Application of Bayesian Neural Network to Materials Diagnosis and Life Assessment†

Hidetoshi FUJII*, Harshad K.D.H BHADEASHIA and Kiyoshi NOGI *****

Abstract

The purpose of this paper is to introduce some examples of the application of Bayesian neural network to materials diagnosis and life assessment. The concept of Bayesian inference is added to the traditional neural network, enabling its reliable application to problems of materials diagnosis and life assessment. As an example, the fatigue crack growth rate of nickel base superalloys has been modeled as a function of some 51 variables, including stress intensity range ΔK, log ΔK, chemical composition, temperature, grain size, heat treatment, frequency, load waveform, atmosphere, R-ratio, the distinction between short crack growth and long crack growth, sample thickness and yield strength. The Bayesian method puts error bars on the predicted value of the rate and allows the significance of each individual factor to be estimated. In addition, it has been possible to estimate the isolated effect of particular variables such as the grain size, which cannot in practice be varied independently.

KEY WORDS: (Material Diagnosis) (Life Assessment) (Neural Network) (Bayesian Inference) (Fatigue Cracks) (Nickel-Base Superalloys)

1. Introduction

Any diagnosis and life assessment method for materials must to some extent depend on a knowledge of the current state and how that might change with future service. The modeling of changes is particularly difficult in the context of engineering materials because the properties can depend on a vast array of potential variables. Such a problem is ideal for neural networks which are capable of realizing a variety of non-linear relationships of considerable complexity. Although the training of the network can be very computer intensive, the trained network can be utilized in real time.

A number of attempts have been made to use a neural network as a control device in applications such as electrical appliances and the control of the welding process. There are, however, difficulties because of statistical noise. In order to redress this, the authors have recently introduced the Bayesian neural network in the field of the materials diagnosis and life assessment[1]. The method yields error bars which are dependent on the reliability of the model in specific set of variavles of interest in the input space.

The method is introduced in the context of some work on the fatigue life of nickel base superalloys. Superalloys have been used in aeroengine gas turbines for about 50 years. The weight percentage of the superalloys in engines has increased over the years to 30%. A typical loading cycle comprises starting up, takeoff and climb, cruising, landing and shut-down. The highest stresses are experienced in the bore of the disc early in the flight cycle, generally while it is in the lower temperature range 200-300°C. Stress in the rim region is lower, but at a higher temperature, 500-600°C[2].

Thus, the requirements for fatigue crack propagation are dependent on the environment where the material is used. Quantification of fatigue crack growth rate is essential in life prediction. The neural network model is more effective than linear regression analysis. When combined with sound statistical and metallurgical theory, the method allows an estimation of fatigue crack growth rate in nickel base alloys and its dependence on particular

† Received on May 19, 1997
* Research Associate
** Reader, University of Cambridge
*** Professor

Transactions of JWRI is published by Joining and Welding Reseach Institute of Osaka University, Ibaraki, Osaka 567, Japan.

variables in isolation. The latter may not be possible to measure in experimental studies where it is rare that a single quantity can be varied in isolation.

2. Neural Networks with Bayesian Statistics

2.1. Biological neuron

The neural network is literally based on the structure of a neuron. **Figure 1** shows an outline of a biological neuron. In the structure of a neuron, many synapses are connected to one soma. When the sum of the stimuli from the synapses exceeds a critical value, the soma sends a pulse to the next soma. For this structure, McCulloch-Pitts[3] suggested the first model of a neuron in 1943 using the following equation:

$$x(t+1) = 1[\sum_j w_{ij} x_j(t) - \theta_i] \qquad (1)$$

where w is the strength of the connection of a synapse to the soma and θ is the critical value for the transmission of a pulse.

1[x] is a step function and can be described as shown in **Fig.2**. However, nowadays non-linear functions such as the sigmoid function is generally used for an artificial neural network because it is a continuous, non-linear function and thus easy to use.

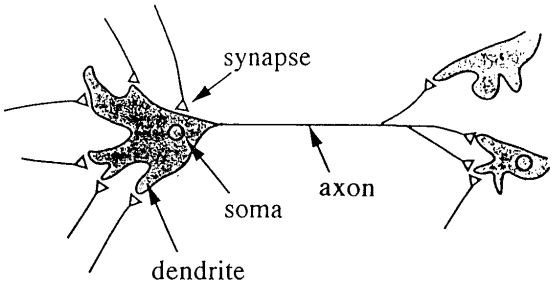

Fig.1 Biological neuron.

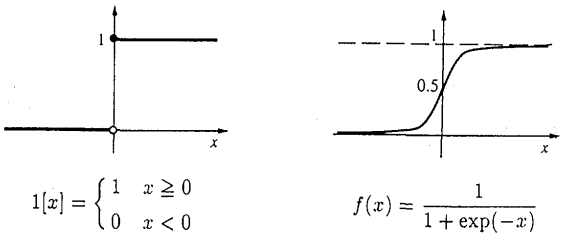

Fig.2 Functions for neural network.

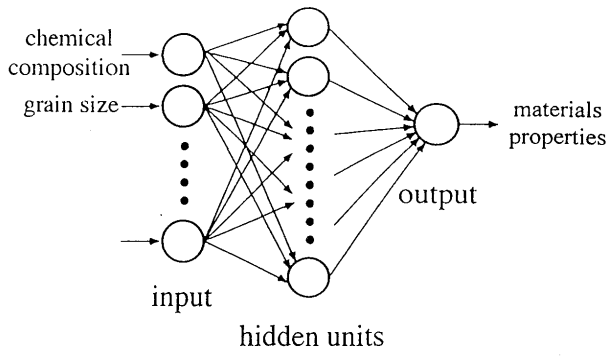

Fig.3 Neural network model used in this study.

2.2 Artificial neural network

The framework used in this study consists of the two ideas of neural network modeling and Bayesian statistics[4-6]. Neural networks are non-linear parallel computational devices inspired by the structure of the brain[3].

Figure 3 shows the structure of the neural network used in our model. Experimental conditions and other factors such as chemical composition, grain size, heat treatment and temperature are input from the left hand side. To predict the output, that is, materials properties, hidden units were used between the inputs and the output so that more complex relationships could be expressed. The transfer function relating the inputs to the *ith* hidden units is given by

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

The relationship between the hidden units and the output is linear:

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

The coefficients w and bias θ of these equations are determined in a such way as to minimize the energy function, as explained later.

3. The Analysis

The neural network is implemented in a such way as to minimize the difference between the values in a data base and the predicted values by the neural network. In order to make a better neural network, more data are required. The data base used in this study consists of 1894 combinations of fatigue crack growth and 51 inputs including the stress intensity factor ΔK, chemical

composition, temperature, grain size, the condition of heat treatment, frequency, loading condition, atmosphere, R-ratio, load waveform, sample thickness and yield stress. All these data are from the published literature[2,7-39]. The details of the data used for the predictions is shown the literature[1]. Log ΔK is included as an input factor because a factor which has a metallurgical meaning helps to find the optimum relationship between the input and the output.

Both the input and output variables were first normalized within the range ±0.5 as follows:

$$x_N = \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 0.5 \qquad (4)$$

where x_N is the normalized value of x, x_{max} is the maximum value and x_{min} is the minimum value of each variable of the original data. This normalization is not essential to the neural network approach but allows a convenient comparison of the influence of individual input variables on outputs.

Using the normalized data, the coefficients (weights) w and bias θ were determined in such a way as to minimize the following energy function[4]:

$$M(\mathbf{w}) = \alpha E_D + \beta E_W \qquad (5)$$

The minimization was implemented using a variable metric optimizes[40]. The gradient of $M(w)$ was computed using backpropagation algorithm[41]. The energy function consists of the error function, E_D and regularization E_W. The error function is the sum squared error as follows:

$$E_D(\mathbf{w}) = \frac{1}{2}\sum_m (y(\mathbf{x}^m;\mathbf{w}) - t^m)^2 \qquad (6)$$

where $\{x^m, t^m\}$ is the data set. x^m represents the inputs and t^m the targets. The m is a label of the pair. The error function E_D is smallest when the model fits the data well, i.e., when $y(x^m;w)$ is close to t^m. The coefficients w and biases θ shown in eqs.(2) and (3) make up the parameter vector w. A number of regularizers $E_{w(c)}$ are added to the data error. These regularizers favor functions $y(x;w)$ are smooth functions of x. The simplest regularization method uses a single regularizer $E_w = 1/2\Sigma w_i^2$. Here, however, we have used a slightly more complicated regularization method known as the Automatic relevance determination model[42]. Each weight is assigned to a class c depending which neurons it connects. For each input, all the weights connecting that input to the hidden units are in a single class. The hidden units' biases are in another class, and all the weights from the hidden units to the outputs are in a final class. $E_{w(c)}$ is defined to the sum of the squares of the weights in class c[5].

$$E_{w(c)}(\mathbf{w}) = \frac{1}{2}\sum_{i\in c} w_i^2 \qquad (7)$$

This additional term favors small values of w and decreases the tendency of a model to 'overfit' noise in the data set. The control parameters α_c and β together with the number of hidden units determine the complexity of the model. These hyperparameters define the assumed Gaussian noise level $\alpha_\upsilon^2 = 1/\beta$ and the assumed weight variances $\sigma_{w(c)}^2 = 1/\alpha_{(c)}$. α_υ is the noise level inferred by the model. The parameter α has the effect of encouraging the weights to decay. Therefore, a high value of σ_w implies that the input parameter concerned explains a relatively large amount of the variation in the output. Thus, σ_w is regarded as a good expression of the significance of each input though not of the sensitivity of the output to that input. The values of the hyperparameters are inferred from the data using the Bayesian methods of ref.42. In this method, the hyperparameters were initialized to values chosen by the operator, and the weights were set to small random initial values (Gaussian with mean 0 and standard deviation 0.3). The objective function $M(w)$ was minimized to a chosen tolerance, then the values of the hyperparameters were updated using a Bayesian approximation given in ref.4. The $M(w)$ was minimized again, starting from the final state of the previous optimization, and the hyperparameters were updated again, repeating 8 times.

As the number of hidden units increases, the difference (σ_υ) between predicted values and experimental values decreases monotonically, as shown

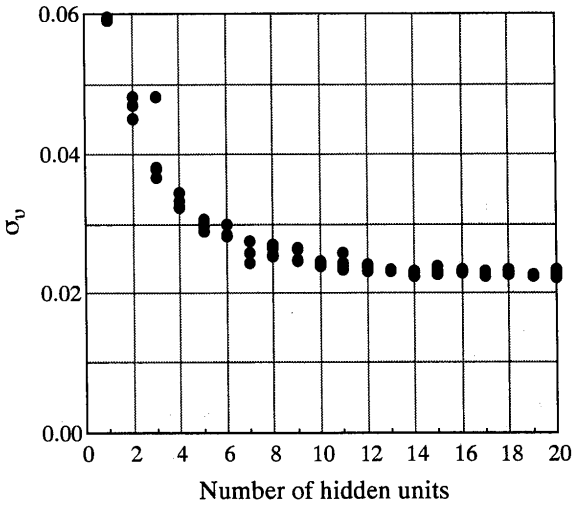

Fig.4 Variation of σ_υ as a function of number of hidden units.

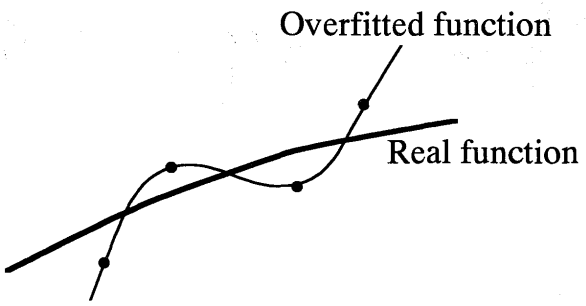

Fig.5 Overfitting of function.

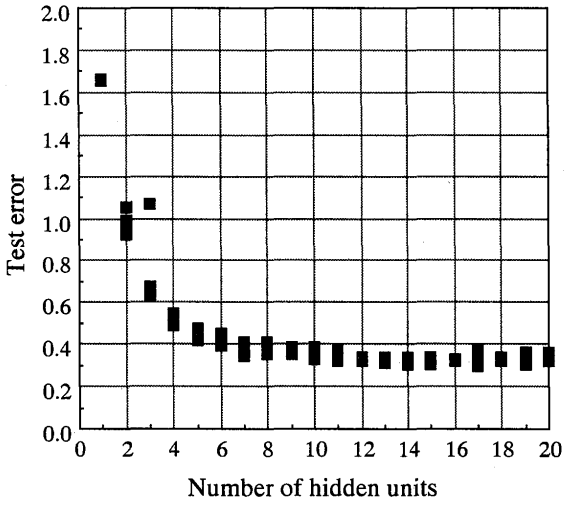

Fig.6 Variation of test error as a function of number of hidden units.

in **Fig.4**. More complex relations can be modeled with a larger number of hidden units. However, the function may then be overfitted, as shown in **Fig.5**, because experimental data always contain errors. In order to reduce overfitting, the test error (the value of the error function for a non-trained data set) was measured, using 947 randomly chosen rows of data which were not included in the training set. **Figure 6** shows the change in the test error as a function of the number of hidden units. There is a minimum at 17 hidden units. The increase in the test error over 17 hidden units indicates that the function may be overfitted. However, the increase is very small. This indicates that the distribution of data is close to the assumed Gaussian distribution and Bayesian modeling worked well. In principle, when the Bayesian modeling is completely optimized, an infinite number of hidden units could be used without overfitting[43].

Figure 7 shows the level of agreement obtained for both the training and test data respectively. In this model, in the test data set, too, the predicted data fit the

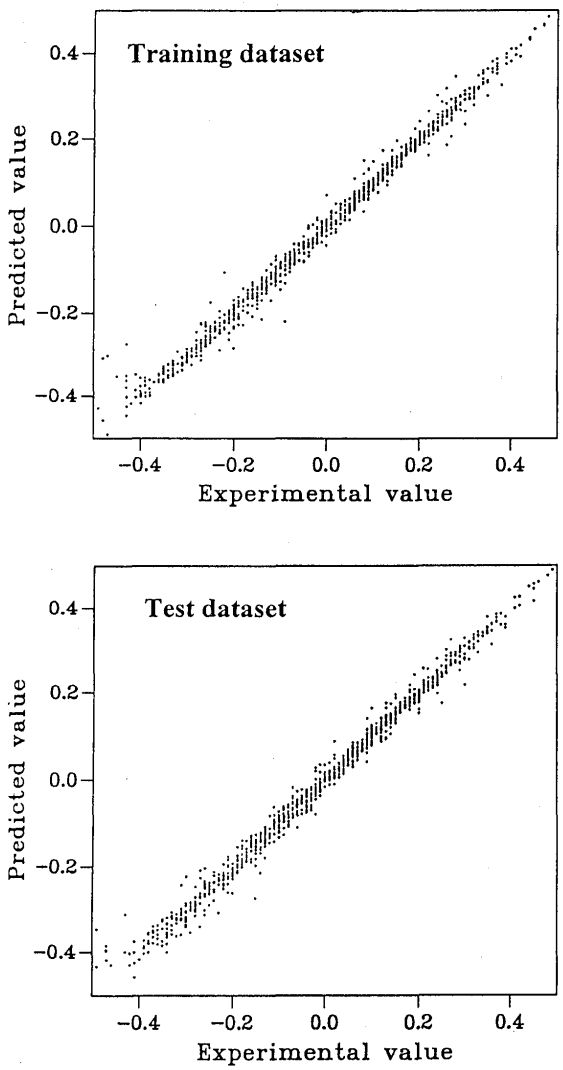

Fig.7 Comparison between predicted and experimental fatigue crack growth rate.

experimental data very well. This indicates that this model can predict the fatigue crack growth rate precisely.

4. Significance of Individual Factors

The neural network model allows an estimation of the "significance" of individual factors in influencing the fatigue crack growth rate using the value of σ_w. A higher value of σ_w implies that the input parameter concerned explains a relatively large amount of the variation in fatigue crack growth rate in the data set. Note that it is not an indication of the sensitivity of fatigue crack growth rate for that input parameter.

As shown in **Fig.8**, logΔK is clearly more strongly linked to fatigue crack growth rate than ΔK on its own. This result coincides with our metallurgical knowledge including Paris' law[44]. Note that these physical

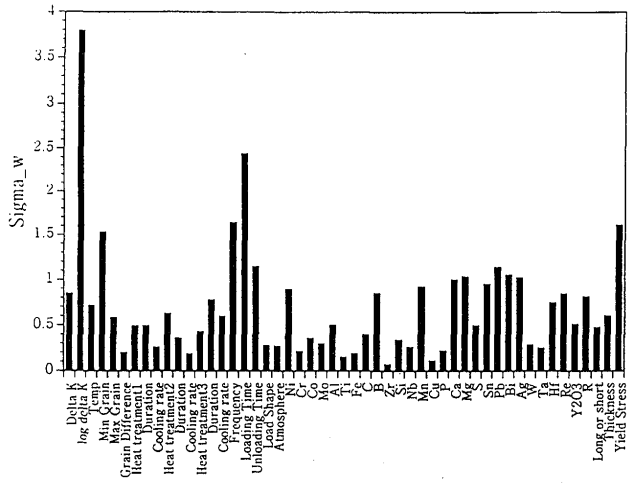

Fig.8 Significance of individual factors on fatigue crack growth rate.

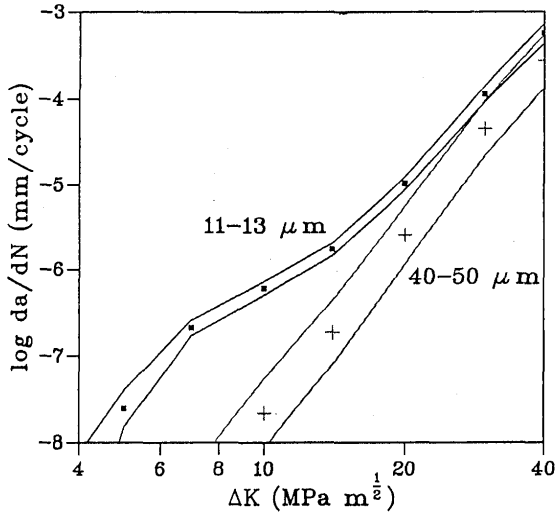

Fig.9 Effect of grain size on fatigue crack growth rate in Astroloy.

relationships were found only from the data and neural network analysis. As expected, the fatigue crack growth rate was found to be sensitive to the minimum grain size, frequency, loading time and yield stress.

5. Effect of Individual Factors

The neural network permits the effect of each factor to be examined individually, which may be impossible to do experimentally.

Figure 9 shows the estimated fatigue crack growth rate of Astroloy at room temperature; the experimental data are from the published leterature[14,15], summarized in **Table1**. The error bar is one sigma, or standard deviation, which means a 67% safety margin. As expected[16] the

Table 1. Main experimental condition inputted for the prediction of Fig.9.

Grain size	11-13μm	40-50μm
1st step heat treatment	1377K, 4h, AC	1423K, 4h, AC
2nd	923K, 24h, AC	923K, 24h, AC
3rd	1033K, 8h, AC	1033K, 8h, AC
Temperature	293K	293K
Atmosphere	Air	Air
R-ratio	0.1	0.1
Frequency	40Hz	40Hz
Yield Strength	1021MPa	954MPa

AC: air cooling

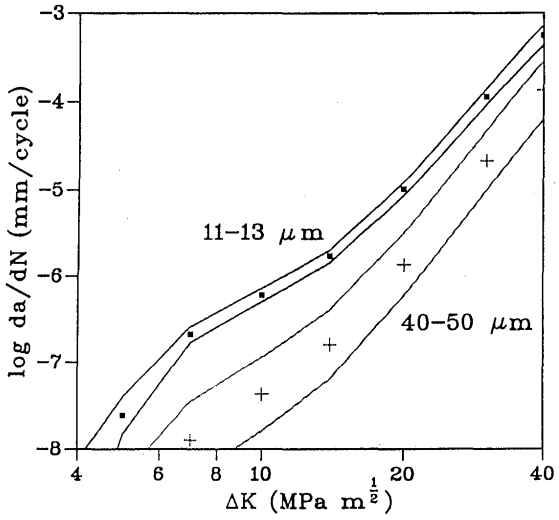

Fig.10 Effect of grain size alone on fatigue crack growth rate.

fatigue crack growth decreases when the grain size increases. Increasing grain size tends to produce more heterogeneous slip[45], resulting in a reduction of fatigue crack growth.

In practice, changes in the grain size are usually achieved by heat treatment. This was the case for the experimental data shown in Fig.9. Of course, heat treatment may affect other features within the material. In fact, in addition to increasing grain size, the heat treatment reduced the yield stress from 1021 MPa to 954 MPa. In the neural network model, the grain size alone can be changed, as shown in **Fig.10**, without altering any of the other inputs. Figure 10 is in this respect proof that an increase in grain size causes a reduction in the fatigue crack growth rate.

In fact, the fatigue crack growth rate increases when only the value the yield stress is changed from 1021 MPa to 954 MPa, as shown in **Fig.11**. In Paris regime, it is considered that fatigue behavior is dependent on crack-tip strain range, or the range of crack opening displacement Δd per cycle[46].

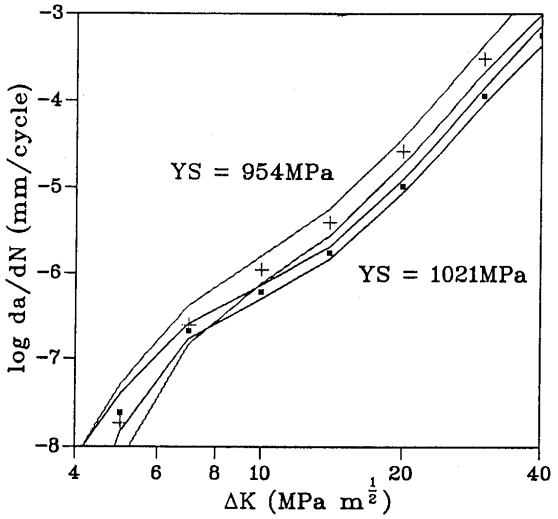

Fig.11 Effect of yield strength alone on fatigue crack growth rate.

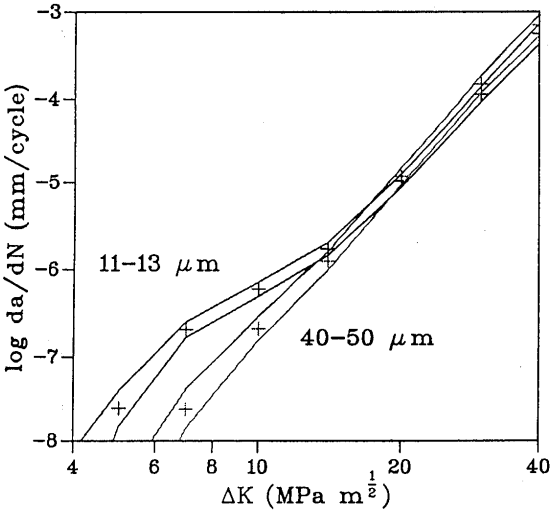

Fig.12 Effect of heat treatment alone on fatigue crack growth rate.

$$\Delta d = Q\frac{\Delta K^2}{\sigma_{cy}E} \tag{8}$$

where Q is constant and σ_{cy} should be the cyclic yield stress (though, monotonic σ_y is often used as an approximation)[2], Accordingly

$$\frac{da}{dN} \propto Q\frac{\Delta K^2}{\sigma_{cy}E}. \tag{9}$$

Thus, when the yield stress is reduced, the fatigue crack growth should increase. This relation indicates that the effect of yield stress alone, shown in Fig.10, is reasonably predicted.

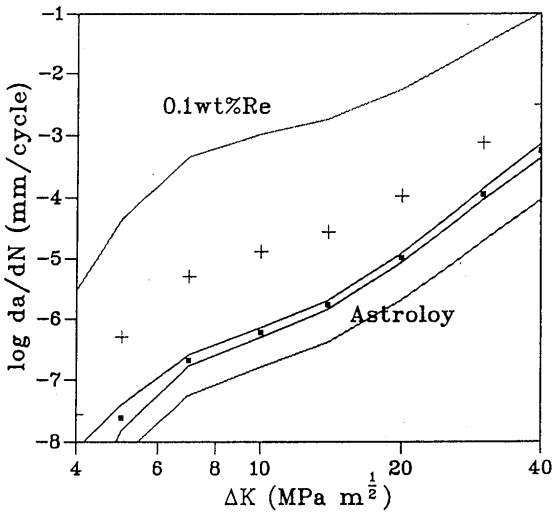

Fig.13 Prediction of effect of addition of 0.1mass% Re.

On the other hand, the case where only the heat treatment is changed, without altering the grain size or yield strength is illustrated in **Fig.12**. Thus, any changes in the crack growth rate should indicate the presence of other factors affecting fatigue crack growth. There is a clear change in the near-threshold region which is the part most sensitive to microstructure.

The effect shown in Fig.12 might therefore be due to the higher solution treatment reducing coarse γ'. Coarse γ' particles have interfaces with the matrix which contain dislocations[14]. Our model does not include direct microstructure because such data have not been reported frequently. This would be a good area for future work.

One can also predict the fatigue crack rate under unusual conditions. The effect of adding 0.1 mass% Re to Astroloy is illustrated in **Fig.13**. Since the error bars are large because of the absence of experimental data, it cannot be concluded that Re really affects fatigue crack growth. This is a good example of the safety of the predictions made by the model, in that the error bars are large when the model is uncertain.

6. Committee Model

A method involving an assessment of predictions from a variety of models (i.e. a committee of models) has been proposed to improve the reliability of predictions[1] though in the present study a simpler model was used.

The same data can be empirically modeled in many ways, for example by varying the number of hidden units or starting value of σ_w. The variety of models thus produced can be ranked according to the magnitude of the test error. The best individual model would then have the

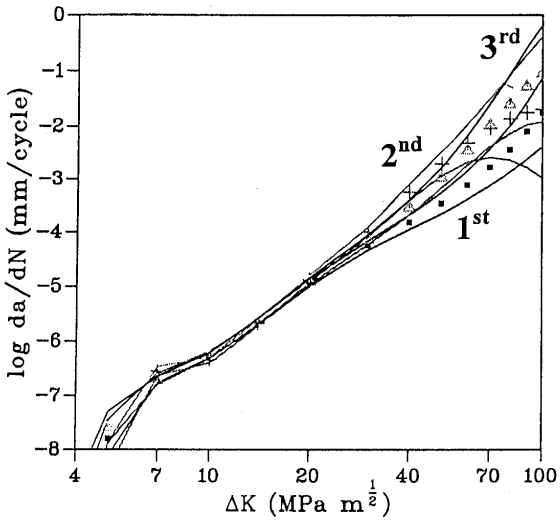

Fig.14 Comparison of best, second best and third best model.

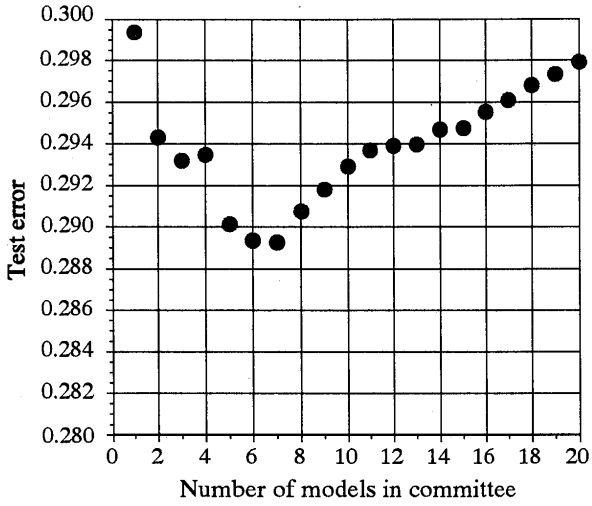

Fig.15 Change in test error with the number of models in the committee.

minimum test error. However, it is possible in principle to reduce the test error further by using the average of predictions from a number of models, i.e. a committee of models.

The test errors of the variety of models produced are not very different from that of the best model, as can be seen in **Fig.6**. This indicates that these models should lead to a similar predictions. However, for some choice of input variables the predictions are nevertheless different because of the limitations in the training data. **Figure 14** shows the predictions using the three best models for Astroloy at room temperature (The main experimental conditions[14,15] are summarized in the first column in Table 1.) When ΔK is less than 40MPa $m^{1/2}$, the difference in the predictions is very small, but when ΔK is more than 40MPa $m^{1/2}$, the difference is large. In the latter region, the error bars are also large, indicating that there are insufficient or imprecise data in that region of input space. The error bars in this case are for 67% confidence.

A 'committee' model was therefore introduced in order to see whether more reliable predictions to be made. The method is as follows:

(1) The individual models are first ranked via their test errors.

(2) A committee of N models is then formed by combining the best N models, where N=1,2,3.....

The mean prediction $\bar{y}$ of the committee is

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \qquad (10)$$

and the associated error σ in $\bar{y}$ is given by

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 + \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2 \qquad (11)$$

Figure 15 shows the decrease in test error by combining models. A committee of seven of the best individual models has a minimum best error. As shown in this figure, the use of the committee model can reduce the test error by 3% compared with the best individual network.

7. Conclusion

(1) As an example of material diagnosis and life assessment, a neural network in a Bayesian framework was applied successfully to estimate the fatigue crack growth rate of nickel base superalloys. The results are found to be consistent metallurgical experience.

(2) The model can be used to examine the effects of each variable in isolation. As a result, it was confirmed that logΔK is more strongly linked to the fatigue crack growth rate than to ΔK, as expected from the Paris Law. Similarly, it was possible to determine the effect of grain size alone. It was confirmed that an increase in the grain size should lead to a decrease in the fatigue crack growth.

(3) The range of the error bar changes in each case predicted. When the data are noisy or nonuniform, the error bar indicated tends to be larger. This gives confidence in the use of the model.

References

1) H.Fujii, D.J.C.MacKay, H.K.D.H.Bhadeshia: ISIJ int. **36**, (1996) 1373-1382.
2) J. E. King: *Mater. Sci. Thechnol.* **3**, (1987) 750-764.
3) W. S. McCulloch, W. H. Pitts: *Bull. Math. Biophys.* **5**, (1943) 115-133.
4) D. J. C. MacKay: *Neural Computation* **4**, (1992) 415-447.
5) D. J. C. MacKay: *Neural Computation* **4**, (1992) 448-472.
6) D. J. C. MacKay: *Neural Computation* **4**, (1992) 720-736.
7) J. E. King, R. A. Venables, M. A. Hicks, The Effect of Microstructure, Temperature and R-ratio on Fatigue Crack Propagation and Threshold Behaviour in Two Ni-Base Alloys., S. R. Valluri, D. M. R. Taplin, P. Rama Rao, J. F. Knott, R. Dubeys, Eds., 6th Int. Conf. on Fracture (ICF6) New, Dehli, India, 1984), pp. 2081.
8) L. Grabowski, J. E. King: *Fatigue Fract. Engng. Mater. Struct.* **15**, (1992) 595-606.
9) M. R. Winstone, K. M. Nikbin, G. A. Webster: *J. Mater. Sci.* **20**, (1985) 2471-2476.
10) J. L. Yuen, P. Roy: *Scripta Metall.* **19**, (1985) 17-22.
11) S. P. Lynch, T. C. Radtke, B. J. Wicks, R. T. Byrnes: *Fatigue Fract. Engng. Mater. Struct.* **17**, (1994) 313-325.
12) H. Ghonem, D. Zheng: *Mater. Sci. Eng.* **A150**, (1992) 151-160.
13) M. Clavel, A. Pineau: *Metall. Trans. A* **9A**, (1978) 471-480.
14) J. E. King: *Metal Science* **16**, (1982) 345-355.
15) C. W. Brown, J. E. King, M. A. Hicks: *Metral Science* **18**, (1984) 374-380.
16) J. Gayda, V. Miner: *Metall. Trans.* **14A**, (1983) 2301-2308.
17) F. Gabrielli, M. Marchionni, G. Onofrio, Time Dependent Effects on High Temperature Low Cycle Fatigue and Fatigue Crack Propagation of Nickel Base Superalloys, K. Salama, K. Ravi-Chandar, D. M. R. Taplin, P. Rama Raos, Eds., Advance in Fracture Research (Pergamon Press, Huston, Texas, 1989), vol. 2, pp. 1149-1163.
18) J. L. Yuen, P. Roy, W. D. Nix: *Metall. Trans. A* **15A**, (1984) 1769-1775.
19) K. Sadananda, P. Shahinian: *J. Eng. Mater. Tech.* **100**, (1978) 381-387.
20) A. Sengupta, S. K. Putatunda: *Scripta Metall. Mater.* **31**, (1994) 1163-1168.
21) K. Sadananda, P. Shahinian: *Metals Tech.*, (1982) 18-25.
22) M. Clavel, A. Pineau: *Metall. Trans.* **9A**, (1978) 471-480.
23) M. Clavel, C. Levaillant, A. Pineau: *Creep-Fatigue-Environment-Interactions* , (1980) 24-45.
24) F. Gabrielli, R. M. Pelloux: *Metall. Trans. A* **13A**, (1982) 1083-1090.
25) F. Gabrielli, G. Vimercati, V. Lupinc, Environmental Effects on High Temperature Fatigue Crack Growth Behaviour of Nickel Base Superalloys, Mechanical Behavior of Materials - V (Pargamon Press, China, 1987).
26) P. Shahinian: *Metal Tech.* **5**, (1978) 372-380.
27) K. Sadanada, P. Shahinian: *Eng. Fract. Mech.* **11**, (1979) 73-86.
28) K. Sadananda, P. Shahinian: *J. Mater. Sci.* **13**, (1978) 2347-2357.
29) K. Sadananda, P. Shahinian, Crack Growth under Creep and Fatigue Conditions, R. M. Pelloux, N. S. Stoloffs, Eds., Creep-Fatigue-Environment Interactions (Metal Soc. of AIME, MIlwaukee, Wisconsin, 1980), pp. 86-111.
30) S. P. Lynch, T. C. Radtke, B. J. Wicks, R. T. Byrnes: *Fatigue Fract. Engng. Mater. Sturct.* **17**, (1994) 297-311.
31) J. S. Crompton, J. W. Martin: *Metall. Trans. A* **15A**, (1984) 1711-1719.
32) L. A. James, W. J. Mills: *Eng. Fract. Mech.* **5**, (1985) 797-817.
33) R. A. Venables, M. A. Hicks, K. J.E., Influence of Stress Ratio on Fatigue Thresholds and Structure Sensitive Crack Growth in Ni-Base Superalloy, D. Davidsons, Ed., Proc. Int. Symp. on Fatigue Crack Growth Threshold Concepts (AIME, 1983), pp. 341-357.
34) R. A. Venables, J. E. King, The Effect of Dynamic Strain Ageing on Fatigue Crack Propagation and Threshold Behavior in a Ni-Base Superalloy, Fatigue 84 1984), pp. 1371-1378.
35) M. A. Hicks, J. A. King: *Int. J. Fatigue* **5**, (1983) 67-74.
36) J. L. Yuen, C. G. Schmidt, P. Roy: *Fatigue Fract. Engng. Mater. Struct.* **8**, (1985) 65-76.
37) J. C. Healy, L. Grabowski, C. J. Beevers: *Fatigue Fract. Engng. Mater. Struct.* **15**,(1992) 309-321
38) B. A. Cowles, D. L. Sims, J. R. Warren, R. V. Miner Jr.: *Trans. ASME* **102**, (1980) 356-363.
39) B. Lawless, S. D. Antolovich, C. Bathias, B. Boursier, The Effect of Microstructure on the FCP and Overload Behavior of Waspaloy at Room Temperature, J. M. Wells, J. D. Landess, Eds., Fracture: Interaction of Microstructure, Mechanisms and Mechanics (Metall. Soc. AIME, Los Angeles, California, 1984), pp. 285-301.
40) W.H.Press, S.A.Teukolsky, W.T.Vetterling and B.P.Flannery: Neumerical Recipes in C, 2nd edition, Cambrige University Press, 418.
41) D. E. Rumelhart, G. E. Hinton, R. J. Williams: *Nature* **323**, (1986) 533.
42) D. J. C. MacKay: *Network: Computation in Neural Systems* **6**, (1995) 469.
43) R. M. Neal: Radford Thesis 'Bayesian Learning for Neural Networks' (1995) Dept. of Computer Science, Univ. of Toronto.
44) P. C. Paris, F. E. Erdogan: *J. Basic Eng. (Trans. ASME, D)* **85**, (1963) 528.
45) B. A. Lerch, Jayaraman: *Mater. Sci. Eng.* **66**, (1984) 151-166.
46) J.F.Knott: 'Fundamental Mechanics' (1973) 240, London, Butterworths.