data fits by Yamamoto and co-workers [56] are shown by the solid line. The agreement is quite good below ≈350 °C but the neural-network nominal prediction deviates from the direct data fits at higher T_i, where negligible $\Delta\sigma_y$ is observed. However, the direct fit line generally falls within the estimated uncertainty limits of the neural-network model. Note that while the direct data fit does not predict a peak in hardening at 3 dpa over the nominal T_i range covered, it does predict one for the maximum $\Delta\sigma_y$ beyond a T_i-dependent saturation dose. It can be postulated that such a hardening peak is physically consistent with the thermal mobility and recombination of defects. At low temperatures the radiation-introduced defects are not mobile enough to form the extended structures such as dislocation loops, and cascade overlap quickly leads to a saturated damage state. At higher temperatures these structures tend to dissolve and the defects annhilate at fixed sinks. In the intermediate temperature regime, a hardening peak is therefore not unexpected.

Fig. 8 shows σ_y versus $T_i = T_t$ neural-network predictions for F82H in the irradiated and unirradiated conditions, compared to data. The predictions tend to fall only slightly below the unirradiated reported by Hishanuma [58] over the entire range of T_t. The nominal predicted σ_y for the irradiations to 10 dpa are somewhat higher than Hishanuma's corresponding data points, but within the uncertainty limits.

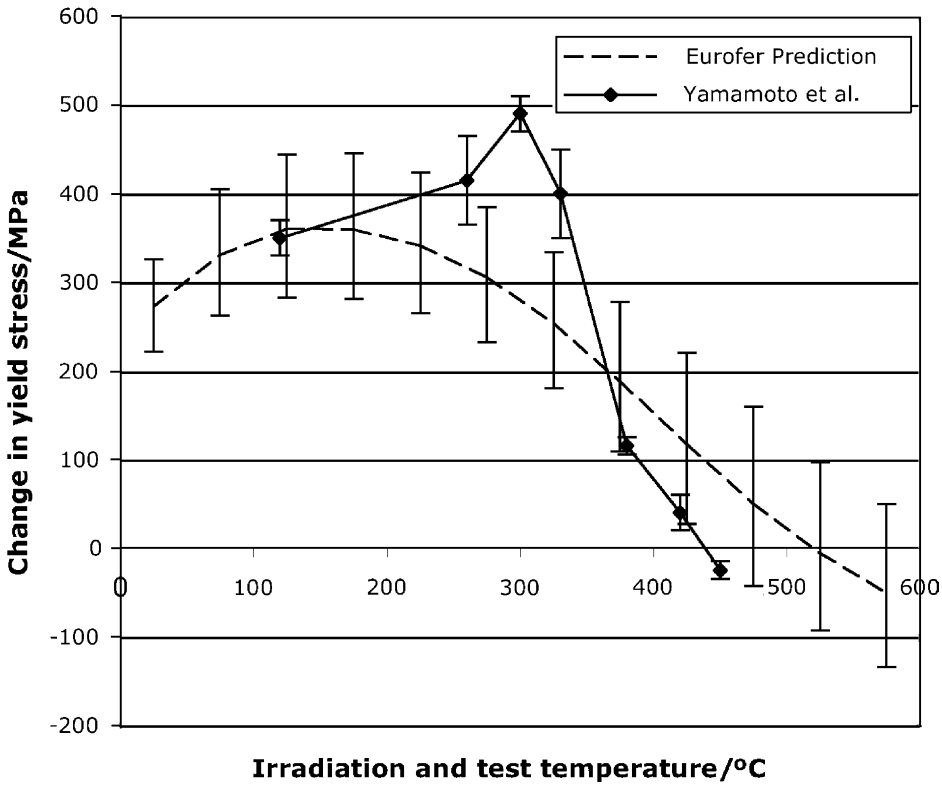

Fig. 7. Model prediction for the yield stress of Eurofer'97 irradiated to 3 dpa at various temperatures, compared with saturation hardening predictions from Yamamoto et al. [56]. A dose of 3 dpa was chosen for the predictions as it generally corresponds with the onset of saturation in Yamamoto's model.

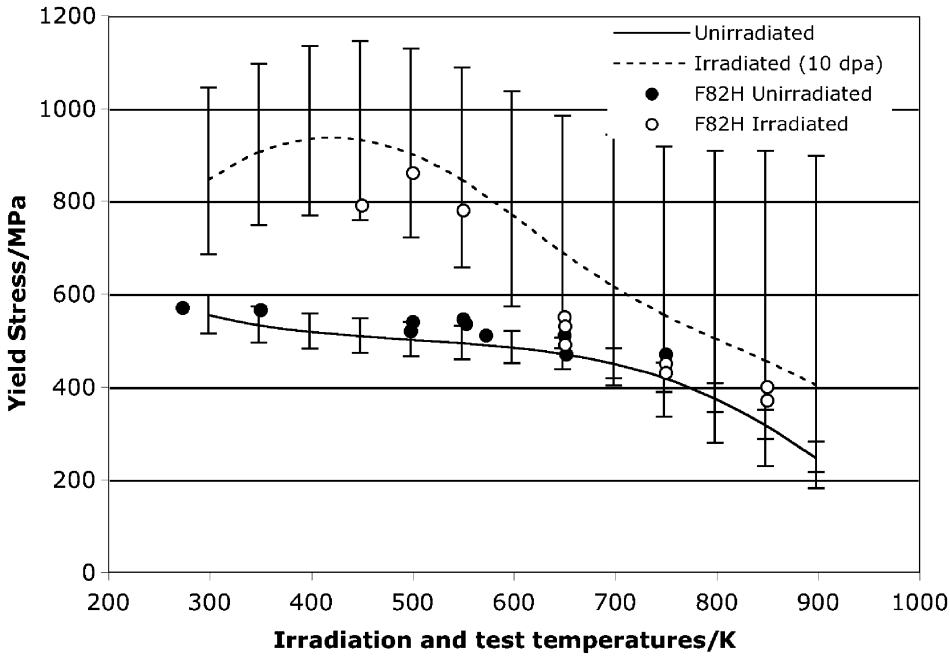

Fig. 8. Model predictions for the behaviour of modified F82H, compared to those reported in Hishinuma [584]. The model predictions are for 10 dpa, $T_{test} = T_{irr}$. The doses in Hishinuma vary from 3 to 34 dpa.

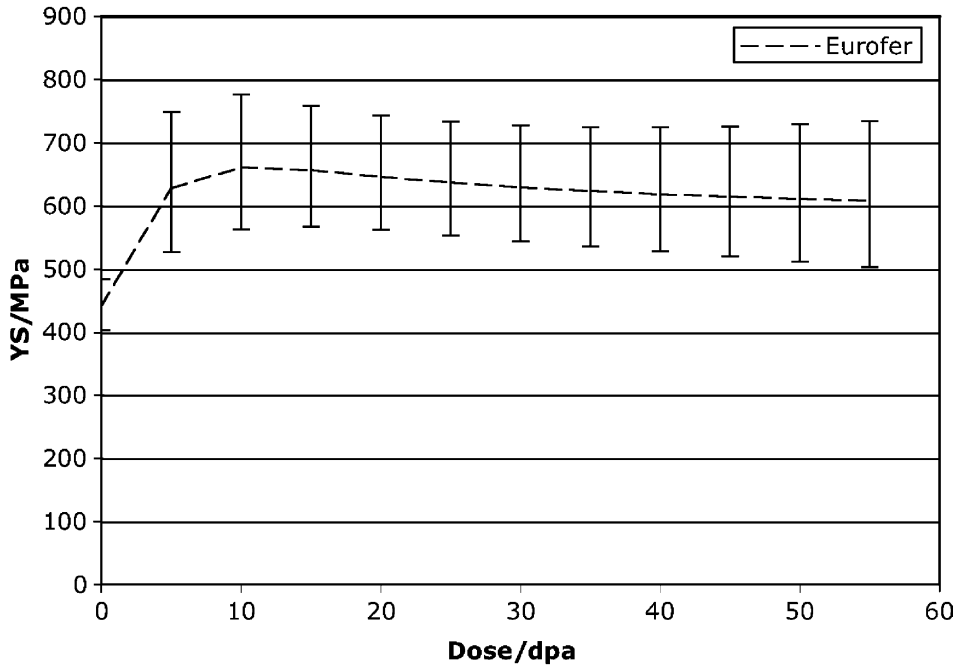

Fig. 9. Model prediction showing irradiation hardening of Eurofer'97. $T_{test} = T_{irr} = 400$ °C.

5.2. Irradiation dose

It is well known that irradiation causes rapid hardening of steels even at low doses, and this low-dose effect is reflected in the model (Fig. 9). The degree to which the hardening will saturate at high doses (as postulated by Makin and Minter [59]) is unknown. Indeed, as observed in some

cases [60], at higher T, the σ_y appears to reach a maximum and then decrease. The model shows a range of behaviours, depending on the alloy, some showing a peak followed by a decrease of σ_y (F82H) while others continue to increase (T91). However, it is not known if these differences are real or an artefact of the database distribution and uncertainties. On average, the ANN-predicted σ_y saturates at a substantial $\Delta\sigma_y$ that is higher than predicted by the direct data fit. It is worth noting, however, that the irradiation parameters for these predictions lie significantly outside the training database and so they are accompanied by large uncertainties. The observed trends, though, are reasonable, but only high-dose data generated in future experiments will help validate and refine the model.

5.3. Chemical influence

The presence of nickel appears to increase the yield stress, both before and after irradiation (Figs. 10 and 11). This influence was observed and described by Klueh and Vitek [61]. However, closer examination of the original data showed that these alloys were subjected to a different heat treatment than others in the dataset, and this may have given rise to this effect.

6. Comparison with other modelling methods

Ideally, and certainly in the long run, mechanical property predictions will be based on hierarchical models of microstructural evolutions linked to structure property relations in a way that will incorporate the effects of all important variables and their interactions, as well as incorporating the known physics and underlying mechanisms. Direct non-linear regression fits to the data using simple phenomenological, but physically motivated models are also useful. For example, the fits by Yamamoto et al. provide a good representation of the existing $\Delta\sigma_y$ database [56]. However, such models often contain assumptions that are sometimes hidden or unrecognised. For example, in the fitting form assumed by Yamamoto et al., high-dose saturation behaviour is assumed and predicted from low-dose trends even in the absence of high-dose data. While uncertainties in such extrapolations can be estimated, the effects of the various assumptions are not easy to quantify in a useful way. In principle, the neural-network approach avoids the need for any assumptions about the form of the fitting equation, and provides error estimates of extrapolations. However, this approach has the corollary disadvantage that it does not allow easy ways to include known physics. Thus the use of

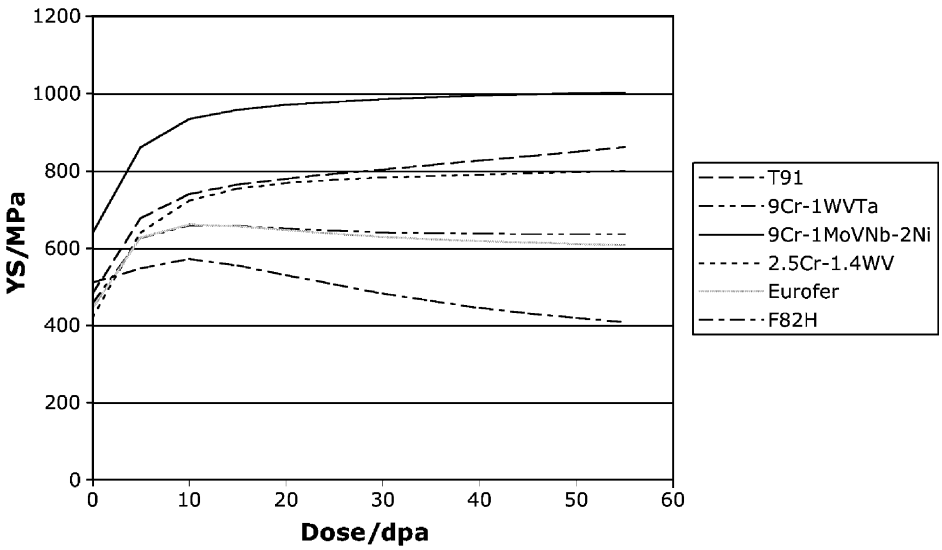

Fig. 10. Comparison of irradiation hardening of various steels showing both saturation and high-dose recovery. $T_{test} = T_{irr} = 400$ °C. The uncertainty bars have been omitted for clarity.

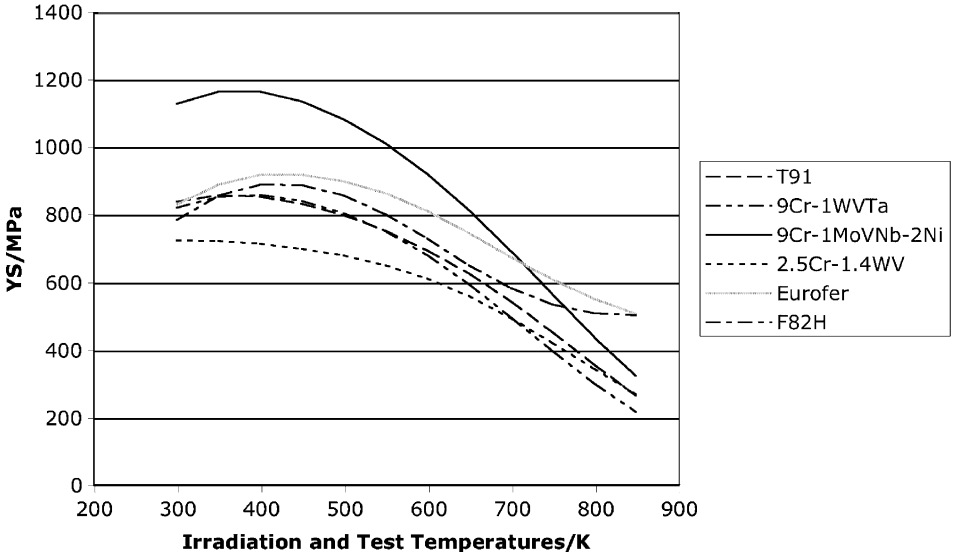

Fig. 11. Comparison between different types of steel irradiated to 2 dpa at different temperatures. The uncertainty bars have been omitted for clarity.

complementary approaches will be needed and perhaps their most important contribution will be to highlight the most significant gaps in our knowledge in a way that can help guide future experiments.

7. Predictions

Fig. 12 shows the neural-network predictions of σ_y and corresponding uncertainty estimates over a range of T_{test}, T_i (300–900 K) and dpa (0–200 dpa).

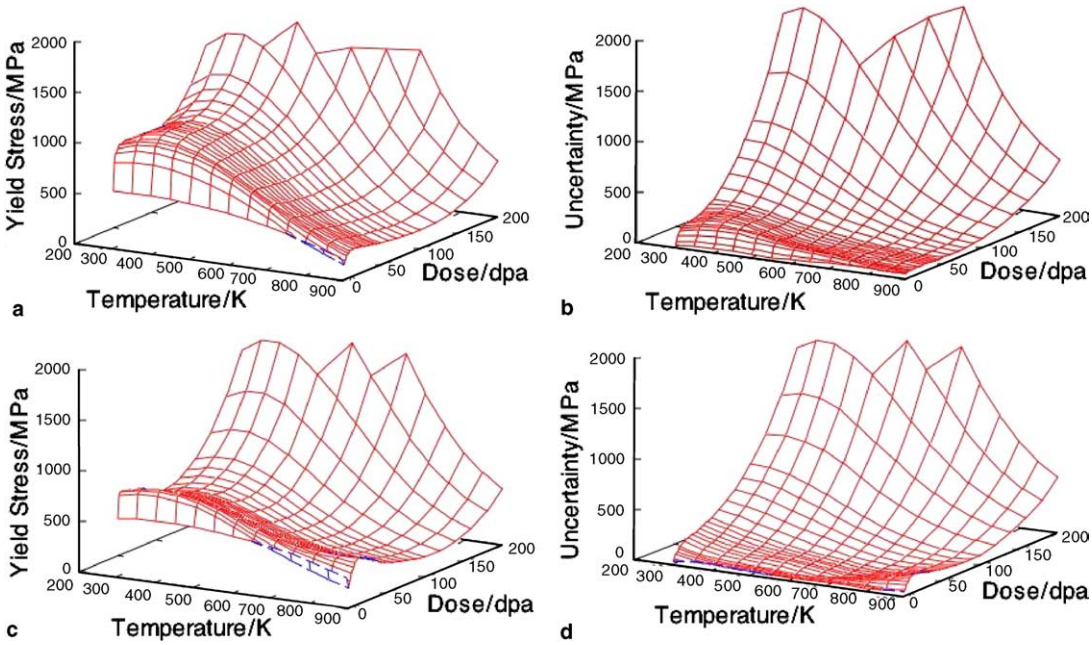

Fig. 12. Hardening predictions (and modelling uncertainties) for Eurofer'97 (top) and F82H (bottom), as a function of temperature ($T_i = T_{test}$) and dose.

The regions well outside the training database have uncertainties that are comparable to or larger than σ_y itself. These uncertainties thus represent areas where experiments might fruitfully be carried out, and provide an indication that predictions in these areas should be regarded sceptically until corroborated.

8. Discussion, summary and conclusions

The yield stresses of RAFM steels have been analysed using a neural-networks method based on a Bayesian framework. The data were obtained from a variety of sources and cover a wide range of compositions and irradiation conditions. Projected trends have been found to be reasonably consistent with those expected, and the quantitative agreement was generally good even when checked against unseen data within the range of the training database.

This paper demonstrates the power and flexibility of a neural-network modelling approach to the problem of analysing irradiation damage, and should be viewed as an early step in a long-term effort. Applied to the specific problem of hardening, one of the first conclusions is that the current database is insufficient to provide the desired refinement of the model that would reduce the uncertainty in predictions of irradiated yield stresses. The construction of an improved database would be helped by future experiments providing sufficient detail of potential important variables, such as dose rate, irradiation energy spectrum, length of irradiation, and providing estimates of error in key variables, such as irradiation temperature.

In spite of these limitations, the model provides a means of making wide-ranging quantitative predictions, including a capacity to indicate uncertainty. This indication of modelling uncertainty also provides information about where it would be useful to concentrate experimental efforts. The high-dose, high-temperature regime is directly relevant to fusion power plants, and yet is short of data – although the difficulty of obtaining this data currently is acknowledged.

Future work will concentrate on expanding this approach to other physical properties affected by irradiation and refinement of the model to take account of improvements in the database.

In summary the main points of this paper are:

1. A database of irradiated RAFM steel properties was assembled from published results.
2. An artificial neural network was successfully trained on this database, and captured non-linear features present in the data.
3. The neural network also predicted unseen data satisfactorily.
4. For complex problems of this kind, ANN models offer a powerful method of identifying trends in datasets and analysing the contributions of a variety of input parameters.
5. A trained ANN model can be interrogated to provide quantitative predictions of material properties as they vary with any input. This can be used to optimise materials across the whole input data space. As the model also provides a measure of uncertainty with its output, it minimises misleading predictions in areas of sparse training data.
6. Experimenters should report complete details of both material manufacture (including chemical analysis including impurities, and any heat treatments used) and experimental conditions: dose rate; irradiation energy; and irradiation temperature. These factors can change markedly depending on the position of the sample in the reactor, and are considered to be important.
7. The current database should be expanded as new experimental results become available, and should incorporate the variables mentioned above, where available.

Acknowledgements

It is a pleasure to thank Ian Cook for encouragement and stimulating discussions. This work was funded by the UK Engineering and Physical Sciences Research Council and EURATOM, and was also supported by the US DOE Office of Fusion Energy Science (Grant # DE-FG03-94ER54275).

Appendix A. The artificial neural network

The operation and construction of such networks have been reviewed in detail elsewhere [50,62], and we summarise the main features here. A simple three layer feed-forward variety, of the type commonly used for many applications, is shown in Fig. A1. The first layer consists of the inputs to the network. The second layer consists of a number of *neurons*, non-linear operators whose arguments are provided by the first layer in the network. The activation

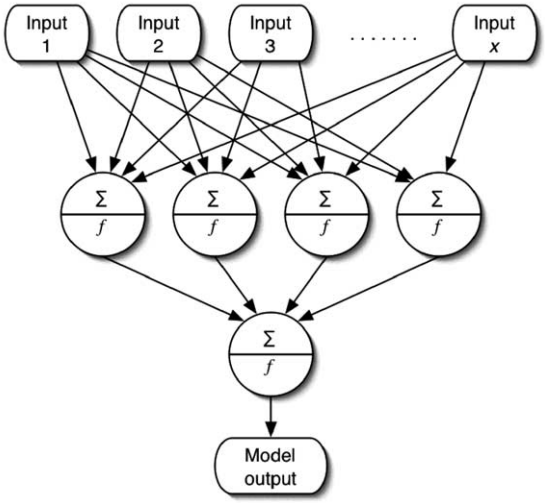

Fig. A1. A schematic diagram of a three layer feed-forward network. The model's complexity is controlled by the number of neurons in the second layer, known as hidden units.

function for the neurons in the second layer, h_i, can be any non-linear continuous and differentiable function; in this work we have used the tanh function (Eq. (A.1)). The overall output function in the third layer, y, can again be any function, and is commonly linear (Eq. (A.2)). The neuron activation function is given by

$$h_i = \tanh\left(\sum_j w_{ij}^{(1)} x_i + \theta_i^{(1)}\right) \tag{A.1}$$

and the output weighting function is

$$y = \sum_i w_i^{(2)} h_i + \theta^{(2)}. \tag{A.2}$$

The x_i are the inputs, and w the weights which define the network. The aim of training a network is to find the optimum set of values for w. This procedure is explained in more detail below. The parameters θ are known as biases, and are treated internally as weights associated with a constant input set to unity. The complexity of such networks scales with the number of 'hidden' units in the middle layer, and it has been shown that with sufficient such neurons, any continuous, differentiable function can be simulated [4].

In order to simplify the weightings, inputs and targets are normalised prior to training within a range of ±0.5. The normalisation function is

$$x_i = \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 0.5, \tag{A.3}$$

where x is the un-normalised input, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values in the database for that input, and x_i is the normalised value. The network is therefore not constrained to a particular range of outputs (for example, positive outputs only) and so the target must be chosen with care to avoid unphysical model outputs. For example, for a property which cannot be less than zero such as σ_y, $\ln(\sigma_y)$ could be used as the network training target.

Because of the inherent flexibility of an ANN, there is the possibility of overfitting a model. Training a network therefore involves finding a set of weights and biases that minimise an *objective function* which balances complexity and accuracy, typically

$$M(w) = \alpha E_w + \beta E_D, \tag{A.4}$$

in which E_w is a *regulariser*; its function is to force the network to use small weights:

$$E_w = \frac{1}{2}\sum_i w_i^2 \tag{A.5}$$

and E_D is the overall error between target output values and network output values:

$$E_D = \frac{1}{2}\sum_i (t^{(i)} - y^{(i)})^2, \tag{A.6}$$

where $t^{(i)}$ is the set of targets for the set of inputs $x^{(i)}$, while $y^{(i)}$ is the set of corresponding network outputs. α and β are control parameters which influence the balance between a simple and an over-complex model.

In this case, the fitting method used was based on a Bayesian approach and treats training as an inference problem, allowing estimates to be made of the uncertainty of the model fit (Fig. A2). Rather than trying to identify the best set of weights, the algorithm infers a probability distribution for the weights from the data presented. In this context, the performances of different models are best evaluated using the *log predictive error* (LPE) rather than the test error. This error penalises wild predictions to a lesser extent when they are accompanied by appropriately large error bars and is defined by

$$\mathrm{LPE} = \sum_m \left[\frac{1}{2}\frac{\left(t^{(m)} - y^{(m)}\right)^2}{\left(\sigma_y^{(m)}\right)^2} + \log\left(\sqrt{2\pi}\sigma_y^{(m)}\right)\right], \tag{A.7}$$

where t and y are as defined above, and $\sigma_y^{(m)}$ is related to the uncertainty of fitting for the set of

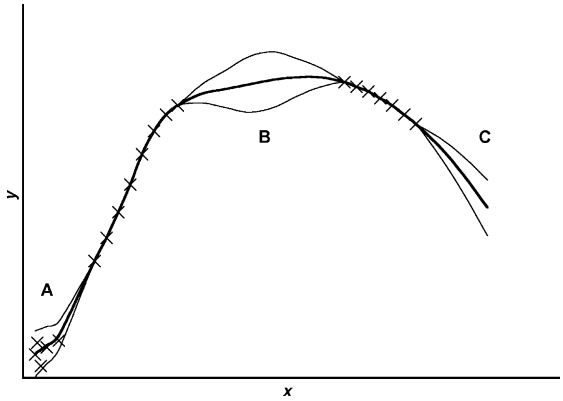

Fig. A2. Schematic illustration of the uncertainty in defining a fitting function in regions where data are sparse (B) or noisy (A). The thinner lines represent error bounds due to uncertainties in determining the weights. Note that, outside the range of data, the extrapolation is increasingly uncertain (C).

inputs $x^{(m)}$. It should be pointed out that, for computational reasons, the training software actually uses a negative version on this function that increases with increasing accuracy (Fig. 3(c)).

To train a useful model, it is necessary to check its ability to generalise to new data. To accomplish this, the training data are randomly split into two sets, a *training set* and a *test set*. The model is trained on the first set, and then its predictions are compared against the second. The error in these predictions is called the *test error*, and the ultimate purpose of training a model is to minimise this error, both against the training data and against unseen data from future experiments.

Of course, models with different numbers of hidden units will give different predictions. These predictions will depend not only on the set of training data, but also the initial guess made for the probability distribution of the weights, the prior (the *seed*). Therefore, optimum predictions can often be made by using more than one network. This is referred to as a *committee*. The prediction $\bar{y}$ of a committee of networks is the average prediction of its members, and the associated uncertainty is

$$\sigma^2 = \frac{1}{L}\sum_l \sigma_y^{(l)^2} + \frac{1}{L}\sum_l (y^{(l)} - \bar{y})^2, \tag{A.8}$$

where L is the number of networks in the committee and the exponent l refers to the model used to give the corresponding prediction $y^{(l)}$. During training, it is usual to compare the performances of increasingly large committees on the testing set of data (Fig. 3(d)). Usually, the error is minimised by using more than one model in the committee. The selected models are then retrained on the entire database.

Although the values of the weights in these models can always be examined, they are difficult to interpret directly, and the easiest way to identify the interactions in a model is to use it to make predictions and observe the behaviour that emerges from various combinations of inputs.

The algorithm we used is due to MacKay and allows the control parameters to be inferred from the data, permitting automatic control of the model complexity [4].

Software

A number of other trends predicted by the model are currently being examined, but the number of possibilities of interactions is such that it is not possible to study them all fully, so it is hoped that other researchers will take the opportunity to help refine the model and the dataset. As the dataset becomes more complete and covers greater ranges of compositions of steel and irradiation conditions, the model will become more accurate. A software suite for carrying out predictions using this model is available online http://www.msm.ac.uk/map/map.html.

References

[1] K. Ehrlich, E.E. Bloom, T. Kondo, J. Nucl. Mater. 283–287 (2000) 79.
[2] K. Ehrlich, A. Möslang, Nucl. Instrum. and Meth. B 139 (1998) 72.
[3] H.K.D.H. Bhadeshia, ISIJ Int. 39 (10) (1999) 966.
[4] D. MacKay, Information Theory, Inference, and Learning Algorithms, 1st Ed., Cambridge University, 2003.
[5] F. Abe et al., J. Nucl. Sci. Technol. 31 (4) (1994) 279.
[6] A. Alamo et al., J. Nucl. Mater. 283 (2000) 353.
[7] N. Baluc et al., J. Nucl. Mater. 283 (2000) 731.
[8] I. Belianov, P. Marmy, J. Nucl. Mater. 263 (1998) 1259.
[9] Y. Dai, X.J. Jia, K. Farrell, J. Nucl. Mater. 318 (2003) 192.
[10] Y. Dai et al., J. Nucl. Mater. 283 (2000) 513.
[11] K. Farrell, T.S. Byun, J. Nucl. Mater. 296 (2001) 129.
[12] K. Farrell, T.S. Byun, J. Nucl. Mater. 318 (2003) 274.
[13] P. Fernandez et al., Fus. Eng. Des. 58–59 (2001) 787.
[14] I.V. Gorynin et al., J. Nucl. Mater. 283 (2000) 465.
[15] J. Henry et al., J. Nucl. Mater. 318 (2003) 215.
[16] J. Henry, M.H. Mathon, P. Jung, J. Nucl. Mater. 318 (2003) 249.
[17] P. Jung et al., J. Nucl. Mater. 318 (2003) 241.
[18] R. Kasada et al., J. Nucl. Mater. 272 (1999) 360.
[19] R. Kasada et al., J. Nucl. Mater. 263 (1998) 1199.
[20] A. Kimura et al., J. Nucl. Mater. 237 (1996) 319.
[21] R.L. Klueh, J. Nucl. Mater. 179 (1991) 728.

[22] R.L. Klueh, D.J. AlexanderEffects of Radiation on Materials: 18th International Symposium, ASTM, 1999.
[23] R.L. Klueh, D.J. Alexander, M. Rieth, J. Nucl. Mater. 273 (2) (1999) 146.
[24] R.L. Klueh, D.J. Alexander, M.A. Sokolov, J. Nucl. Mater. 304 (2–3) (2002) 139.
[25] R.L. Klueh, J.J. Kai, D.J. Alexander, J. Nucl. Mater. 225 (1995) 175.
[26] R.L. Klueh, P.J. Maziasz, J. Nucl. Mater. 187 (1992) 43.
[27] R.L. Klueh et al., J. Nucl. Mater. 283 (2000) 478.
[28] R.L. Klueh, J.M. Vitek, J. Nucl. Mater. 132 (1985) 27.
[29] R.L. Klueh, J.M. Vitek, J. Nucl. Mater. 161 (1989) 13.
[30] Y. Kohno et al., J. Nucl. Mater. 272 (1999) 145.
[31] A. Kohyama et al., J. Nucl. Mater. 155 (1988) 896.
[32] H. Kurishita et al., J. Nucl. Mater. 215 (1994) 730.
[33] R. Lindau et al., J. Nucl. Mater. 272 (1999) 450.
[34] E. Lucon, M. Decreton, E. van Walle, Fus. Eng. Des. 69 (2003) 373.
[35] S.A. Maloy et al., J. Nucl. Mater. 296 (2001) 119.
[36] E.I. Materna-Morris, M. Rieth, K. Ehrlich, in: Effects of Radiation on Materials: 19th International Symposium, ASTM, 2000.
[37] J. Rensman et al., J. Nucl. Mater. 307 (2002) 245.
[38] A.F. Rowcliffe et al., J. Nucl. Mater. 263 (1998) 1275.
[39] J. Russell Hawthorne, in: Effects of Radiation on Materials, ASTM, 1985, p. 580.
[40] V.V. Rybin, I.P. Kursevich, A.N. Lapin, J. Nucl. Mater. 263 (1998) 1324.
[41] J.L. Séran et al., J. Nucl. Mater. 215 (1994) 588.
[42] J.L. Séran et al., in: Effects of Radiation on Materials: 15th International Symposium, ASTM, 1992.
[43] V.K. Shamardin et al., J. Nucl. Mater. 307 (2003) 229.
[44] K. Shiba, JAERI Database.
[45] K. Shiba et al., Properties of the Reduced Activation Ferritic Steels F82H IEA Heat, JAERI, 1997.
[46] T. Shibayama, A. Kimura, H. Kayano, J. Nucl. Mater. 237 (1996) 270.
[47] P. Spatig et al., J. Nucl. Mater. 263 (1998) 1345.
[48] E. van Osch et al., in: Effects of Radiation on Materials: 19th International Symposium, ASTM, 2000.
[49] E.V. van Osch et al., in: Tensile Properties of 2.5 dpa 300 °C Neutron Irradiated RAFM Plate, Powder HIP and EB and TIG Weld, NRG-Petten, 2002.
[50] D.J.C. MacKay, Network 6 (3) (1995) 469.
[51] R.G. Carter, N. Soneda, Workshop on dose rate effects in reactor pressure vessel materials, in: Electric Power Research Institute Conference Proceedings, EPRI, 2002.
[52] G.R. Odette, T. Yamamoto, D. Klingensmith, Philos. Mag. 85 (4–7) (2005) 779.
[53] F. Abe et al., J. Nucl. Mater. 191 (1992) 845.
[54] J.P. Hirth, J. Lothe, Theory of Dislocations, 2nd Ed., John Wiley & Sons, 1982.
[55] S. Murakami, A. Miyazaki, M. Mizuno, J. Eng. Mater. Tech. 122 (1) (2000) 60.
[56] T. Yamamoto et al., Compilation and preliminary analysis of a irradiation hardening and embrittlement database for 8Cr martensitic steels, ORML report DOE/ER-0313/35, 2003, ORNL, p. 100.
[57] A.D. Brailsford, J. Nucl. Mater. 84 (1979) 245.
[58] A. Hishinuma et al., J. Nucl. Mater. 258–263 (1998) 193.
[59] M.J. Makin, F.J. Minter, Acta Met. 8 (1960) 691.
[60] A. Kohyama et al., J. Nucl. Mater. 233–237 (1996) 138.
[61] R.L. Klueh, J.M. Vitek, J. Nucl. Mater. 150 (1987) 272.
[62] C.M. Bishop, Neural Networks for Pattern Recognition, 1st Ed., Clarendon, Oxford, 1995.