Materials and Manufacturing Processes 24 (2009) 16-21.

Optimization of Neural Network for Charpy Toughness of Steel Welds

Junhak Pak, Jaehoon Jang, H. K. D. H. Bhadeshia and L. Karlsson*

Graduate Institute of Ferrous Technology (GIFT)
Pohang University of Science and Technology
Pohang, Republic of Korea

**ESAB AB, Central Research Laboratories*
40277 Goteborg, Sweden

Abstract

By their very nature, empirical models must be treated with care in order to avoid predictions which are not physically possible. One example is the calculation of the Charpy impact toughness of steel welds as a function of composition and processing, where the impact energy should not be negative. However, there is nothing to prevent a user from implementing inputs which lead to nonsensical results. We examine here whether a scheme used in kinetic theory can be generalised to create neural networks which are bounded. It is found that such procedures lead to bias. In the process of doing this work, some interesting trends have been discovered on the role of process parameters in determining the toughness of steel welds.

Keywords Bias; Charpy energy; Neural networks; Welds.

Introduction

Empirical equations are mathematical functions which suffer from the disadvantage that they can be misused to produce unphysical values of outputs. The problem becomes acute when such models are made available to the general user, who may not appreciate the subtleties of the technique. In structural engineering, the Charpy toughness is an important quality control parameter which is used to gain confidence in the integrity of components. It represents the energy absorbed by a standard sample during fracture. The higher the energy, the safer the component is likely to be in the context of brittle fracture.

The Charpy energy can never be negative since that implies that the sample contributes energy to the test machine. Problems like these can in principle be

circumvented by avoiding the direct modelling of the output, but instead using a logarithm or double logarithm of output during the training of an empirical model [1-3]. It is the purpose of this paper to examine this issue and at the same time to discover trends in the Charpy energy of weld metals as a function of certain solutes and welding parameters. The neural network method itself is not discussed [4-8]; the details are available elsewhere along with numerous examples of applications [9-17]. The way in which the Charpy model was designed using the network has also been explained thoroughly in an accessible thesis [18], so only relevant details are included here for the sake of brevity.

The Data and Network

The variables included in the dataset accumulated for analysis are listed in Table 1 along with the means and standard deviations. The full set consisted of some 5973 experiments collected from industrial collaborators and published data as explained elsewhere [18].

The Bayesian network [5-7] itself always consisted of just three layers, the input, hidden and output layers. An individual such model may include up to 20 hidden nodes, and five different seeds were used in each case to ensure that the starting values of the weights does not lead to a loss of performance. In addition, we do not simply use a best-fit model but a committee of models in order to ensure better generalisation; the membership of the committee is ensured by ranking all the models in terms of their error and then testing various combinations to see whether the generalisation can be improved on unseen data. The specific details are in [18] and the actual models are fully documented and available freely on www.msm.cam.ac.uk/map.

One mechanism by which overfitting was avoided was by dividing the dataset equally and at random into a training and a test set. The latter was not used in creating the model but instead to check how the model generalises with unseen data. The training and test errors should be about the same in order to avoid overfitting. The typical performance of an individual model on a test and training dataset is illustrated in Fig. 1.

Bias in Models

Since a neural network is a regression method, there is a risk of producing unphysical relationships. In modelling the volume fraction of retained austenite in cast irons, Yescas *et al.* [1,2] used a combination of logarithms to confine the fraction between 0 and 1. The double logarithm function used in that work is consistent with Avrami theory [19-21] for the kinetics of solid-state transformations, in which the fraction ξ varies as $\xi = 1 - \exp(-kt^n)$, where k and n

are constants and *t* is the time. It follows according to this physical model that $\ln[-\ln(1-\xi)]$ should vary with $n \ln t$. The volume faction is therefore naturally confined between 0 and 1 so it is justified to use this double logarithmic function to model the volume fraction. Indeed, Yescas *et al.* discovered that it is only when the output is this function, rather than the raw volume fraction, that the variation in the latter can be correctly captured by the neural network [1-2]. In another context, Sourmail *et al.* in their work on the elevated temperature properties of steels modelled the logarithm of creep rupture life rather than the life itself [3]. There is again, a physical basis for this because in a plot of stress versus life, the life extends dramatically as the stress is decreased towards zero. The life should be infinite at zero stress, but it clearly cannot be negative and this is consistent with the use of a logarithm of the life.

The question then arises as to whether it is acceptable to use the logarithm or double logarithm functions purely in order to confine the output to be greater than zero, or between 0 and 1, respectively, even when there is no physical justification for the form of the function.

To assess this scenario, the following equation was to represent the Charpy impact energy *y* as the output in a neural network model:

$$y' = -\ln\left\{-\ln\left(\frac{y-y_{min}}{y_{max}-y_{min}}\right)\right\} \qquad \{1\}$$

where y_{min} and y_{max} are the minimum and maximum set values of the impact toughness. Compared with the original equation of Yescas [1,2], the term on the right was multiplied by -1 in order to retain a direct proportionality between y and y'. The value of y_{min} was set as zero, the least physical value of impact toughness. Because of a lack of physical basis for *y'*, there arises a difficulty in setting the value of y_{max}; although this must be larger than the maximum value in the dataset, it can take on any value beyond the maximum measured value.

Thus two different models were created. In the first model, y_{max} was set to the maximum measured impact toughness of 357 J, and in the second model, it was arbitrarily set to ten times this value at 3570 J. Some calculations using randomly selected inputs from the database (Table 2, Data 1) are illustrated in Fig. 2. Since the purpose was to highlight the influence of y_{max}, the extrapolation was performed to impractically high test temperature regimes. Note that any error bars of prediction by neural networks in this work include $\pm 1\sigma$ and fitting uncertainty as calculated by the Bayesian framework [4-8].

The deliberate exaggeration of extrapolated results in Fig. 2 reveals dramatically different results for the two methods. With the first one (y_{max}= 357 J), the limiting value at high temperatures converges artificially to y_{max}. In contrast, y_{max} = 3570 J, the extrapolation leads to approximately zero toughness at high temperatures. This implies strongly that the selection of y_{max} plays a role in biasing during modelling or during the undoing of double logarithms in the final step to obtain toughness.

The toughness should roughly follow a sigmoidal trend as a function of the test temperature, with low values typical of cleavage at low temperatures, and high values associated with ductile fracture at elevated temperatures. This is roughly the trend shown in Fig. 2a, but the upper limit becomes 357 J, which cannot be exceeded due to the form of equation 1. This also cannot be physically justified. However, if the upper limit is set artbitrarily high, then the experimental data on which the model is trained fall into the low-toughness regime in the output space. Only a tenth of the region defined for Fig. 2b would be occupied by the training data. The model perceives this to indicate that the toughness will always be much smaller than its capability to predict, and as a consequence, the extrapolation shown in Fig. 2b tends towards zero.

The metallurgical problem considered here is such that this kind of bias cannot be justified, unlike the kinetic analysis by Yescas [1,2]. It is not acceptable to set bounds on the toughness. Thus, the use of equation 1 was abandoned. An alternative to use $y' = \ln y$ was also excluded because it practically allows an infinite value of y_{max} which cannot be justified either. In conclusion it was decided to use just the raw toughness values as the output in all the modelling and to be careful in interpreting the outputs.

2. Interpass temperature

The input space was thoroughly explored in order to discover significant phenomena worthy of careful study. Interesting relationships associated with the interpass temperature were found, based on the input conditions listed in Table 2, Data 2 [23]. They approximately correspond to the compositions of the 0.5Mn alloy and the 2Mn alloy reported elsewhere [23-26].

The interpass temperature is that at which a weld is allowed to cool to, before depositing another layer. The biggest physical effect of the interpass temperature during welding is to influence the cooling rate. This can be seen from the Rosenthal equation [27-28]:

$$T - T_0 = \frac{P}{2\pi\kappa r}\exp\left\{-\frac{v(\xi+r)}{2a}\right\} \quad \{2\}$$

where T is local temperature due to a point source of power input P moving at velocity v. T_0 is the far-field plate temperature which is equivalent to the interpass temperature in this case. κ is the thermal conductivity and a is the thermal diffusivity. r is a polar co-ordinate measured from a reference frame attached to the moving heat source, related to the stationary frame by: $r = \sqrt{\xi^2 + y^2 + z^2}$ where ξ is a coordinate measuring the translation of the heat source, y and z simply being coordinate axes orthogonal to ξ. Thus the cooling rate is expected to decrease as the interpass temperature increases. Empirical equations [29] for arc welding have been derived based on the form of equation 2, and were used to evaluate the time ($t_{8/5}$) taken for an ISO2560 weld metal to cool over the temperature range 800-500°C as a function of the interpass temperature. The calculation was performed using the welding current, voltage and speed of 180 A, 34 V and 0.004 m s^{-1} (heat input 1.53 kJ mm^{-1}) respectively. for arc voltage, current and welding speed, respectively. The time $t_{8/5}$ was found to be 11.7, 16.6, 26.3 and 53.1 s for interpass temperatures fo 100, 200, 300 and 400°C respectively. The interpass temperature therefore significantly influences the weld cooling rate, and through that, possibly the microstructure and mechanical properties.

We shall use this information in interpreting the results presented below.

3. Results and Discussion

Using the variables listed in Table 2 (Dataset 2), the influence of interpass temperature was analysed, with a particular focus on manganese and nickel since these are the solutes of greatest interest in the design of low-transformation temperature weld metals [23-26]. Many other trends have been reported elsewhere [18]. Fig. 3 shows the results for the 0.5Mn and the 2Mn compositions (with the other variables as listed in Table 2 (Dataset 2). The contour plots for the two cases show a striking similarity. Each plot has regions which have been labelled “high” and “intermediate” respectively, to indicate the sensitivity of the toughness to the interpass temperature at a constant nickel concentration.

We begin by summarising the observations that need explanation. At low manganese concentrations, an increase in the interpass temperature leads to a deterioration of toughness when the nickel concentration is less than 6

wt% (Fig. 3a). In the high manganese case, this deterioration occurs at concentrations less than 2 wt%, beyond which the alloy becomes more tolerant to interpass temperature and there is even an increase in toughness with interpass temperature at large nickel concentrations (Fig. 3b).

The second observation is that the intermediate sensitivity region for the high Mn concentrations as shown in Fig. 3b is shifted to smaller nickel concentrations when compared with the low manganese concentrations as shown in Fig. 3a.

These results can in principle be understood in terms of the cooling of a weld metal in the context of a continuous cooling transformation (CCT) diagram, as shown in Fig. 4. The microstructure of a weld, and hence its properties, will be sensitive to the interpass temperature if the alloy hardenability is such that the cooling curve intersects the CCT diagram at different temperatures when the cooling rate is changed.

Therefore, with a high manganese and high nickel concentration, all cooling curves intersect the CCT diagram at the same temperature (martensite-start), making the alloy insensitive to the interpass temperature, as is the case in Fig. 3b. A reduction in the nickel concentration at high manganese permits slower cooling curves to intersect the bainite part of the CCT diagram, thus increasing the sensitivity of the microstructure to the interpass temperature. When both manganese and nickel are in small concentrations, the alloy becomes very sensitive to the cooling rate.

This analysis explains why the intermediate region (Fig. 3) is shifted to lower nickel concentrations when the manganese concentration is increased, since both nickel and manganese are austenite stabilisers. The predicted deterioration in toughness at low nickel concentrations when the interpass temperature is increased is because the transformation-start temperature increases as the interpass temperature is rasied. Coarser microstructures are associated with higher transformation temperatures re are coarser microstructures generated [15, 30, 31].

4. Conclusions

The attempt to bound predictions of neural networks between upper and lower limits in order to avoid unphysical outputs using logarithmic representations of the output proved unworkable. The method introduced

bias in the models produced. A Charpy impact energy model for steel weld metals was therefore produced using the raw energy rather than a functional representation of that energy. This revealed rather interesting results with respect to the sensitivity of the impact toughness to the interpass temperature, nickel and manganese concentrations. The greatest sensitivity to the interpass temperature is associated with domains in which the microstructure is dependent on the cooling rate. This would not be the case, for example, where the microstructure is predominantly martensitic, because the martensite-start temperature does not for ordinary steels depend on the cooling rate.

Acknowledgements

The authors are grateful Professor Hae-Geon Lee for the provision of laboratory facilities at GIFT.

References

1. M. Yescas Gonzalez, H. K. D. H. Bhadeshia and D. J. C. MacKay, Retained austenite in austempered ductile cast irons, Materials Science and Engineering A, Vol. 311, 2001, 162-173.
2. M. Yescas Gonzalez and H. K. D. H. Bhadeshia, Model for the maximum fraction of retained austenite in austempered ductile cast iron, Materials Science and Engineering A, Vol. 333, 2002, 60-66.
3. T. Sourmail, H. K. D. H. Bhadeshia and D. J. C. MacKay, Neural network model of the creep strength of austenitic stainless steels, Materials Science and Technology, Vol. 18, 2002, 655-663.
4. D. J. C. MacKay, Bayesian interpolation, Neural Computation, Vol. 4, 1992, 41 5-447.
5. D. J. C. MacKay, A practical Bayesian framework for backpropagation network s, Neural Computation, Vol. 4, 1992, 448-472.
6. D. J. C. MacKay, Bayesian non-linear modelling for the energy prediction comp e t i t i o n , American Society of Heating and Refrigeration Engineers (ASHRAE) Transacti ons Vol. 100, 1994, 1053-1062.
7. D. J. C. MacKay, Information theory, inference and learning algorithms, Cambri dge University Press, U. K., 2003.
8. H. K. D. H. Bhadeshia, Neural networks in materials science, ISIJ International, Vol. 39, 1999, 966-979.
9. M. Mukherjee, S. B. Singh and O. N. Mohanty, Neural network analysis of strain induced transformation behaviour of retained austenite in TRIP-aided steels, Materials Science and Engineering A, Vol. 434, 2006, 237-245.
10. M. Mukherjee, S. B. Singh and O. N. Mohanty, Strain induced transformation of retained austenite in TRIP aided steels: a neural network model, Materials Sci ence and Technology, Vol. 23, 2007, 338-346.
11. S. Datta and M. K. Banerjee, Optimizing parameters of supervised learning tec hniques (ANN) for precise mapping of the input--output relationship in TMCP steels, Scandanavian Journal of Metallurgy, Vol. 33, 2004, 310-315.
12. R. Kemp and G. A. Cottrell and H. K. D. H. Bhadeshia and G. R. Odette and T. Yamamoto and H. Kishimoto, Neural network analysis of Charpy transition te mperature of irradiated low-activation martensitic steels, Journal of Nuclear M aterials, Vol. 367-370, 2007, 603-609.
13. C. Capdevilla and C. Garcia-Mateo and F. G. Caballero and C. Garcia de Andres , Neural network analysis of the influence of processing on strength and ductil ity of automotive low carbon sheet steels, Journal of Materials Science, Vol. 3 8, 2006, 192-201.
14. H. K. D. H. Bhadeshia, D. J. C. MacKay and L.-E. Svensson, The impact toughness of steel welds, Materials Science and Technology, Vol. 11, 1995, 1046-1051.

15. S. B. Singh and H. K. D. H. Bhadeshia, Estimation of bainite plate thickness, M aterials Science and Engineering A, Vol. 245, 1998, 72-79.
16. T. Cool and H. K. D. H. Bhadeshia, The yield and ultimate tensile strength of s teel welds, Materials Science and Engineering, Vol. 223, 1997, 186-200.
17. S. Chatterjee and H. K. D. H. Bhadeshia, δ-TRIP steel, Materials Science and T echnology, Vol. 23, 2007, 819-827.
18. J. Hak Pak, Modelling of impact toughness of weld metals, Master of Engineering Thesis, Pohang University of Science and Technology, Pohang, Korea, 2007. www.cml.postech.ac.kr
19. M. Avrami, Kinetics of phase change, Journal of Chemical Physics, Vol. 7, 1939, 1103-1112.
20. M. Avrami, Kinetics of phase change, Journal of Chemical Physics, Vol. 8, 1940, 212-224.
21. M. Avrami, Kinetics of phase change, Journal of Chemical Physics, Vol. 9, 1941, 177-185.
22. J. Fridberg, L.-E. Torndahl and M. Hillert, Diffusion in iron, International Metals Reviews, Vol. 4, 1981, 185-212.
23. M. Murugananth, H. K. D. H. Bhadeshia, E. Keehan, H. O. Andren and L. Karlsson, Mathematical Modelling of Weld Phenomena VI, Strong and tough steel welds, eds H. Cerjak and H. K. D. H. Bhadeshia, Institute of Materials, London, 2002, 205-230.
24. E. Keehan, L. Karlsson, H.-O. Andren and H. K. D. H. Bhadeshia, Influence of C, Mn and Ni on strong steel weld metals: increased impact toughness, Science and Technology of Welding and Joining, Vol. 11, 2006, 9-18.
25. E. Keehan, L. Karlsson, H.-O. Andren and H. K. D. H. Bhadeshia, Influence of C, Mn and Ni on strong steel weld metals: increased strength, Science and Technology of Welding and Joining, Vol. 11, 2006, 18-24.
26. E. Keehan, L. Karlsson, H.-O. Andren and H. K. D. H. Bhadeshia, New developments with C-Mn-Ni high strength steel weld metals: microstructure, Welding Journal, Vol. 85, 2006, 200-210.
27. D. Rosenthal, Mathematical theory of heat distribution during welding and cutting, Welding Journal, Vol. 20, 1941, 220s-234s.
28. D. Rosenthal, Theory of moving sources of heat and its application to metal treatments, Transactions of the A.S.M.E., B11, Nov. 1946, 849-866.
29. L.-E. Svensson, B. Gretoft and H. K. D. H. Bhadeshia, A Bayesian neural network analysis, Scandanavian Journal of Metallurgy, Vol. 15, 1986, 97-103..
30. M. Lord, Design and modelling of strong steel weld deposits, Ph.D. Thesis, University of Cambridge, 1999.
31. R. C. Reed and H. K. D. H. Bhadeshia, Reconstructive austenite-ferrite transformation in low alloy steel, Materials Science and Technology, Vol. 8, 1992, 421-435.

Materials and Manufacturing Processes 24 (2009) 16-21.

Table 1: The dataset used in creating the neural network models.

Variable	Range	Mean	Standard deviation
C/wt%	0.008-0.19	0.07	0.02
Si/wt%	0-1.63	0.35	0.14
Mn/wt%	0-2.31	1.20	0.40
S/wt%	0.002-0.14	0.01	0.01
P/wt%	0-0.25	0.01	0.01
Ni/wt%	0-12.40	0.88	1.83
Cr/wt%	0-19.50	0.35	1.19
Mo/wt%	0-2.43	0.20	0.31
V/wt%	0-0.53	0.01	0.03
Cu/wt%	0-2.18	0.08	0.21
Co/wt%	0-0.092	0.003	0.01
W/wt%	0-3.86	0.004	0.11
O/ppmw	25-1700	429	161
Ti/ppmw	0-770	71	115
N/ppmw	0-1000	95	67
B/ppmw	0-200	8	27
Nb/ppmw	0-1770	32	109
HI/kJmm^{-1}	0.21-16.36	1.49	0.83
IT/℃	20-350	182	39
PWHTT/℃	20-940	198	287
PWHTt/h	0-100	1.3	4.2
D_{Fe}	0-3.68 × 10^{12}	6.22 × 10^{10}	4.47 × 10^{11}
TT/℃	-196-136	-34.5342	36
Charpy toughness/J	0.1-356	85	50

ppmw: Part per million by weight
HI: Heat input IT: Interpass temperature
PWHTT: Post-weld heat treatment temperature
PWHTt: Post-weld heat treatment time
D_{Fe}: A variable for iron diffusion during post-weld heat treatment
TT: Test temperature

Table 2: Input sets used for the calculations presented in Fig. 1 and in the interpass temperature calculations. D_{Fe} is given by the product of the post-weld heat treatment time and exp{-Q/RT}, where Q is the activation energy for the self-diffusion of iron at 260,000 J mol^{-1} [22]; it is a measure of the potency of the heat treatment.

Input	Data 1	Data 2
C	0.034	0.025
Si	0.27	0.37
Mn	2.14	0.65
S	0.008	0.006
P	0.01	0.013
Ni	7.3	6.6
Cr	0.5	0.21
Mo	0.62	0.4
V	0.011	0.011
Cu	0.03	0.03
Co	0.009	0.009
W	0.005	0.005
O	0.033	0.038
Ti	0.008	0.008
N	0.012	0.0180
B	0.0001	0.0001
Nb	0.001	0.001
Heat input / kJ mm^{-1}	1	1
Interpass temperature / °C	250	250
Post-weld heat treatment temperature / °C	20	20
Post-weld heat treatment time / °C	0	0
D_{Fe} / s	0	0
Test temperature / °C	-60	-60

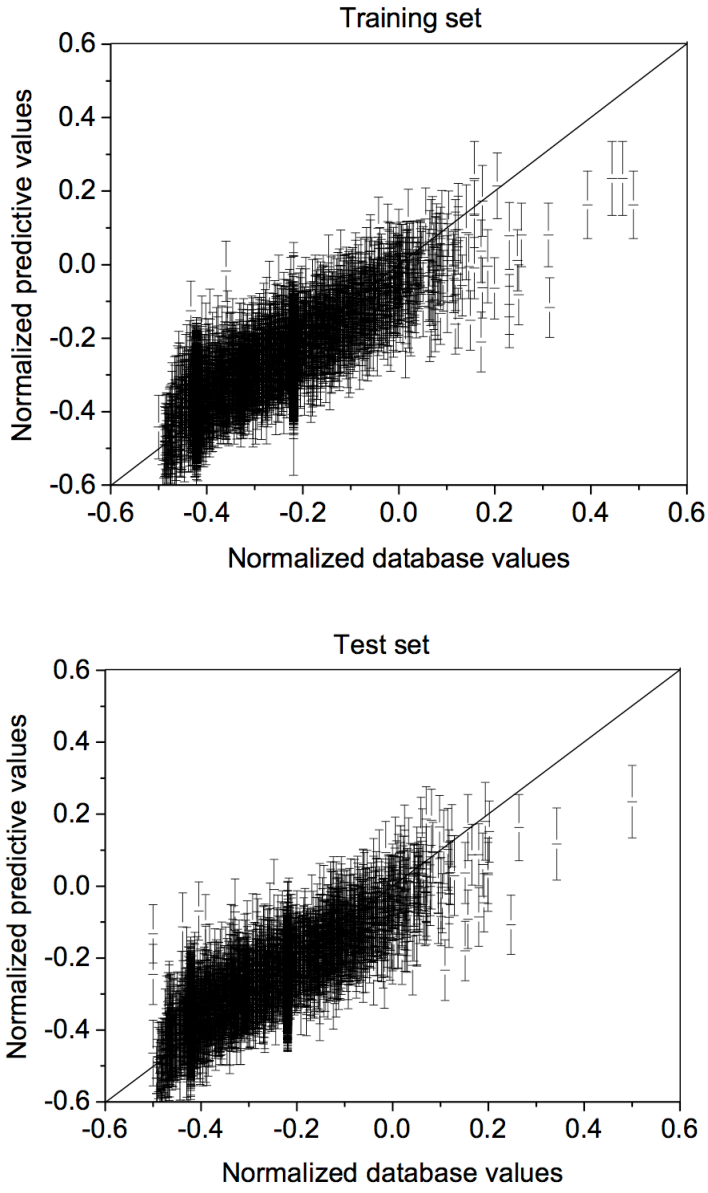
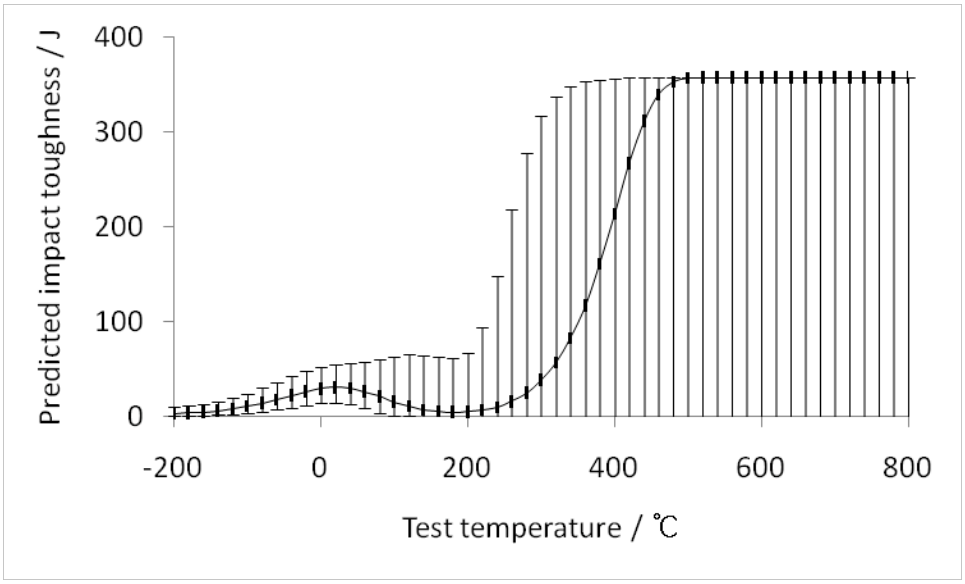
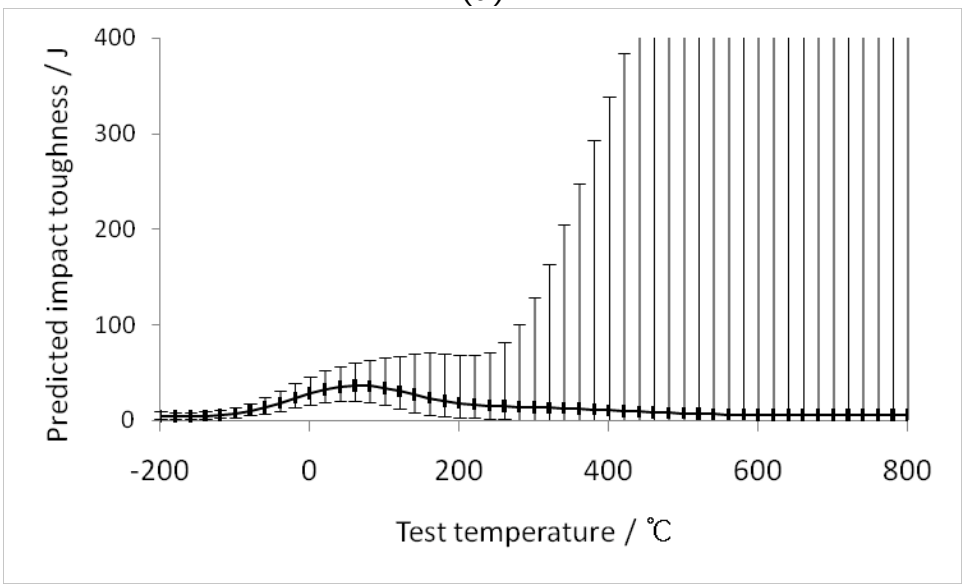


Fig. 1: The typical performance of the trained model on training and test (unseen) data.



(a)



(b)

Fig. 2: Extrapolation results illustrating the influence of the choice of y_{max} on the behaviour of the neural network . (a) y_{min}=0 and y_{max}=357. (b) y_{min}=0 and y_{max}=3570, with vertical scale limited 400J because of the upper bounds can be as large as 3570.

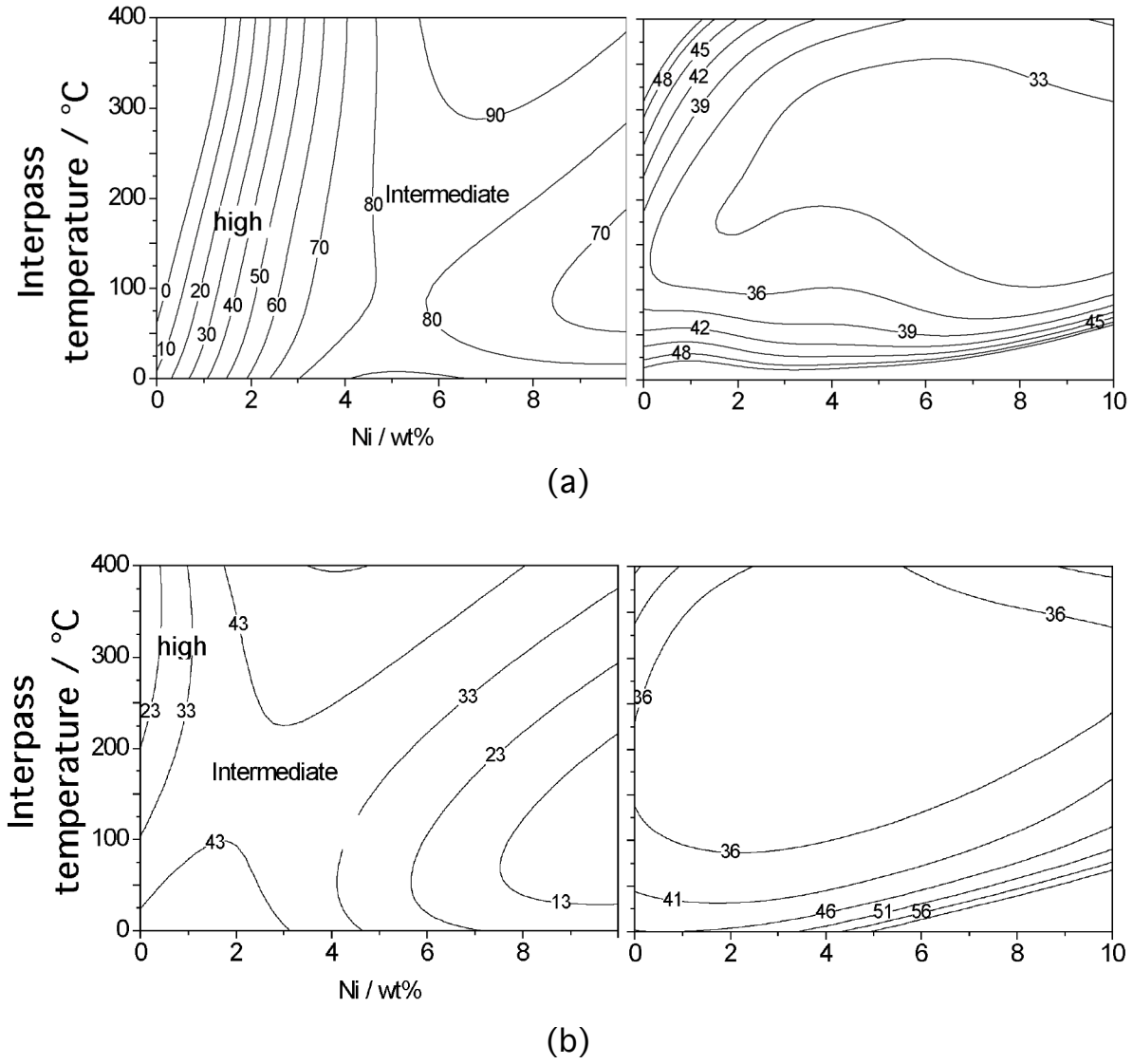


Fig. 3: Influence of the interpass temperature and nickel on the impact toughness. The contours represent the impact energy in Joules, and the diagram on the right in each case gives the corresponding contours for the uncertainty ($\pm 1\sigma$) in that energy. (a) Low manganese (0.5Mn), (b) high manganese (2Mn).

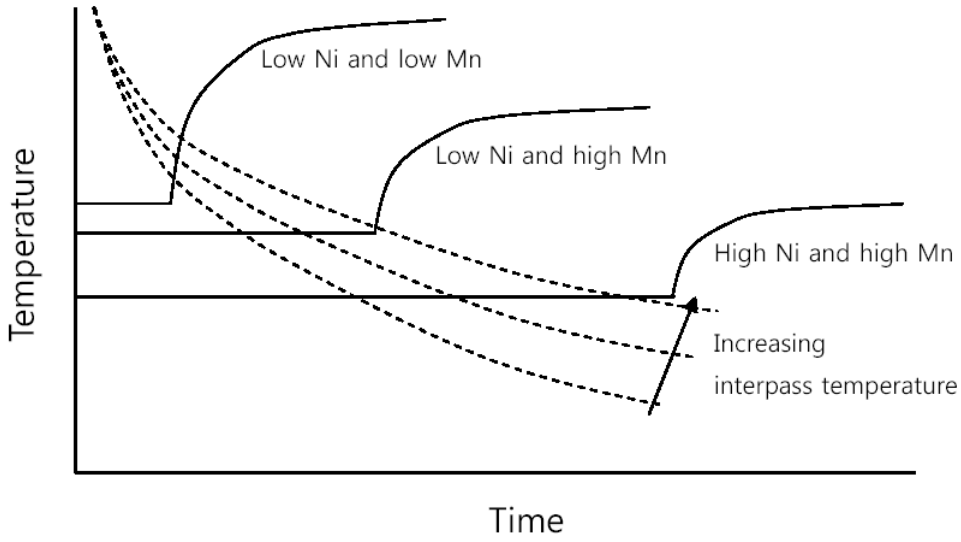


Fig. 4: Schematic continuous cooling transformation diagram illustrating the effect of solutes and interpass temperature on transformation temperature.