

Comparison of Artificial Neural Networks with Gaussian Processes to Model the Yield Strength of Nickel-base Superalloys

F. TANCRET, H. K. D. H. BHADSHIA and D. J. C. MACKAY¹⁾

Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

1) Department of Physics, Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK.

(Received on March 19, 1999; accepted in final form on April 26, 1999)

The abilities of artificial neural networks and Gaussian processes to model the yield strength of nickel-base superalloys as a function of composition and temperature have been compared on the basis of simple well-known metallurgical trends (influence of Ti, Al, Co, Mo, W, Ta, of the Ti/Al ratio, γ' volume fraction and the testing temperature). The methodologies are found to give similar results, and are able to predict the behaviour of materials that were not shown to the models during their creation. The Gaussian process modelling method is the simpler method to use, but its computational cost becomes larger than that of neural networks for large data sets.

KEY WORDS: neural network; superalloys; Gaussian process; modelling.

1. Introduction

Modern metallic materials usually contain a large number of alloying elements added to satisfy an array of design criteria. The influence of individual solutes on their mechanical properties can easily be measured and understood under given conditions; sometimes simple interactions between two or three elements can be rationalised, but taking all the interactions into account at the same time is generally not possible. Recent papers^{1–3)} have emphasised the possibility to use “learning machines”, such as artificial neural networks (ANN), to model the behaviour of complex materials as a function of their composition and/or thermomechanical treatments. Artificial neural networks perform a non-linear parametric regression of an output (yield strength, ultimate tensile strength, elongation, creep properties...) as a function of a number of inputs (composition, thermomechanical treatments, temperature, etc.). For the modelling to be effective, one must use a database containing a large number of measurements, covering a wide range of alloy compositions. ANN are “trained” on the dataset by fitting a complex parametrical function, consisting of a weighted sum of hyperbolic tangents, to the data. Training involves an optimisation of the parameters through a progressive reduction in the sum of squared errors derived by comparing predictions against the measured values of the output parameter. This error tends to become smaller as the complexity of the fitting function increases. There is thus a danger of overfitting the data and creating models which generalise badly. To control the complexity of the function produced by the network, additional terms, called regularizers or weight decay functions, are added to

the sum of squared errors. These regularizers penalise functions of high complexity, with the strength of the penalty being controlled by a set of parameters called hyperparameters or weight decay rates. In the Bayesian interpretation of neural networks optimization, these hyperparameters control the prior probability distribution over functions. They express the degree of non-linearity expected in the function. The Bayesian framework allows the complexity of a model to be controlled automatically.⁴⁾

While it is possible to optimise the parameters and hyperparameters of just one neural network, it is usually found that better results are obtained by optimising multiple neural networks differing in the number of parameters and in the random initial condition of the optimisation. A subset of these networks is selected to form a committee of models, the average of whose predictions is taken. The selection of the models in the committee is made by testing the performance of each network on a validation set, that is, a set from the data points held out from the training set. Because of all these procedures, the creation of good neural network predictions involves the study of a large number of models, and subsequently, the choice of a good committee, making the entire process very laborious.

An alternative, similar, but simpler approach is the Gaussian process method.^{5,6)} As in the Bayesian neural network model, a prior distribution is assumed on the function underlying the data, but instead of defining this prior distribution in terms of a prior on the parameters of the function, we put a Gaussian process prior directly on the function itself. This prior expresses smoothness properties and long-range correlations that the function is expected to have. Given such a prior, and assuming a

noise model that is Gaussian, predictions of new data values, given a data set of size N , can be obtained by inversion of an $N \times N$ matrix. Typically, we will not know precisely what smoothness properties or complexity to expect the function to have, so, as in the neural network method, we introduce hyperparameters that control the expected complexity of the function. And, as in the neural network method, these hyperparameters can be inferred from the data using Bayesian methods. The results obtained from a Gaussian process optimised in this way are usually found to be satisfactory without the creation of a committee of Gaussian process models, so the need for a validation set is eliminated.

However, while ANN have proved to be very efficient in modelling the behaviour of metallic alloys,¹⁻³⁾ Gaussian processes have rarely been used in metallurgy,^{7,8)} and their ability to recognise metallurgical laws is still unknown.

The aim of this paper is to compare the performance of GP with that of ANN in modelling the tensile properties of nickel-base superalloys over a wide range of compositions, and to assess their respective abilities to recognise well-known metallurgical trends, as a function of test temperature, solute elements, *etc.* Finally their ability to perform extrapolations will also be discussed.

2. The Database

The database consists of the yield strength of thirty-six INCO wrought nickel-base superalloys tested in traction between 21°C and 1093°C⁹⁾ (Astroloy; D-979; Hastelloy X; Inconel 600, 601, 617, 625, 690, 706 and X750; Unitemp AF2-1D; IN587; M-252; Nimonic 75, 80A, 81, 90, 105, 115, 263, 942, PE11, PE16 and PK33; Pyromet 860; Rene 41 and 95; RGT 4 and 13; Udimet 400, 500, 520, 630, 700 and 710; Waspaloy). The composition range is defined in **Table 1**. The database contains a total of 192 measurements. The inputs are the concentrations in various elements (in weight %) and the test temperature. All the inputs and outputs have been normalised linearly between -0.5 and $+0.5$, corresponding respectively to the lowest and the highest value of each input or output in the database. This allows an easier visualisation of the importance of each variable in explaining changes in the output parameter.

3. ANN and GP Training and Optimisation

3.1. Artificial Neural Networks

All the results obtained with ANN come from reference 10), where a complete description of the model can be found. A committee of five ANN individually trained and optimised on one half of the data and tested on the other half had been used.

3.2. Gaussian Processes

GPs have a statistical structure that needs to be briefly exposed. Let's consider the data, D , as constituted of N L -dimensional input vectors $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N\} = [X_N]$ and corresponding outputs, or targets, $\{t_1, t_2, \dots, t_N\} = \vec{t}_N$. Consider now the joint probability distribution, in an

Table 1. Input and output (yield strength) ranges for the data.

Input	Minimum value	Maximum value	Unit	
Ni	38	76	weight %	
Cr	12	30		
Co	0	20		
Mo	0	10		
W	0	6		
Ta	0	1.5		
Nb	0	6.5		
Al	0	4.9		
Ti	0	5		
Fe	0	40		
Mn	0	0.5		
Si	0	0.5		
C	0.03	0.35		
B	0	0.16		
Zr	0	0.2		
T	21	1093		°C
Output				
Yield strength	28	1310	MPa	

N -dimensional space, of the N output values in the database given the N inputs, $P(\vec{t}_N | [X_N])$. Similarly, the joint probability distribution of both the N data points and of the single new point with input vector \vec{x}_{N+1} , for which we want to predict the output t_{N+1} , is $P(t_{N+1}, \vec{t}_N | \vec{x}_{N+1}, [X_N])$. What we are looking for is the one-dimensional probability distribution over the predicted point, $P(t_{N+1} | \vec{x}_{N+1}, D)$, given that we know the corresponding input vector, \vec{x}_{N+1} , and the data $D = \{\vec{t}_N, [X_N]\}$. The relationship between the above quantities is⁵⁾:

$$P(t_{N+1} | \vec{x}_{N+1}, D) = \frac{P(t_{N+1}, \vec{t}_N | \vec{x}_{N+1}, [X_N])}{P(\vec{t}_N | [X_N])} \dots (1)$$

In Bayesian neural networks, this joint distribution is a complex function involving integration over all possible settings of the parameters weighted by their prior probability distribution. In the Gaussian process approach we define this distribution directly to be a Gaussian process. The GP model assumes that the joint probability distribution of any N output values is a multivariate Gaussian,

$$P(\vec{t}_N | [X_N], \Theta) \propto \exp\left(-\frac{1}{2} (\vec{t}_N - \vec{\mu})^T [C_N]^{-1} (\vec{t}_N - \vec{\mu})\right) \dots (2)$$

where $\vec{\mu}$ is the mean, $[C_N]$ a covariance matrix which is a function of $[X_N]$, and Θ a set of parameters which will be discussed later. Consequently, a similar equation—with $N + 1$ variables—holds for $\vec{t}_{N+1} = (t_{N+1}, \vec{t}_N)$, and Eq. (1) reduces to a univariate Gaussian of the form⁵⁾

$$P(t_{N+1} | \vec{x}_{N+1}, D) = \frac{1}{\sqrt{2\pi\sigma_f^2}} \exp\left[-\frac{(t_{N+1} - \hat{t})^2}{2\sigma_f^2}\right] \dots (3)$$

where \hat{t} is the posterior mean and σ_f the standard deviation given by:

$$\hat{t} = [k]^T [C_N]^{-1} \vec{t}_N \dots (4)$$

and

$$\sigma_f^2 = k - [k]^T [C_N]^{-1} [k] \dots (5)$$

where

$$[k] = [C(\vec{x}_1, \vec{x}_{N+1}), C(\vec{x}_2, \vec{x}_{N+1}), \dots, C(\vec{x}_N, \vec{x}_{N+1})] \dots (6)$$

and

$$k = C(\vec{x}_{N+1}, \vec{x}_{N+1}) \dots (7)$$

Equation (3) gives the probability distribution of the output, t_{N+1} , given the set of inputs, \vec{x}_{N+1} , and the data, D . Both the prediction, \hat{t} , and its standard deviation, $\sigma_{\hat{t}}$, depend on the covariance matrix, $[C_N]$, whose elements C_{ij} are given by the covariance function, C . The form of this function is essential since it embodies our assumptions about the nature of the underlying input-output function we want to model. Basically, it defines how strongly any input will influence the value of the output and the lengthscales of trends in the function underlying the data. The covariance function we use is

$$C(\vec{x}_i, \vec{x}_j | \Theta) = \theta_1 \exp \left[-\frac{1}{2} \sum_{l=1}^L \frac{(x_i^l - x_j^l)^2}{r_l^2} \right] + \theta_2 + \sigma_n^2 \delta_{ij} \dots (8)$$

where $\Theta = \{r_l (l=1 \text{ to } L), \theta_1, \theta_2, \sigma_n\}$.

This function gives the covariance between any two outputs, t_i and t_j , with corresponding input vectors \vec{x}_i and \vec{x}_j . The closer the inputs, the smaller the exponent of the first term in Eq. (8) the larger the first term, and the stronger the outputs will be correlated, making it probable that they have similar values. This first term also includes the length scale, r_l , over which the function varies in the l th input dimension, which is an indication of the smoothness of the interpolant in this dimension. This covariance function expects there to be no long-range correlations in the data on lengthscales much bigger than r_l .

The second term, θ_2 , is an offset, allowing the functions to have a non-zero mean value. In our case, this parameter has been set to zero. The last term, $\sigma_n^2 \delta_{ij}$, is the noise model, with δ_{ij} being equal to 1 if $i=j$ and to 0 otherwise. We have thus an input-dependent noise model of variance σ_n^2 for the output, and we are assuming the inputs to be noise-free.

The set of parameters $\Theta = \{r_l (l=1 \text{ to } L), \theta_1, \theta_2, \sigma_n\}$ are called the hyperparameters because they define the probability distribution over functions rather than the interpolating function itself. They are equivalent in role to the hyperparameters in a Bayesian neural network. The hyperparameters, Θ , the dataset, $[X_N]$, \vec{t}_N , and the new input vector, \vec{x}_{N+1} , define completely the value of the prediction, \hat{t} , and of its standard deviation, $\sigma_{\hat{t}}$. The optimum values of the hyperparameters are inferred during the training of the model by maximizing the probability of the hyperparameters given the data, $P(\Theta | D)$, which is done numerically within a Bayesian framework.⁵⁾

As a matter of comparison with the ANN model, the Gaussian process program has been trained on one half of the data and tested on the other half. Whereas the ANN method uses the test error to rank the models for inclusion in a committee, the GP method used here has no need for a test set.

4. Making Predictions

Both the ANN committee model and the GP were used to perform some predictions on situations previously "unseen" by the model, and to verify some well-known metallurgical trends:

- influence of temperature between 21°C and 1 093°C on the Nimonic 115 alloy (60Ni-14.3Cr-13.2Co-3.3Mo-4.9Al-3.7Ti-0.15C-0.16B-0.04Zr, wt%).
- influence of the γ' phase volume fraction at 21°C and 700°C in the Ni-Al system.
- individual effects of Ti and Al additions on a Ni-20Cr-10Co-1Al-1Ti-0.03C wt% alloy at 21°C.
- influence of the Ti/Al atomic ratio on a Ni-13.7Cr-0.046C-0.0052B wt% alloy at 21°C when the total atomic content (Ti+Al) is kept constant and equal to 7 at%.
- effect of reducing the Co content on a Ni-19Cr-4Mo-1.3Al-3Ti-0.035C-0.03B-0.006Zr wt% alloy (Waspaloy) at 21°C and 538°C.
- individual effects of Mo, W and Ta on a Ni-20Cr-10Co-3.3Al-3.3Ti-0.03C wt% alloy at 21°C.

In general, the influences of the variables have been tested on the maximum range covered by the database, *i.e.* normalised values from -0.5 to +0.5.

5. Results and Discussion

5.1. Training and Test

Figure 1 shows comparisons between the measured yield strength and that predicted by the models (ANN and GP) for the training and the test sets. A small dispersion of the points around the " $x=y$ " line means a small error, *i.e.* a good prediction. It can be seen that in both cases the dispersion is similar.

Error bars are also very similar for ANN and GP. The error bars given by ANN include the predictive uncertainty, the typical error between the prediction and the supposed "true" value of the output, and the 1σ error due to the inferred level of noise in the database. Gaussian processes give error bars corresponding to the 1σ dispersion of the calculated probability distribution of the output, that include both the predictive uncertainty and the noise level, which is equivalent to the error bars given by the ANN model.

Consequently, ANN and GP seem to have very similar levels of confidence in predicting values of the training and test sets.

5.2. Temperature

The predicted influence of temperature on the yield strength of Nimonic 115 is presented in Fig. 2, as well as the actual measured values. Predictions indicate a slight decrease in the yield strength with temperature, then a peak and a strong drop above 800°C. This is almost the actual variation, due to a competition between the decreasing and increasing yield strengths of the γ matrix and of the γ' inclusions with temperature up to 900°C, respectively.¹¹⁾ However, there is a significant discrepancy between the ANN and GP predictions in the range 21-500°C, associated with larger error bars. Indeed, the measurements included in the

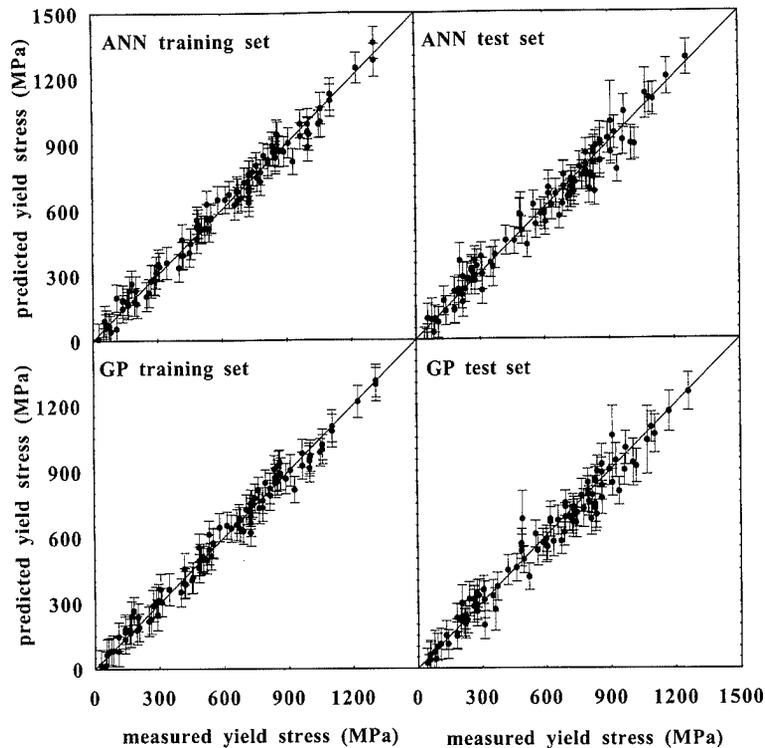


Fig. 1. Predicted versus measured yield strengths in the case of ANN and GP (training and test sets).

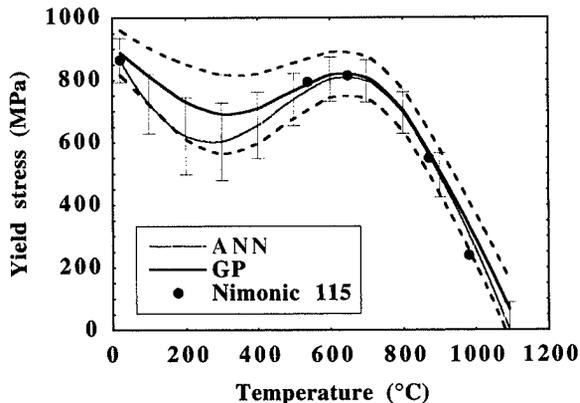


Fig. 2. Predicted and actual evolution of the yield strength of a Nimonic 115 alloy as a function of temperature (the experimental points were included in the database).

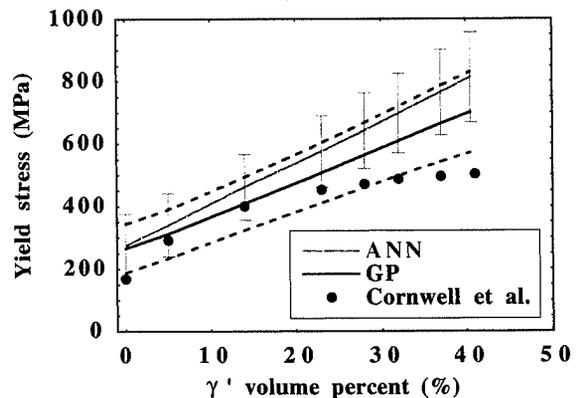


Fig. 3. Predicted and experimental influence of the γ' volume percent on the yield strength in the Ni-Al system at 21°C (experimental points from Cornwell *et al.*⁹⁾).

database have been made at 21°C and above 538°C. As a consequence, error bars are larger in the range 21–538°C, acting as a warning signal indicating that the models are not confident in this domain. Nevertheless, ANN and GP give similar results and describe correctly the actual variation in the range covered by the database.

5.3. γ' Volume Fraction

Both models predict an increase in room temperature yield strength with the γ' volume percent, V_f , in the Ni-Al system (Fig. 3). This is almost in agreement with results published by Cornwell *et al.*,¹²⁾ as well as with classical precipitation strengthening models.^{13,14)} However, the actual trend is not perfectly predicted, because the Ni-Al system is rather far from the complex commercial alloys of the database. The GP result is closer to the experimental trend than the ANN one.

5.4. γ' Forming Elements: Al and Ti

The γ' forming elements usually have the largest effect on the yield strength of nickel-base superalloys, since precipitation hardening is the most efficient strengthening mechanism in these materials. The predicted influences of Ti and Al additions on the yield strength of a Ni-20Cr-10Co-1Al-1Ti-0.03C wt% alloy at 21°C are presented in Fig. 4. Once more the predictions of ANN and GP are very similar. It can be seen that the models predict a stronger influence of titanium. This fact is confirmed by Fig. 5, where the influence of the Ti/Al ratio on the yield strength of a Ni-20Cr-10Co-0.03C wt% alloy at 21°C is plotted. In this graph, the total atomic content of Ti+Al has been kept constant and equal to 7 at%, as in the experiment of Miller *et al.*,¹⁵⁾ whose results are also presented in Fig. 5.

First, both models are less well defined for the low and high values of Ti/Al (large error bars), reflecting the

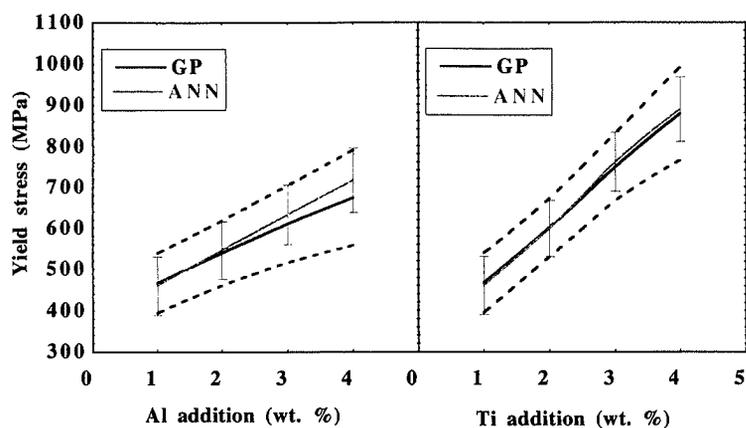


Fig. 4. Predicted individual effects of Al and Ti additions on the yield strength of a Ni-Cr-Co-Ti-Al-C alloy at 21°C.

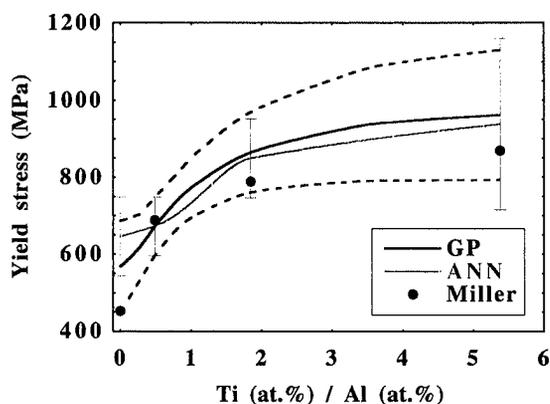


Fig. 5. Predicted and actual influence of the Ti/Al ratio on the yield strength of a Ni-Cr-Co-Ti-Al-C alloy at 21°C (experimental points from Miller *et al.*¹²⁾).

lack of data in these ranges: only a few alloys of the database contain either a lot of Ti and very little Al, or a lot of Al and no Ti.

Secondly, even if the experimental results of Miller *et al.* lie within the predicted error bars, the values are frequently lower than those predicted. This can be explained by the fact that the materials used by Miller *et al.* were simple cast alloys, whereas the alloys in the database were commercial wrought alloys whose properties have been optimized by complex heat treatments.

Finally, as both titanium and aluminum partition mainly to the γ' phase, it can be assumed, as a first approximation, that the γ' volume fraction is constant. This figure thus indicates that, for a constant γ' volume fraction, replacing an aluminum atom by a titanium atom gives an increase in the yield strength. This is in agreement with theoretical considerations.

Since titanium atoms are bigger than aluminum (+4%), they induce an increase of the γ' lattice parameter, of the γ/γ' lattice mismatch, and thus of the strain fields. Moreover, it has been shown that titanium increases the anti-phase boundary energy of the γ' phase, which renders more difficult the cutting of inclusions by dislocations.^{13,16)} Titanium is thus expected to give more effective strengthening effects than aluminum, and this has been recognised by both the artificial neural networks and the Gaussian processes.

5.5. Cobalt

The predicted role of cobalt on the yield strength of a Ni-19Cr-4Mo-1.3Al-3Ti-0.035C-0.03B-0.006Zr wt% alloy (Waspaloy) at 21°C and 538°C is presented in Fig. 6 for both ANN and GP. Once more the results are similar. Cobalt is expected to have a small influence on the yield strength of this alloy at 21°C and 538°C, in agreement with the experimental data published by Maurer *et al.*¹⁷⁾ (Fig. 6). However, although having no strengthening effect, cobalt is often necessary, in spite of its high price, to improve the hot corrosion resistance of the alloys.¹⁸⁾

5.6. The Refractories: Mo, W, Ta

The predicted effects of molybdenum, tungsten and tantalum additions to a Ni-20Cr-10Co-3.3Al-3.3Ti-0.03C alloy at 21°C are shown in Fig. 7. ANN and GP predictions are still in quite good agreement in the case of Mo and W, but the trend is not very well defined in the case of Ta. It can be seen that error bars are longer in the case of Ta and W, due to the small number of alloys in the database containing these elements. These error bars are very large in the case of Ta, indicating a low level of confidence in both cases.

Mo, W (and Ta) are predicted to induce an almost linear strengthening. However, since they partition differently to the γ and γ' phases,¹⁰⁾ their strengthening mechanisms are also different. Molybdenum partitions mainly to the γ phase, hence giving a solid solution strengthening of the matrix. The predicted effect calculated from Fig. 7 is about 40MPa/at%, which is close to 35MPa/at% found by Mishima *et al.* for Mo in pure Ni.¹⁹⁾ Tungsten partitions about equally to γ and γ' , and yields both a solid solution strengthening of the matrix¹⁹⁾ and an increase of the anti-phase boundary energy of the inclusions.²⁰⁾ The role of tantalum is more complex. It partitions mainly to the γ' phase, thus increasing its volume fraction.²¹⁾ It also increases the anti-phase boundary energy of the γ' phase^{16,20)} as well as the γ/γ' misfit by increasing the γ' lattice parameter.²²⁾ A small fraction of Ta partitions to the matrix, yielding a little solid solution strengthening effect.²¹⁾ Tantalum should thus be expected to be the most efficient strengthener, which has been inferred by the ANN,

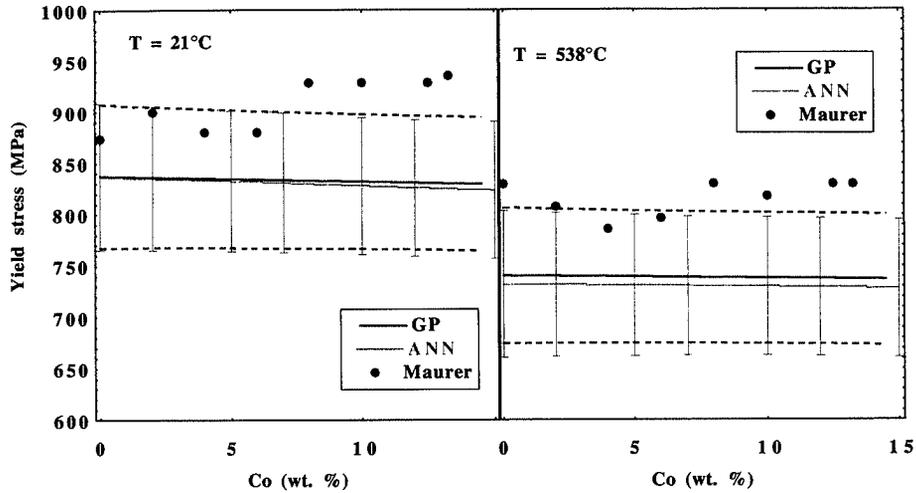


Fig. 6. Predicted and actual effects of reducing the cobalt content of a Ni-Cr-Co-Mo-Al-Ti-C-B-Zr alloy at 21°C and 538°C (experimental points from Maurer *et al.*¹⁴⁾).

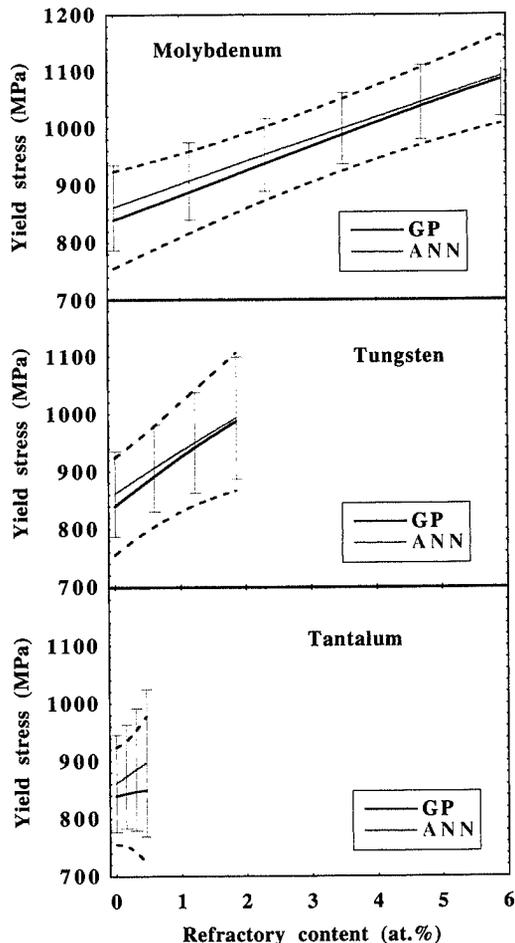


Fig. 7. Predicted individual influences of Mo, W and Ta additions on a Ni-Cr-Co-Ti-Al-C alloy at 21°C.

but not by the GP. However, it should be kept in mind that these predictions are associated to large error bars, *i.e.* a very low level of confidence.

5.7. Extrapolation

The difference between neural networks and Gaussian processes for extrapolation outside the database is of theoretical concern. An ANN will extrapolate by keeping the trend it had inferred just at the end of the database.

This can produce good results if the inferred trend is reasonable, but may also be wrong in the case of a deviation of the true trend from the inferred one beyond the database. A GP will keep this “last” trend for a while, but the covariance function that we use assumes that trends do not persist beyond lengthscales defined by the hyperparameters, $\{r_i\}$, so in the case of extrapolation to large distances the predictions tend to a default value determined by the prior (zero in the present case—in normalised values). In both cases, predictions become unreliable, and the models indicate low levels of confidence through large error bars. If an extrapolation is to be made, the results must be understood as a simple indication, and experimental verification will always be necessary. However, ANN are more likely to perform better extrapolations in the case of monotonic trends, and, if GP extrapolations are disatisfactory, the covariance function of the GP can be altered so that the model expects trends of the desired form.

5.8. Labour and Computing Time

Because the two techniques seem to give similar results, the choice between them could often be a question of how much labour and/or computing time is needed to optimize the model and to make predictions.

The first part of the problem is optimizing—or training—the model. This step is always very laborious in the case of neural networks, since it implies training and testing many different networks with different initial conditions and complexities, choosing the best ones amongst them, and making a committee of them. In these conditions, computing time can often be neglected with respect to the human labour, even with databases of several thousands of points. Typically, optimising a good committee of neural networks would need an “incompressible” human work of several days or weeks.

By contrast, training Gaussian processes is relatively simple, and in our experience it is sufficient to train just one model. However, the optimisation computing time increases approximately as the cube of the database size. Thus, optimisation times can be very fast (minutes or hours) for small databases (a few hundred points) but

can become very long and prohibitive for databases of several thousand points (for example, in another problem, approximately 20 days for a database with 2000 points and 30 inputs on a 300 MHz CPU).

The second part of the problem is making predictions. Once they are trained, neural networks don't use the database to perform predictions, but just a mathematical equation and a set of optimised parameters, or "weights". Thus, predictions can be very fast, the computing time is independent of the database size, and just depends on the number of inputs, the number of parameters in the model and, of course, the number of predictions to be made. In the present problem (16 inputs), the committee of ANN makes about 560 predictions per minute on a 300 MHz CPU.

Gaussian processes use the database to make predictions, with a computing time increasing either as the cube, the square or linearly with the database size, depending on the implementation (if we choose not to ask for error bars, the complexity can be reduced to linear).⁵⁾ In the present case, with 16 inputs, and using only the training set (96 points) as the database to make predictions, the model performs about 1500 predictions per minute, which is faster than the committee of neural networks. But if the whole database is used to make predictions (192 points), the speed drops down to 240 predictions per minute, and down to 10 predictions per minute when using a database of 384 points.

6. Conclusions

The abilities of a committee of artificial neural networks and of a Gaussian process to produce an empirical model of the yield strength of commercial nickel-base superalloys as a function of composition and test temperature have been compared. The comparison criterion was the ability to recognise and predict some well-known metallurgical trends. Both ANN and GP are able to predict correctly, with a similar level of confidence, the influences of test temperature, of a number of individual alloying elements (Al, Ti, Co, W, Mo), of the γ' volume fraction, and of the Ti/Al ratio. The predictions are in good agreement with published measurements and theoretical expectations.

In all cases, ANN and GP give similar results, and display a low level of confidence in the regions where data are missing or widely separated (temperature range 21–538°C, effect of Ta). In these regions of the database, error bars act as a warning signal.

The main advantage of Gaussian processes is to avoid the labour of selecting optimum models, as is necessary in the case of artificial neural networks. Moreover, our method made use of a test set to choose the committee of ANN. The GP method has no need for a test set, so the GP can be trained just once, on the whole database.

However, GP can be computationally time-consuming for large databases, since the computing time grows as

the cube of the database size for optimisation and as the square for prediction (in our case). On the other hand, the calculation time for ANN to make predictions is independent of the training database size. Consequently, because they are simple to use, GP can be preferred for small databases (typically less than one thousand input–output pairs), and ANN preferred for large databases. Nevertheless, due to the rapid increase in computer power and speed, Gaussian processes may become the method of choice in the future.

Acknowledgements

We are grateful to the Engineering and Physical Sciences Research Council for funding this work, and to Alstom Energy Ltd., British Steel plc., INCO Alloys Ltd., Mitsui Babcock Energy Ltd., Rolls-Royce Allen Steam Turines and the University of Wales, Swansea for partnership. The authors thank Doctor Joy Jones for the provision of data. It is a pleasure to acknowledge all the help and support provided by Professor Colin Humphreys in the Technology Foresight Programme.

REFERENCES

- 1) H. Fujii, D. J. C. MacKay and H. K. D. H. Bhadeshia: *ISIJ Int.*, **36** (1996), No. 11, 1373.
- 2) K. Ichikawa, H. K. D. H. Bhadeshia and D. J. C. MacKay: *Sci. Technol. Joining*, **1** (1996), No. 1, 43.
- 3) T. Cool, H. K. D. H. Bhadeshia and D. J. C. MacKay: *Mater. Sci. Eng.*, **A223** (1997), 186.
- 4) D. MacKay: *Neural Computation*, **4** (1992), 448.
- 5) M. N. Gibbs: PhD Thesis. University of Cambridge, UK, (1998).
- 6) C. K. I. Williams and C. E. Rasmussen: *Advances in Neural Information Processing Systems 8*, MIT Press, Cambridge, MA, (1996).
- 7) C. A. L. Bailer-Jones, D. J. C. MacKay, T. J. Sabin and P. J. Withers: *Aust. J. Intelligent Information Processing Systems*, **5** (1998), No. 1, 10.
- 8) C. A. L. Bailer-Jones, H. K. D. H. Bhadeshia and D. J. C. MacKay: *Mater. Sci. Technol.*, **15** (1999), 287.
- 9) High-temperature High-strength Nickel Base Alloys, the Nickel Development Institute, Birmingham, UK, (1987).
- 10) J. Jones: PhD thesis, University of Cambridge, UK, (1997).
- 11) P. Beardmore, R. G. Davies and T. L. Johnston: *Trans. Metall. Soc. AIME*, **245** (1969), 1537.
- 12) L. R. Cornwell, J. D. Embury and G. R. Purdy: Proc. the 3rd Bolton Landing Conf. on Ordered Alloys, Claitors Publishing Co., Baton Rouge, (1970), 387.
- 13) D. Raynor and J. M. Silcock: *J. Met. Sci.*, **4** (1970), 121.
- 14) L. M. Brown and R. K. Ham: *Strengthening Methods in Crystals*, Elsevier Publishing Co. Ltd., London, (1971), 9.
- 15) R. F. Miller and G. S. Ansell: *Metall. Trans.*, **8A** (1977), 1979.
- 16) A. P. Miodownik and N. Saunders: Report to Rolls-Royce plc. UK Thermotech Ltd., (1993), 24133.
- 17) G. E. Maurer, L. A. Jackman and J. A. Domingue: *Superalloys 1980*, ASM, Metals Park, OH, (1980), 43.
- 18) A. K. Jena and M. C. Chaturvedi, *J. Mater. Sci.*, **19** (1984), 3121.
- 19) Y. Mishima, S. Ochiai, N. Hamao, M. Yodogawa and T. Suzuki: *Trans. Jpn. Inst. Met.*, **27** (1986), No. 9, 656.
- 20) J. P. Collier, P. W. Keefe and J. K. Tien: *Metall. Trans. A*, **17A** (1986), 651.
- 21) Z. Meng, G. Sun and M. Li: *Superalloys 84*, TMS, Warrendale, PA, (1984), 563.
- 22) M. V. Natal and L. J. Ebert: *Metall. Trans. A*, **16A** (1985), 1863.