Improving cryogenic grade copper with nonlinear models

Summary

Residual resistivity ratio is a standard measure of the purity and low temperature conductivity of copper. Oxygen free copper for cryogenic applications can have a conductivity two to three orders of magnitude higher at 4.2 K than at room temperature. There are several papers in literature indicating the effects of various impurities on the conductivity or resistivity of copper. Quantitative effects are usually described as linear variations. The general wisdom regarding the low temperature conductivity has favoured higher and higher levels of purity. However, at even lower concentrations of impurities, their effects are not always negative. Secondly, the effects are not very linear in terms of either conductivity, resistivity or residual resistivity ratio.

New techniques of nonlinear modeling have come up in the last ten-twelve years which have made the development of nonlinear empirical models of a higher quality possible, without the necessity of knowing the type and severity of the nonlinearities present in the relations. These techniques have been used successfully for a variety of applications in materials science, particularly for the prediction of material properties from the composition of the material. This article describes nonlinear modelling for residual resistivity ratio of oxygen free copper for cryogenic applications from Outokumpu Poricopper's production and experimental data.

Introduction

While copper is used for a large variety of purposes including wiring, plumbing and architectural applications, it has alternatives in the common applications. The level of know-how required to produce copper for these purposes is also relatively low. Among the highest value-added copper materials are oxygen free cryogenic grades of copper, which require a much higher level of know-how to produce and process without losing its low temperature conductivity, and do not have any good alternatives. Such grades of copper are used as stabilizers for superconductors], needed for large magnet windings. The amount of copper used is several times that of the actual superconducting material, which makes the demand for these grades of copper very significant.

The key quality measure of this copper is the residual resistivity ratio (RRR), which is defined as the ratio of the resistance at room temperature, 293 K, to the resistance at helium's boiling point, 4.2 K. The resistivity of copper at room temperature is not very sensitive to impurities. Typical values of RRR of oxygen free, cryogenic grade copper range between 200 and 700. RRR values of several thousand have also been reported. Most of the earlier literature shows linear effects of impurities and temperature on resistivity. Every impurity in solution was said to reduce the electrical conductivity. However, some of the impurities separate out as oxides forming another phase, which then do not influence the conductivity of the solution. Oxygen, thus, even reduces the effects of the impurities and has a positive effect on RRR upto a certain limit. Impurities like zinc which form such oxide inclusions do not reduce the conductivity. An interesting attempt to empirically model the RRR is based on atomic sizes of the impurities, in which the extra resistivity contributed by each impurity is a sum of linear and exponential terms of the atomic volume size factor. The main merit of this model is that the coefficients used on the linear and exponential terms are constants, and do not depend on the impurity.

New techniques of nonlinear modelling which have come up within the last ten-twelve years, have permitted the development of highly sophisticated nonlinear empirical models, without knowing the type and severity of nonlinearities present in the relations. It is also possible to combine process knowledge with this kind of empirical models, which often leads to better models. These new techniques have opened up new possibilities. It is now possible to develop accurate and reliable nonlinear models relating composition with material properties like RRR or tensile strength from production data, when the production data has sufficient information content.

There are hardly any processes or materials in this world which have absolutely linear characteristics. It is therefore wise to treat the nonlinearities rather than ignore them. To treat the nonlinearities, one can use new techniques of nonlinear modelling, like artificial neural networks. The proponents of linear techniques draw on their simplicity and the possibility of adding nonlinear terms in linear regression. Often this is not done, and is not efficient even if it is done. Nature does not follow the simplicities that we try to fit it in, using linear techniques.

Neural networks, on the other hand, have the so-called universal approximation capability which make them suitable for most function approximation tasks we come across in process industries. The user does not need to know the type and severity of nonlinearities while developing the models.

Production data

Besides RRR, concentrations of over 20 different elements are measured during the wire rod production of oxygen free copper. Most of these are of the order of 1 ppm or less, and therefore, have limited measurement accuracy. Some of the input variables show visible trends when RRR is plotted against them. Figure 1 shows that on an average, RRR decreases with increasing antimony content. Some of the concentrations are strongly correlated, as can be seen from Figure 2. These variables can cause a lot of trouble in the modeling, since they are not exactly independent variables, and make the Jacobian matrices (derivative matrix of prediction errors with respect to free parameters of the model) rank-deficient, and the parameter estimation problem ill-conditioned. If the variables are as strongly correlated as in Figure 2, the two variables contain more or less the same amount of information as one of those two variables. In such cases, it is often better to drop one of the two variables. For reasons of confidentiality, only a part of the data has been shown in these figures.

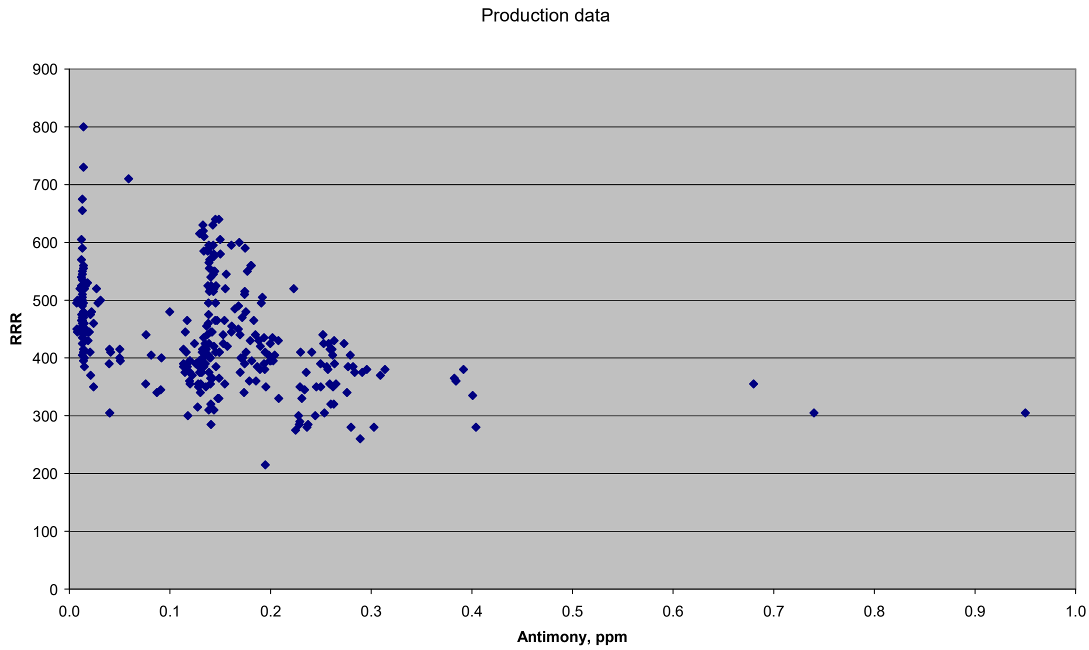

Figure 1. A part of the data plotted here shows how RRR decreases on an average when antimony content is higher.

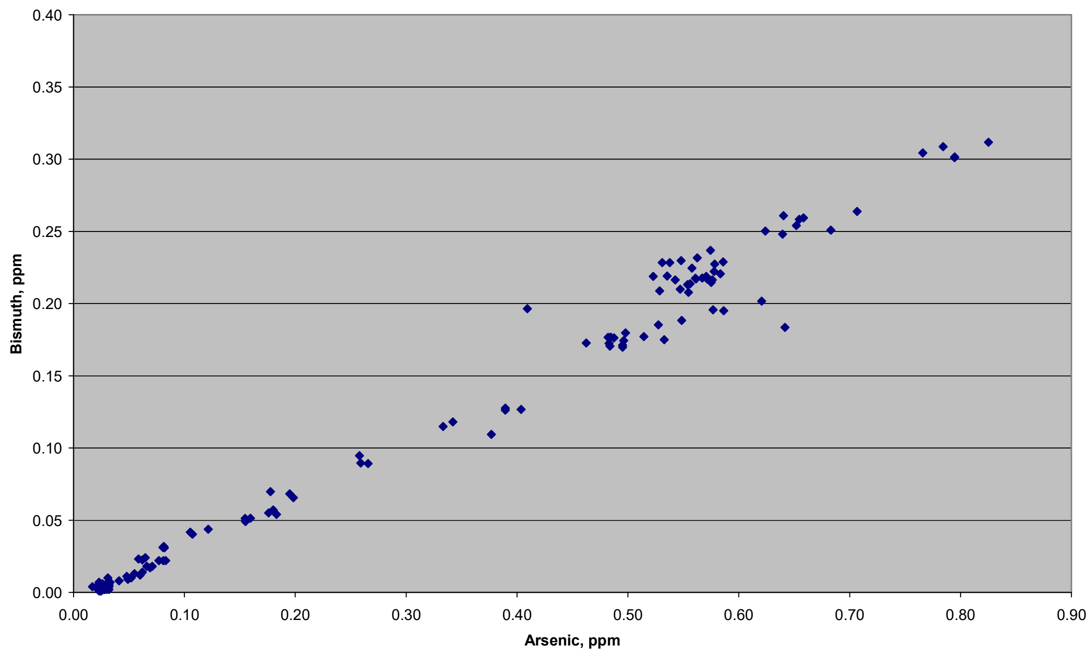

Figure 2. Some of the potential input variables were strongly correlated.

Production data is typically preprocessed before it is used for nonlinear modeling. NLS Preprocessing Software, which was used for this work, allows the user to see the basic statistics of the data, including the ranges, averages, standard deviations, etc. One can plot one variable against another, and see the covariance. Several advanced features include clustering of the data. It also performs division of the data set into training, test and validation sets, with specified conditions, and shows the imbalance in the division. It can also be used to identify possibly erroneous observations, and to study the accuracy and repeatability of the measurements, if suitable sets of observations are found in the data. Measurement accuracy and repeatability limit the accuracy of empirical models, at least when measured in terms of prediction error variance or standard deviation.

Nonlinear modeling and neural networks

Nonlinear modeling can be roughly defined as empirical modeling which takes at least some nonlinearities into account. Nonlinear modeling can be performed in many ways. The simpler ways include polynomial regression and linear regression with nonlinear terms. One can also use basis functions and splines, and in cases where the form of the nonlinearities is known, nonlinear regression can be used. Artificial neural networks are a set of efficient tools for nonlinear modeling, for reasons mentioned earlier, particularly the universal approximation capability of feed-forward neural networks.

Artificial neural networks resemble structurally and to a smaller extent functionally the networks of neurons in biological systems. Like the networks of neurons in the brains, artificial neural networks also consist of neurons in layers directionally connected to others in the adjacent layers (see Figure 3).

There are many different types of neural networks, and some of them have practical uses in process industries [5, 6]. Neural networks have been in use in process industries for about ten years. The multilayer perceptron is a kind of a feed-forward neural network. Most neural network applications in industries [5-12] ranging from concrete [10] to optical fibre cables [12] are based on them.

The output of each neuron *i* in a feed-forward neural network is given by

$$z_i = \sigma\left(\sum_{j=0}^{N} w_{ij} x_j\right)$$

where the activation function is often the logistic sigmoid, given by

$$\sigma(a) = \frac{1}{1+e^{-a}}$$

The incoming signals to the neuron are x_j, and w_{ij} are the weights for each connection from the incoming signals to the i^{th} neuron. The w_{i0} terms are called biases. This results in a set of algebraic equations which relate the input variables to the output variables. Thus, for each observation (a set of input and output variables), the outputs can be predicted from these equations based on a given set of weights. The training procedure

(Figure 4) aims at determining the weights which result in the smallest sum of squares of prediction errors. There are a variety of training methods in use today.

It is also possible to combine neural networks with physical models or other empirical models, which often lead to better solutions. This is referred to as hybrid modeling.

processMax+ systems

Advanced and complicated models like the ones often implemented in *processMax+* systems can be cumbersome to use in their raw form. Unlike simple regression models, the equations may look clumsy and the free parameters don't tell you very much unless you are very familiar with these kinds of nonlinear models. *processMax+* systems make it easy for the user to utilise the models without needing to understand the technology in detail. At the same time, this is no black box. If the user wants to see what the models are doing, a few clicks of the mouse will also show what is going on inside. Over the years, we have understood better and better the facilities which make the system more valuable for the user, and therefore, the *processMax+* systems of today are a result of almost continuous improvement of its features.

The second version of the *processMax+* systems can implement, in principle, almost any kind of empirical models, but they are particularly designed for nonlinear models based on neural networks. The models might also be combinations of neural networks with other kinds of models. Plain linear regression models can also be implemented.

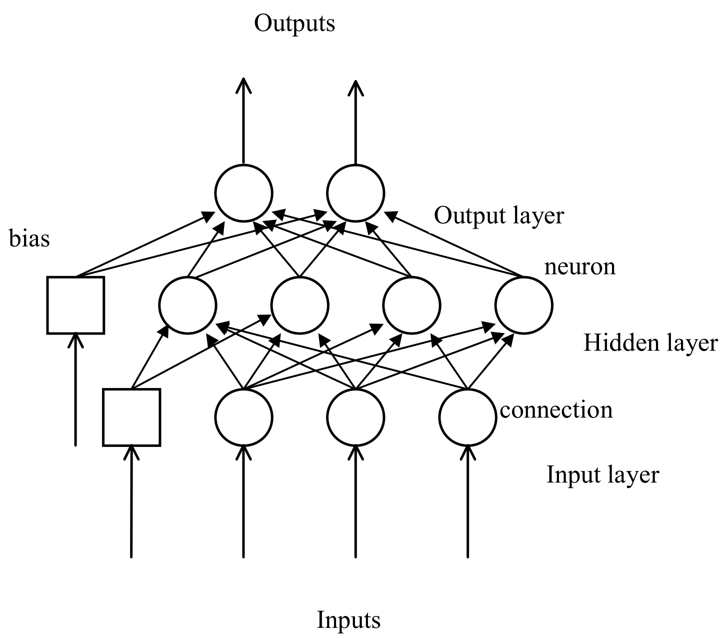

Figure 3. A typical feed-forward neural network has an input layer, an output layer and zero, one or two hidden layers.

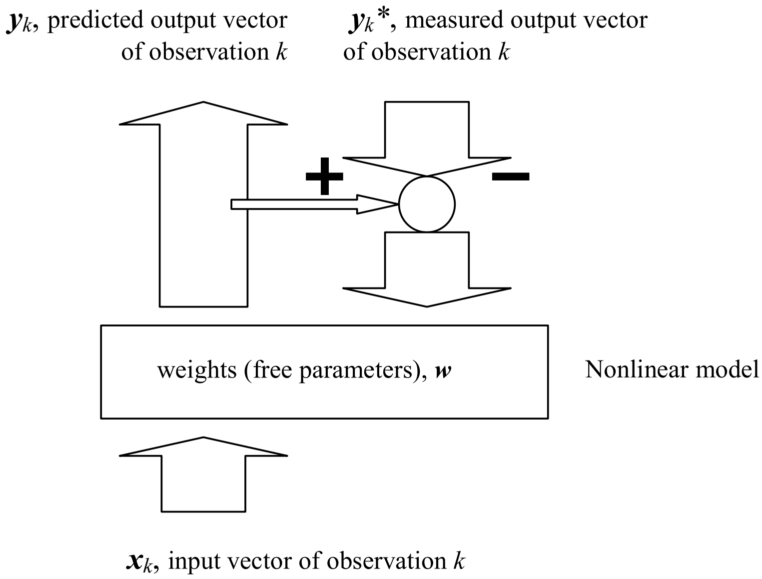

Figure 4. Training is the process of determining the weights, the free parameters of the model.

One general principle has been maintained from the beginning of its development. The *processMax+* systems will be extremely easy to use without a user's manual, and will take very little or no effort to get used to it. Today, people in industries are overwhelmed with new software of all kinds. *processMax+* systems are designed to make their life easier, and not add to the list of things they have to learn.

The models are available to every window in the system. Since the calculation of outputs from the models typically take something like a millisecond on today's PCs, many windows recalculate outputs several times while leaving the ”operating condition” (the set of inputs on the prediction window) unchanged. Almost all the windows have facilities for printing the results on the window. In most cases, the print command will resize the window before sending it to the printer, and then reset it to its standard size.

As readers of this magazine know, *processMax+* systems have earlier been used for mechanical properties [7, 8]. The first version was designed for prediction of tensile strength of steel products like wire rods, strips and plates. *processMax+* systems were then generalized to be able to implement other also material properties.

A few neural network models of RRR were implemented in the second version of *processMax+* systems, and Figures 5 and 6 are based on one of those models.

The first version had a prediction window (Figure 5) and an analysis window (Figure 6) for seeing the effects of the input variables. The prediction window shows the input variables and output variables, where the user can feed in the values of the input variables either manually or from a file, and a click on the “predict” command results in the prediction of the output variables. The analysis window lets the user select an input variable for the horizontal axis and an output variable for the vertical axis, and a click on the “plot” command results in a plot of the kind shown in Figure 6.

The second version of *processMax+* systems includes several advanced features. It is possible to feed in the desired values of the output variables, and the acceptable limits on the input variables, and the *processMax+* system can find suitable input variables. If there is no solution within the given limits, it shows the closest solution it can find. Alternatively, one can ask the *processMax+* system to determine the minimum and maximum values of the output variables within specified limits of input variables. This makes it a very useful tool for research and development engineers for product design. Ensuring that the product will have desired values of the output variables can also be used for quality control.

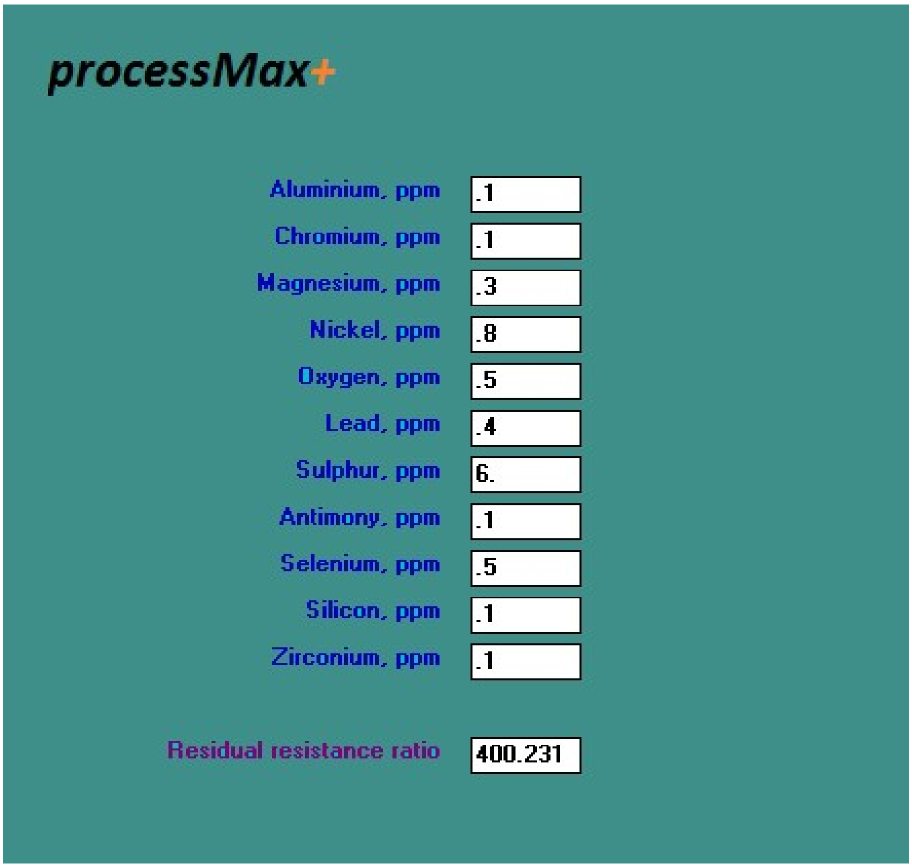

Figure 5. A typical prediction from one nonlinear model implemented in the *processMax+* system

Conclusions

Nonlinear models of RRR of oxygen free copper were developed from production data and implemented in a *processMax+* system for facile use. The nonlinear models were based on feed-forward neural networks. This is now seen as a useful asset in further research and development work on cryogenic grades of oxygen free copper, including in product development.

An important observation from this work, is that some elements are not detrimental to the low temperature conductivity in ppm amounts, while some elements like oxygen can actually increase the RRR values. This provides a possibility of increasing RRR more easily than just by increasing the purity.

This work has generated additional interest and enthusiasm to study more systematically the effects of each of the impurities by well planned experiments.

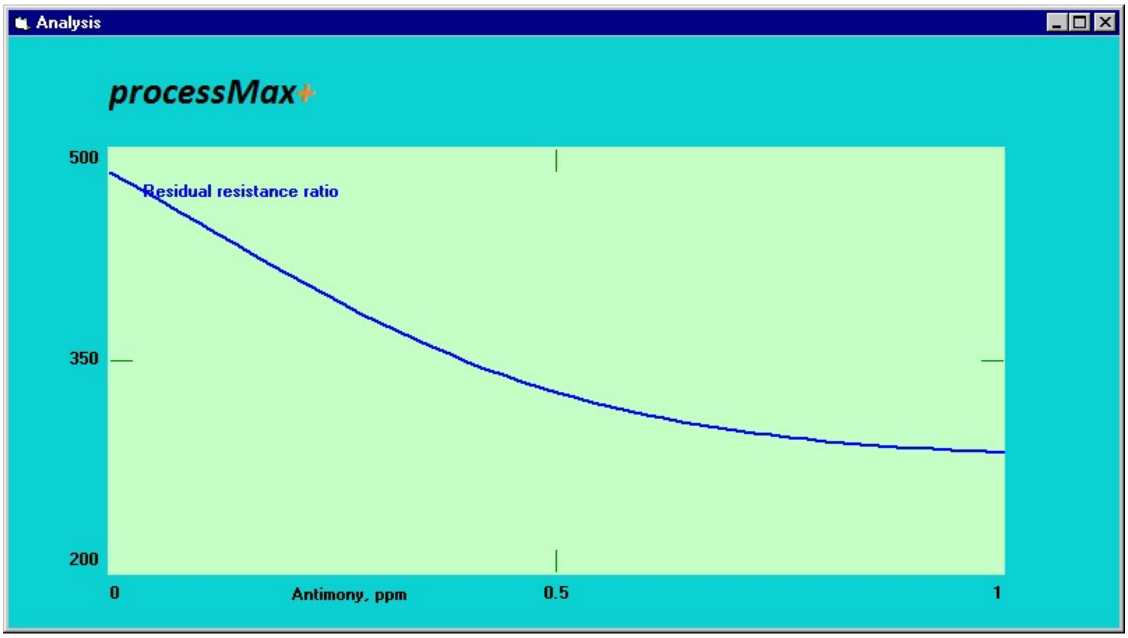

Figure 6. The plot shows the effect of antimony content predicted by the model, keeping other variables constant