Proposal of an empirical formula for the austenitising temperature

Carlos Capdevila, Carlos García-Mateo, Francisca G. Caballero and Carlos García de Andrés

Department of Physical Metallurgy, Centro Nacional de Investigaciones metalurgicas

(CENIM), Consejo Superior de Investigaciones Cientificas (CSIC), Avda. Gregorio del

Amo, 8, 28040 Madrid, Spain.

Corresponding autor: Carlos García de Andrés. Tel.: +34 915538900 (ext. 280)

Fax.: +34 915347425. E-mail: cgda@cenim.csic.es

Abstract

A formula is proposed to calculate the chemical austenitising temperature derived by means of a neural network model. A wide range of compositions were used, allowing to predict changes of tendency in the slopes of the austenitisation temperature, consistent with the existence of a eutectoid composition.

Keywords

Austenite, neural network, thermodynamics, no-linear regression

1. Introduction

One of the most important factors to take into account in designing a steel is the temperature at which austenite is the only stable phase present. This matter has been objective of research for long, the most well known done by Andrews [1], who proposed an empirical formula that works extremely good in a limited range of compositions. The aim of this paper is to extend the chemical compositions where the formula is applicable, as well as to find an expression that could account for the change of tendency when an eutectoid point is present. For

this purpose new modelling techniques like neural network and the latest of assessed thermodynamic data for steels, MTDATA, are used [2].

In normal regression methods, the analysis begins with a prior choice of a relationship (usually linear) between the output and the input variables. A neural network is capable of performing a greater variety of no-linear relationships of considerable complexity. Data are presented to the network in the form of input and output parameters, and the optimum non-linear relationship is found by minimising a penalized likelihood. In effect the network tries out many kinds of relationships in its search for an optimum fit. As in regression analysis, the results consist of a function that relates the inputs to the outputs by a series of coefficients called "weights". In spite of its apparent sophistication, the method is as blind as regression analysis, and neural networks can be susceptible to overfitting. This paper deals with the analysis of the influence of alloying elements such as Mn, Cr, Ni, Si and Mo, separately and together.

First of all it is necessary to define the chemical austenitising temperature T_γ. Two possible definitions for the T_γ could be drawn: i) the minimum temperature at which only austenite exists (no carbides at all); ii) the temperature when ferrite disappear from the microstructure. Fig. 1 shows an example of an isopleth corresponding to the Fe-C-Ni alloy with 0.85% of C. This simple example illustrates how different could be T_γ values and its tendencies. If the first T_γ definition is assumed, open squares, it increases as Ni concentration does, while the tendency is opposite, decreasing T_γ with increasing Ni contents, if the second definition is used, solid squares in Fig. 1.

The experimental data used to create the neural network model was collected from published data [3-6], where precipitates were detected although not identified, thereby the first T_γ definition will be use.

2

2. Experimental procedure: Construction of the model

2.1. The technique

Neural networks are parameterized non-linear models used for empirical regression and classification modelling. Their flexibility enables them to discover more complex relationships between the data than traditional linear statistical models.

A neural network is 'trained' on a set of examples of input and output data. The outcome of the training is a set of coefficients (referred to as weights) and a specification of the functions which in combination with the weights relates the input to the output. The training process involves a search for the optimum non-linear relationship between the input and output data and is computer intensive. Once the network is trained, estimation of the outputs for any given inputs is very rapid.

One of the difficulties with blind data modelling is that of 'overfitting', in which spurious details and noise in the training data are overrated by the model. This gives rise to solutions that generalize poorly. MacKay [7-8] has developed a particularly useful treatment of neural networks in a Bayesian framework, which allows for the calculation of error bars representing the uncertainty in the fitting parameters. The method recognises that there are many functions which can be fitted or extrapolated into uncertain regions of the input space, without unduly compromising the fit in adjacent regions which are rich in accurate data. Instead of calculating a unique set of weights, a probability distribution of sets of weights is used to define the fitting uncertainty. The error bars therefore become large when data are sparse or locally noisy.

The Bayesian framework for neural networks has two further advantages. First, the significance of the input variables is automatically quantified. Consequently, the significance

perceived by the model of each input variable can be compared against metallurgical theory. Second, the network's predictions are accompanied by error bars which depend on the specific position in input space. These quantify the model's certainty about its predictions.

2.2. The experimental database

A complete description of the chemical composition and the transformation temperature is required to ideally model the T_γ temperature in steels. A search of the literature [3-6] allows us to collect 700 individual cases where detailed chemical composition and decomposition temperature were reported. Table 1 shows the list of 6 input variables used for the T_γ temperature analysis.

Table 1. Variables that influence T_γ temperature in steels. SD is standard deviation. Concentrations are in wt.%.

	Minimum	Maximum	Average	SD
C	0.00	2.09	0.41	0.31
Mn	0.00	20.00	0.78	1.59
Si	0.00	3.40	0.21	0.41
Ni	0.00	40.00	1.32	3.49
Cr	0.00	18.39	1.71	3.80
Mo	0.00	5.09	0.28	0.76
T_γ,(°C)	440.00	1290.00	798.20	145.29

2.3. The neural network analysis

The aim is to be able to estimate T_γ temperature as a function of the input variables listed in Table 1. The network consists of three layers: the first layer consisting in six input nodes; the second one, a number of hidden nodes and finally an output node representing T_γ temperature. The network was trained using 350 examples randomly chosen from a total of 700 available, the remaining 350 examples being used as 'new' experiments to test the trained network.

Chemical composition of each alloy element (x_i) defines the inputs nodes, and the T_γ temperature, y, the output node. The function for a network of i hidden units is given by,

$$y = \sum_i w_i^{(II)} h_i + \theta^{(II)}$$

(1)

where

$$h_i = tanh\left(\sum_j w_{ij}^{(I)} x_i + \theta_i^{(I)} \right)$$

(2)

where $w_i^{(II)}$ and $w_{ij}^{(I)}$ are random weights, and $\theta^{(II)}$ and $\theta_i^{(I)}$ are constants.

Notice that the complexity of the function is related to the number of hidden units but one hidden-unit model may not be sufficiently flexible. The availability of a sufficiently complex and flexible function means that the analysis is not as restricted as in linear regression where the form of the equation has to be specified before the analysis. Fig. 2 (a) shows that as expected, the inferred noise level of data (σ_v) [7] decreases monotonically as the number of hidden units increases. However, a high degree of complexity owing to a large number of hidden units may not be justified, and in an extreme case, the model may meaninglessly attempt to fit the noise in the experimental data. MacKay [8] has made a detailed study of this problem and has defined a quantity (the 'evidence'), which comments on the probability of a model [8]. In circumstances where two models give similar results over the known data set, the more probable model would be predicted to be the simplest one; this simple model would have a higher value of evidence. The evidence framework was used to control the regularization constants and σ_v [8].

To find out the optimum number of hidden units of the model, the following procedure was used. The experimental data were partitioned equally and randomly into a test dataset and a training dataset. Only the latter was used to train the model, whose ability to generalize was

examined by checking its performance on the unseen test data. The test error (T_{en}) [7] is a reflection of the ability of the model to predict the output value in the test data:

$$T_{en} = 0.5 \sum_n (y_n - t_n)^2$$

(3)

where y_n is the set of predictions made by the model and t_n is the set of target (experimental) values. In Fig. 2(b), it can be seen that the calculated test error for this T_γ temperature model goes through a minimum at eleven hidden units, T_{en} =0.285. Therefore, the optimum model is that which considers only eleven hidden units. The level of agreement for the training and the test data is shown in Fig. 3(a) and (b); good predictions occur in both instances.

However, it is possible that a *committee* of models can make a more reliable prediction than an individual model. A *committee* of seven best models gives a value of $T_{en} = 0.273$. Therefore, the neural network model used to calculate the T_γ temperature in this paper is a committee of seven best models. Comparison between the calculated and measured values of T_γ temperature for the training and test data is shown in Fig.4 for the best committee (consisting of seven best models).

2.4. Significance of input variables

Fig.5 illustrates the significance (σ_w) [7] of each of the input variables, as perceived by the neural network, in influencing the T_γ temperature. The carbon content together with chromium content clearly has a large intrinsic effect, which is consistent with experimental evidences reported in the literature for decades. The content in manganese in the steel also have a significant effect on T_γ temperature, as it was well established in literature [3], meanwhile the influence of silicon, nickel and molybdenum is moderate.

3. Results and discussion

3.1. Use of the model

A neural network model was used to analyse, in a first step, the evolution of the T_γ temperature in binary systems. In this sense, Fig. 6(a) shows the evolution of T_γ temperature with carbon content for Fe-C steels. Solid line represents the model predictions meanwhile dashed lines represent the error bounds.

It is clear from Fig. 6(b) that no phase transition between $\gamma \rightarrow \alpha$ it is possible for chromium content higher than 14 wt.% since error bounds increase considerably. Likewise, a similar effect with increasing molybdenum content is detected, i.e. no $\gamma \rightarrow \alpha$ transformation for molybdenum content higher than 3.8 wt.% is produced which is in accordance with phase diagram assessments, Fig. 6(c).

Likewise, it is worth mentioning the change in tendency in T_γ temperature in Fig. 6(b) with chromium content. In this sense, a decrease in T_γ temperature is expected as chromium content increases up to 8 wt.%, meanwhile a rise in T_γ temperature values is obtained for further amounts of chromium, which is consistent with Fe-Cr phase binary diagram predictions.

Fig. 6(e) shows evolution of T_γ temperature with Mn content, meanwhile Fig. 6(f) shows the evolution of T_γ increasing the content in nickel. It is shown that T_γ temperature monotonically decreases as Mn and/or Ni content increase, which is in agreement with respective binary phase diagrams.

4. Validation of the neural network model

The created neural network model has been validated using MTDATA [2] together with the NPL database, the latest of assessed thermodynamic data for steels. A series of equilibrium calculations on Fe-C-X ternary alloys, where X stands for Cr, Mo, Ni, Mn and Si, were selected

for the validation. Table 2 shows the range of chemical compositions used for the calculations. The evolution of phase fractions for steels at a variety of temperatures was calculated allowing for the potential existence of cementite, ferrite and austenite in all cases. In some systems others phases were included, such as CMo_2 and M_6C in Mo containing alloys and M_7C_3 and $M_{23}C_6$ in Cr alloys, where M stands for substitutional elements.

Table 2. Theoretical range of chemical compositions studied (wt.%). Fe is set to balance.

C	Mn	Ni	Si	Mo	Cr
0.08-0.85	0-2	0-9	0-2.1	0-5	0-18

Results thus obtained were plotted against the predictions obtained by the model, Fig. 7, showing the good level of agreement achieved, especially in the systems containing Si, Mn and Ni. Due to the complex phase diagrams of some Mo and Cr steels, is not surprising that they exhibit a higher discrepancy, but in all cases it lays within the error bars limits, therefore validating the created model.

5. Determination of a new formula for the T_γ temperature

Determination of an equation that summarises the knowledge provided by the neural network model created will allow making a much easier use of it.

In order to obtain a formula for the T_γ temperature is necessary to disclose the effect of different alloying elements. The same compositions shown in Table 2, but increasing the C up to 1.5 wt.%, were used. A paraboloid type equation, rather than lineal, is found to be a better approximation to take into account the different tendency of the T_γ temperature once the eutectoid point is pass, equation 4

$$T_\gamma \ (^\circ C) = To - t_1 \ [C]^{1/2} + t_2 \ [X] + t_3 \ [C] \ [X] + t_4 \ [C]^2 + t_5 \ [X]^2 \qquad (4)$$

where To stands for the T_γ temperature for pure-iron, [C] and [X] stands for the mass percentage of carbon and other alloying elements, and t_i (i=1…5) represent fitting constants.

The equation keeps the square root dependence for carbon proposed by Andrews [1], and it also includes terms that take into account the interaction of C-alloying element, C X, and a square dependence of both.

Due to the fact that Fe and C have the largest effect on the T_γ, before everything else, their fitting constants were determined. After a first set of calculations a mean value of 922 ^{o}C for the independent term To, and -233 and 67.5 for the carbon coefficients t_1 and t_4, was obtained. Fixing these values, a second set was perform to reveal the values for the remaining constants, see Table 3.

Table 3. Fitting parameters obtained for equation 4. To = 922, t_1 = -233 and t_4 = 67.5 were fixed values.

X	t_2	t_3	t_5
Cr	-5.7	30.2	1.3
Ni	-26.3	34.2	0.5
Si	64.7	-65.7	4.1
Mo	41.9	-4.2	10.9
Mn	-14.3	32.7	-8.4

Equation 4 was tested against the original data used by Andrews [11-12], which range of applicability is up to 0.6 wt.% for C, 2 wt.% for Mn, 5 wt.% for Ni, 1.78 wt.% for Si, 1 wt.% for Mo and 4.5 wt.% for Cr. The extent of agreement is reasonable good, (R^2= 0.84) given the range of compositions and temperatures involved, as illustrates Fig. 8 (a). Table 4 was used to test the reliability of the model, and to compare with predictions obtained by the classical Andrews equation.

Table 4. Range of chemical compositions in (wt.%).

Fe-C-Mn-X	C	Mn	X
Cr	0.2	1.5	1-10
Ni	0.6	1.5	1-9
Mo	0.2	1.5	0.5-3.5
Si	0.8	1.5	0-2

As it is clear from Fig. 8(b) and (c) a change of slope in the T_γ is predicted, which is consistent with the existence of an eutectoid concentration. Such prediction is impossible using a linear type equation like Andrew's.

Finally it is worth mentioning that equation 4 allows us to extend the T_γ calculation to heavily alloyed steels, where Andrews equation is unable to perform reliable predictions since those steels are out of its range of applicability, Fig. 8(b) to (e).

6. Conclusions

Modelling of the austenitising temperature, T_γ, is carried out in this paper. An artificial neural network method based on a Bayesian framework has been used to rationalize the published data on T_γ in steels. The analysis is empirical but after appropriate training, it is found to reliably reproduce known metallurgical experience. The method is useful because the optimised network summarizes knowledge in a quantitative manner and can be retrained as new data became available. This model is different to those empirical and semi-empirical models created by fitting equations to experimental data.

A new equation to predict T_γ temperature has been proposed. The range of applicability of the proposed equation widen that proposed by Andrews [1] in close to 4 wt.% for Ni and Mo, about 14wt.% for Cr and 1 wt.% for C. This wide composition range, allow predicting change of tendency in the slope of T_γ consistent with the existence of eutectoid compositions.

Acknowledgements

References

[1] K. W. Andrews, JISI 203 (1965) 721.

[2] MTDATA National Physical Laboratory, Teddington, U.K.

[3] Atlas of Isothermal Transformation Diagrams, United States Steels (USS) Corporation, Pittsburgh, 1959.

[4] M. Economopoulos, N. Lambert, L. Habraken, Diagrammes de Transformation des Aciers Fabriques dans le Benelux, Centre National de Recherches Metallurgiques (CRM), Bruxelles, 1967.

[5] F. Maratray and R. Usseglio-Nanot, Atlas of Transformation Characteristics of Chromium and Chromium-Molybdenum White Irons, Climax Molybdenum S.A., Paris, France, 1996.

[6] M. Atkins, Atlas of Transformation Diagrams for Engineering Steels, British Steel Corporation (BSC), Sheffield, England, 1985.

[7] D. J. C. MacKay, Neural Comput. 4 (1992) 415-422.

[8] D. J. C. MacKay, Neural Comput. 4 (1992) 448-460.

[9] D. J. C. MacKay: Neural Comput. 4 (1992) 698-705.

[10] D. J. C. MacKay: Darwin college J.3 (1993) 81-93.

[11] G. Delbart and A. Coutant, Courbes de transformation des de fabrication Francaise, I-IV, 1953-60, St Germain-en-Laye, F IRSID

[12] F. Wever and A. Rose, Atlas für Wärmbehandlung der Stähle, 1954, Dusseldorf, Verlag Stahleisen, MBH.

Figure captions
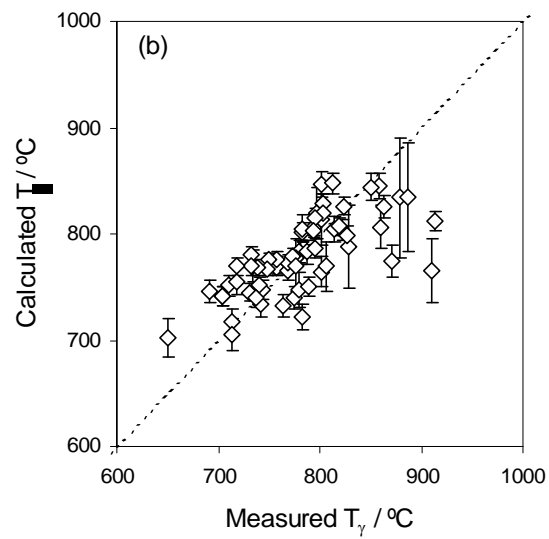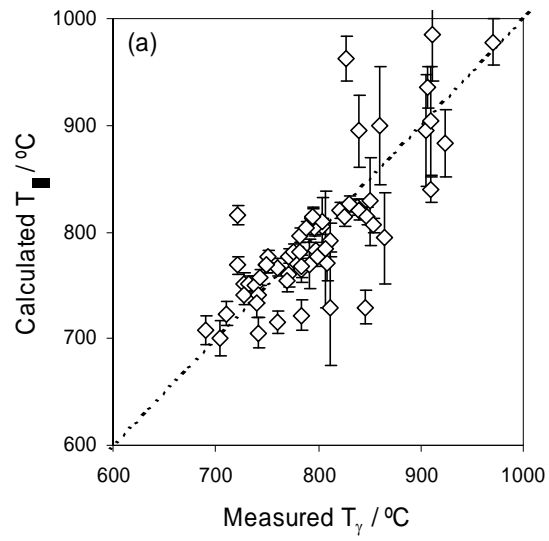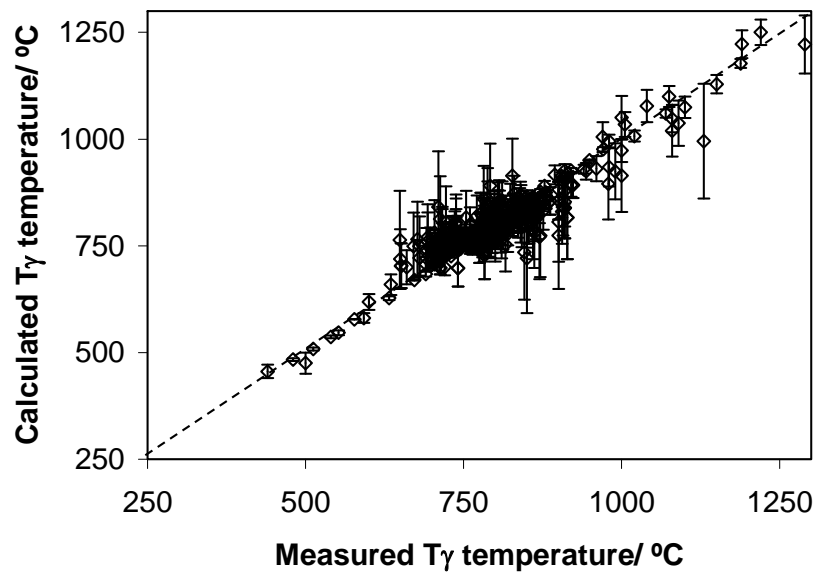
Figure 1. Carbon isopleth of the Fe-C-Ni system .C is 0.85 wt.% and Ni varies from 0 to 9 wt.%, Fe is set to balance. FCC_A1 and BCC_A2 stand for austenite and ferrite respectively.

Figure 2. Variation of (a) inferred noise level (σ_v), and (b) test error (T_{en}) as a function of the number of hidden units.
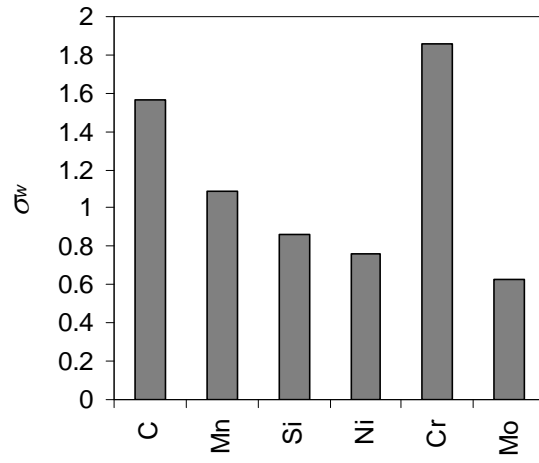
Figure 3. Comparison between calculated and measured T_γ temperature values: (a) training data and (b) test data.
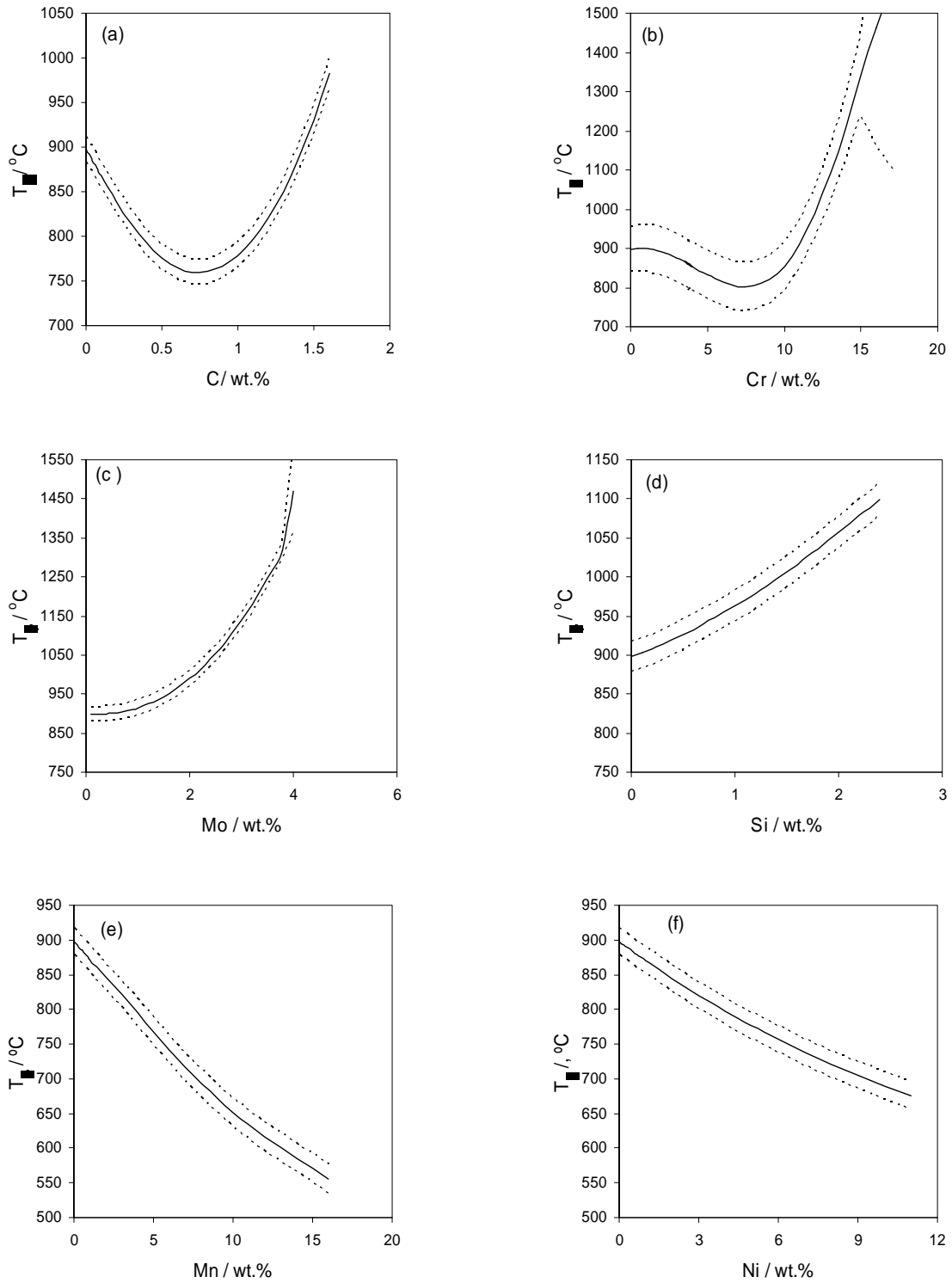
Figure 4. Comparison between the calculated and measured values of T_γ temperature using the seven models committee.
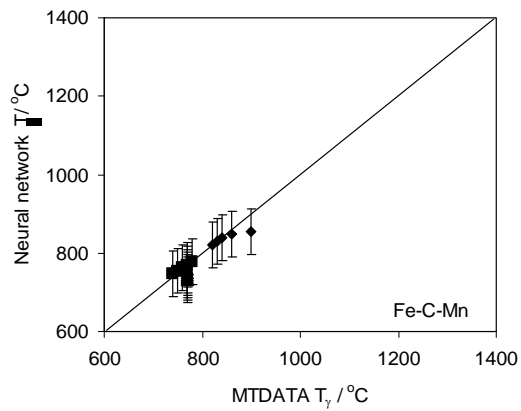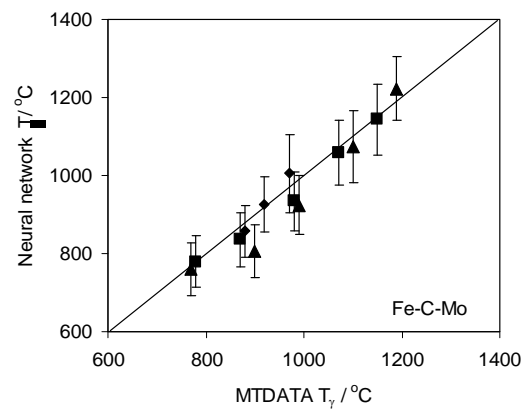
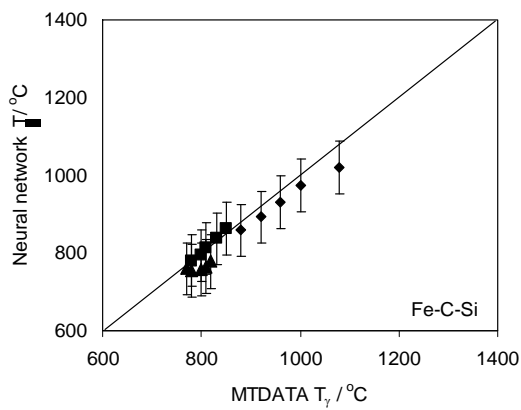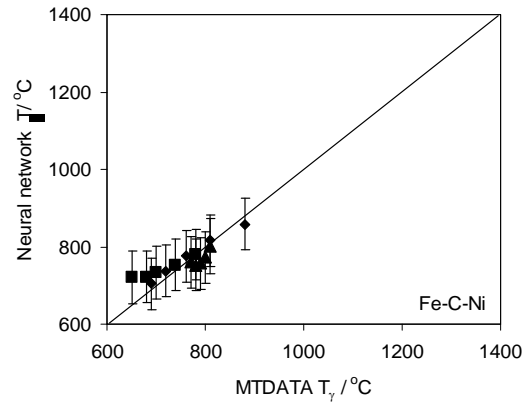**Figure 5.** Bar chart showing the significance (σ_w) of input variables influencing in T_γ temperature perceived by the model.

Figure 6. Evolution of T_γ with (a) carbon content in a Fe-C phase diagram, (b) chromium content in a Fe-Cr phase diagram, (c) molybdenum content in a Fe-Mo diagram, (d) silicon content in a Fe-Si diagram, (e) manganese content in a Fe-Mn diagram and (f) nickel content in a Fe-Ni diagram.

Figure 7. Comparison between MTDATA calculations and Neural Networks predictions for the T_γ in some ternary systems.

Figure 8. (a) Plot of the data measured versus calculated according to equation 4 and Table 3. Comparison of the results obtained by MTDATA, equation 4 and those predicted by Andrews in quaternary alloys according to Table 4, (b) Cr, (c) Ni, (d) Mo and (e) Si containing quaternary alloys

Figures

Figure 1. Carbon isopleth of the Fe-C-Ni system .C is 0.85 wt.% and Ni varies from 0 to 9 wt.%, Fe is set to balance. FCC_A1 and BCC_A2 stand for austenite and ferrite respectively.

Figure 2. Variation of (a) inferred noise level (σ_v), and (b) test error (T_{en}) as a function of the number of hidden units.

Figure 3. Comparison between calculated and measured T_γ temperature values: (a) training data and (b) test data.

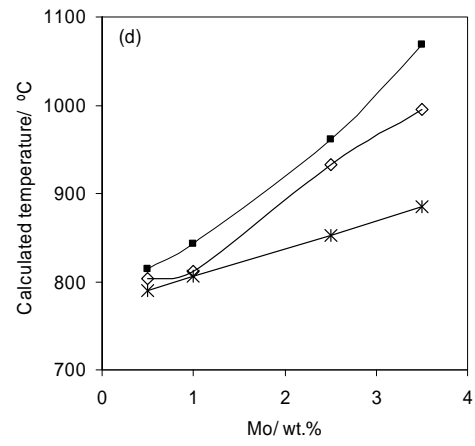*Figure 4.* Comparison between the calculated and measured values of T_γ temperature using the seven models committee.
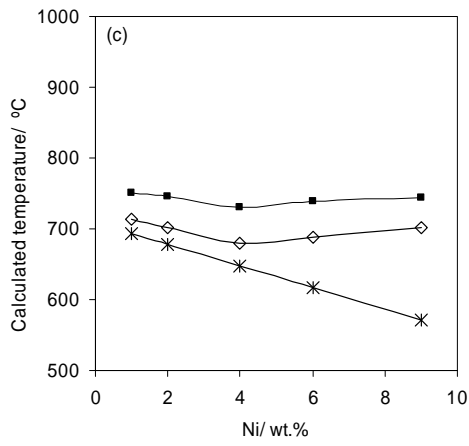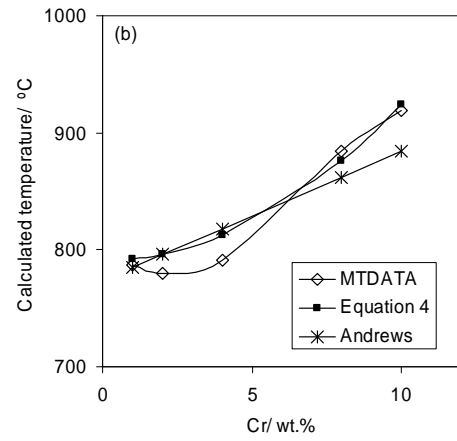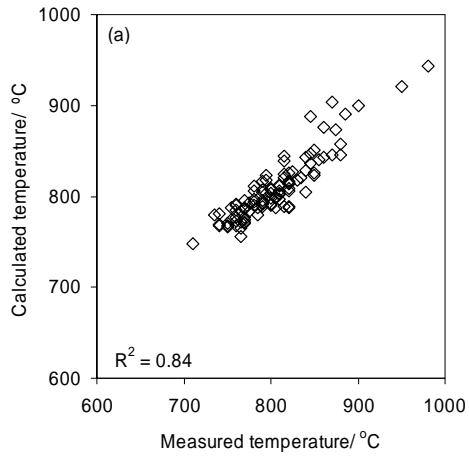
Figure 5. Bar chart showing the significance (σ_w) of input variables influencing in T_γ temperature perceived by the model.

Figure 6. Evolution of T$_\gamma$ with (a) carbon content in a Fe-C phase diagram, (b) chromium content in a Fe-Cr phase diagram, (c) molybdenum content in a Fe-Mo diagram, (d) silicon content in a Fe-Si diagram, (e) manganese content in a Fe-Mn diagram and (f) nickel content in a Fe-Ni diagram.

Figure 7. Comparison between MTDATA calculations and Neural Networks predictions for the T_γ in some ternary systems.

Figure 8. (a) Plot of the data measured versus calculated according to equation 4 and Table 3. Comparison of the results obtained by MTDATA, equation 4 and those predicted by Andrews in quaternary alloys according to Table 4, (b) Cr, (c) Ni, (d) Mo and (e) Si containing quaternary alloys.