Integration of Adaptive Neuro Fuzzy Inference Systems and Principal Component Analysis for the control of tertiary scale formation on tinplate at a hot mill

M. Evans* and **J. Kennedy**

*Corresponding author
College of Engineering, Swansea University, Singleton Park, Swansea, UK, SA2 8PP.
m.evans@swansea.ac.uk

ABSTRACT

Scale is highly detrimental to surface quality for tinplate products. There are a large number of process variables at a typical hot mill and principal component analysis is a well-known technique for reducing the number of process variables. This paper estimates the principal components associated with the hot mill process variables and puts these through an Adaptive Neuro Fuzzy Inference System (ANFIS) to find those hot mill running conditions that will minimise the amount of scale observed on the bottom of the rolled strip. It was found that the variation observed in all the hot mill process variables could be captured through the use of just six principal components, and that using just three of these in an ANFIS was sufficient to identify those operating conditions leading to coils being produced with a consistently low scale count. Specifically, it was found that the best operating conditions for the hot mill were when the first component was lower than -0.098 the second lower than 0.8058 and the third higher than -0.482. These ranges in turn corresponded to certain hot mill temperatures that depended to some extent on the base chemistry of the incoming slab.

Keywords: Scale; Hot Mill; Fuzzy Logic; Neural Networks; Adaptive Neuro Fuzzy Inference System; Principal Components.

1. Introduction

For high-end flat steel products scale is clearly detrimental to surface quality. Scale is formed at the hot mill but it has its greatest impact on quality after further processing. Scale can have a particularly negative impact on the quality of tinplate products. This occurs due to the scale interfering with the interface between the steel substrate and the tin, resulting in a product surface that is not suitable for the tight tolerances required by customers of tinplate products. A typical hot mill has many process variables associated with it including geometries and chemistries associated with the incoming slab and various temperatures associated with the different mills - all of which could potentially influence the final scale count on the finished coil. The aim of this paper is to demonstrate that a combination of well-known data mining techniques can be used to extract important information from a multidimensional industrial data set that enables engineers to control the hot mill process and in particular to control the formation of scale at the hot mill.

There are many well used and understood data mining techniques available for analysing process data. However, the hot mill process studied in this paper has a number of characteristics that prevent any one of these, on their own, being able to successfully model and predict the scale count on coils produced at the hot mill – see Fig. 1. First, the hot mill process has many process variables associated with it and many of these are highly correlated. One obvious solution to this problem is too combine them all into a smaller number of principal components that by construction are uncorrelated with each other. However, the usual approach of then regressing the scale count on these components is going to be inadequate here because count data of this nature are well known to have non normal distributions. More typically, count data is traditionally modified using the Poisson distribution or a suitable generalisation of it.

Secondly, the relationship between the hot mill scale count and the principal components is not the same over all values for the process variables. This would suggest merging the technique of regression trees with principal components, as this would enable the principal components to have different effects on the scale count over different sub regions of the full data set. However, this straight forward merger would tend to create very artificial discontinuities in the modelled response surface that is not representative of the scale count data. It is more appropriate to allow the relationship between the hot mill scale count and principal components to change slowly around certain values for some or all of the principal components. All of these requirements can be meet by using a fuzzy neural network where the inputs to the network are the principal components and where the data is spit into various sub divisions and in each subdivision the mean of the Poison distribution for scale count is related to the principal components. The parameters relating the mean count to the principal components are then optimised by allowing the fuzzy network to maximise the log likelihood of observing all of the scale counts in the data set. Sharp discontinuities between each sub division of the data are avoided by fuzzyfying where these breaks occur using S shaped membership functions.

The novelty of this paper is that the resulting model is a generalisation of the Adaptive Neuro Fuzzy Inference System (ANFIS), first put forward by Jang (1993), to be referred to as ANFISPC in this paper. The unique features of this generalisation over the original ANFIS is the use of principal components as inputs, the use of a distribution for the variable being modelled (n this case scale count) and optimisation using a likelihood function

(associated with this distribution) instead of the average squared prediction error. This modified ANFIS will be of use in modelling count data obtained from other processes or fields of study, and offers the potential to produce more accurate count predictions. Further, with its emphasis on distributions, these predictions will be more useful and comprehensive in nature. Instead of a simple mean count prediction for a particular set of conditions, the model will produce a predicted distribution for the count so that assessments can be made on product quality or consistency in production.

To achieve this aim, the paper is structured as follows. First, a review of scale and its formation is presented, and this is followed by a description of the hot mill at the Port Talbot works and how the scale data is collected there. This is followed by a section describing what principal components are and how they can be constructed. The section after this then describes how these components can be used within an Adaptive Neuro Fuzzy Inference system. The performance of this system is then reviewed in the penultimate section together with a discussion on how this model can be used to control scale formation at the hot mill. Some recommended operating conditions are given in the conclusions section.

2. Literature review on scale type and formation

2.1 Scale

Scale is an oxide layer that builds up on the surface of the steel when it is exposed to environmental conditions. Oxide can form at any point during the life cycle of the product; however it is particularly likely to occur at the hot mill. This is due to the high temperatures at the hot mill, which increases the speed of any chemical reaction. Bolt (2003) states that the important properties to consider for the oxide layer are thickness, composition, adhesion and structure.

The thickness of the oxide layer will also be determined by its processing conditions. Bolt (2003) states that the most important parameters for thickness are the finishing temperatures and the coil gauge. Depending on the temperature of the strip, oxide could be formed at the run-out table (ROT) and also during coiling but that scale formation at these stages is rather modest. This is supported in the work by Yang et al. (2008), where they discuss the influence of Silicon on thickness. They investigated steels with Silicon contents ranging from 0.01-1.91 wt. % and concluded that oxidation rates and scale thickness decreased with increasing Silicon. They also concluded that different oxides are formed at different percentages of Silicon. This work is important because it demonstrates the importance of limiting the chemistry range to avoid scales with different properties forming. Taniguchi et al. (2001) further demonstrated the importance of limiting the chemistry range. They found that when Silicon is at high percentages (1.14 wt. % and above), the penetration depth of the scale increases with increasing Silicon. The work of Munther and Lenard (1999) concluded that thick scale provides a degree of lubrication to the work rolls, while a thin scale has a higher coefficient of friction and is more adherent to the metal substrate, which makes it harder to remove. However, Silicon is not likely to play an important part in this analysis due to the low levels present in the steel investigated.

The composition of these formed oxides influences the properties of the scale. The classical three-layer scale model reveals that when pure iron is oxidised under normal conditions, Wustite (FeO) is formed closest to the steel substrate followed by magnetite (Fe_3O_4) and hematite (Fe_2O_3). Each of these oxides has their own individual properties. Bolt

(2003) states that at hot rolling temperatures, Wustite has a high plasticity, but is extremely brittle at room temperature, which means that a large amount of Wustite will result in poor scale cohesion at room temperatures. Hematite is extremely hard and brittle and is highly detrimental during all stages of processing due to the increased work roll wear it causes and the poor strip quality it creates. Magnetite is the phase that is least brittle at room temperature. This phase can be tolerated in small quantities, because it is not as hard as hematite and shows some plasticity at hot rolling temperatures. Thinner scales are more adherent than thicker scales and the reason for their increased adherence is the higher amounts of hematite and magnetite. This is because magnetite and hematite are a lot harder than Wustite and so is harder to remove through further processing of the coil.

2.2 Temperature

The thickness of the scale is proportional to temperature and time. Sun et. al. (2004) observed that as temperature and oxidation time increases the scale thickness also increases. It was observed that for the first 20 seconds the scale increased in thickness according to linear growth, and then by parabolic growth after that. They explain this observation by the transport mechanism of oxidation. Initially the scale is very thin which allows rapid diffusion of oxygen with the metal at the iron-air interface. As the scale layer increases in thickness the transport rate will reduce and the diffusion will obey parabolic laws.

The composition of the scale that develops is also highly temperature dependent. At temperatures greater than 570° C the main composition of scale is Wustite, which is generally the thickest layer. This is then followed by magnetite and hematite. The literature on what composition of scale is observed at different temperatures is highly variable, mostly due to different hot mill layouts, different types of steel, different experimental methods, and different temperatures. The literature review of Bolt (2000) concludes that as the temperature is increased the more likely it becomes that the scale formed will be composed of a higher percentage of magnetite and hematite at the expense of Wustite.

2.3 Phosphorus

If the diffusion of iron into the scale is suppressed then magnetite and hematite will form in preference to Wustite, and this makes it harder to remove the scale. The composition of the scale can be influenced by the steel composition. For tertiary scale formation, Bolt (2000) states that Phosphorus is of particular importance. This is because Phosphorus can suppress diffusion of iron into the scale. This suppression can, locally, completely cut off the diffusion, which means that as the steel is oxidised magnetite and hematite will form in preference to Wustite.

3. The hot mill process

There is no single process variable that can be attributed to scale formation, so to be able to determine variables that are significant in scale formation at the hot mill, it is important to identify all the process variables through a detailed description of the hot mill process. A schematic view of this process is given in Fig. 2. As illustrated in this figure, slab (typically about 0.25 m thick) formed at the continuous caster enters the hot mill to be rolled

into thin sheet steel. The slab entering the hot mill will also have a specific chemistry that was determined at an early stage of production (i.e. before the slab was cast).

The purpose of the hot mill is to take this thick slab and roll it into thin sheet steel suitable for forming packaging cans, car body panels and other products. At the end of the hot mill the rolled steel is stored as coils of thin sheet steel. Before rolling, the slab must be "softened" by heating it up (to over 1000^0C) in reheat furnaces – shown at the start of Fig. 2. As temperature is one of the most important determinants of oxidation, it is at this stage of the hot mill process that scale starts to be observed on the surface of slab. To be able to understand the impact of this critical area on scale formation, multiple temperature readings are made in the reheat furnaces because these furnaces are very large and consequently have a non- uniform temperature distribution within them.

In an attempt to remove the scale created in the reheat furnaces, the heated slabs pass through the horizontal scale breakers shown in Fig. 2. The slabs then proceeds to the interstand spray system for further scale to be removed by this spray. The slab then proceeds to the rougher mills to be rolled into coils. At the edging and rougher mills the gauge (thickness) of the incoming slab is reduced from around 250 mm to 35 mm (and the processed steel at this stage is often termed the transfer bar). The hot mill at the Tata works in Port Talbot (UK) has a reversing rougher mill, hence several passes are necessary to achieve the required transfer bar thickness (five passes are most likely for the product under investigation). The transfer bar is then held at the coil box to maintain the heat of the bar prior to entering the finishing mill. Various temperature measurements are made across the coil at the rougher mill.

The crop shear temperature is the temperature just before the transfer bar enters the finishing mill. The crop shear temperature is likely to be of vital importance for predicting scale formation because tertiary scale is typically formed during the early stages of the finishing mill. Again, various temperature measurements are made across the coil just before the coil enters the finishing mill. On entry into the finishing mill the slab encounters the descalers again. The purpose of the finishing mill is to further reduce the gauge to between 1.4 - 18.0 mm depending upon the intended market for the finished steel. The finishing mill also controls the shape and temperature of the strip that in turn controls the metallurgical properties achieved by the strip. Various temperature measurements are also made across the coil at the finishing mill.

The phase transformations within the rolled steel occur on the run-out table (ROT) and the cooling is controlled with a water cooling system. The temperature control of the run out table is vital in controlling the austenitic phase transformation. The temperature, cooling rate and cooling path will affect the microstructure present, which in turn will control the mechanical properties of the steel. Finally, the rolled sheet steel is formed into a coil for storage purposes. Table 1 list all the hot mill process variables identified by this discussion of the process.

4. Measuring scale count

The Parsytec inspection system is a camera operated system that detects defects on the surface of the coil. The detection of defects is achieved by the identification of all non-homogeneous sections of the strip. When the affected sections are detected then the Parsytec software classifies the detected sections into features. These features are compared to the

online defect catalogue and then the best match is used for labelling the defects. It is possible for defects to be misclassified or not classified. This can occur because the environment around the cameras has a lot of potential contamination including:

- Water- from the descaler and cooling sprays at the finishing mill.
- Overlapping defects.
- Substrate appearance- every steel composition will have a different surface appearance under the parsytec cameras that makes detection rates variable.

The Parsytec system monitors the top and the bottom of the strip, but they are located in different locations in the mill. The top surface monitor is located at the start of the ROT, while the bottom surface monitor is located at the end of the ROT. The Parsytec system counts the number of formed scales at the top and bottom of each coil on the ROT. Due to the fact that under high coiling temperatures scale can continue to grow on the ROT and that due to water and other contamination after leaving the finishing mill it was determined that the bottom surface monitor for scale would give more reliable results. The dataset is constructed using data associated with coils from Tata Steel Europe's Port Talbot Hot Mill, collected using the Parsytec system described above. Each of the coils in the dataset is intended for a high surface quality tinplate application, so scale levels must be low. To minimise the effect of different heating/cooling conditions only coils of a single gauge of 2.1mm are included in the data set.

Such coil data was collected between September 2009 and March 2010 and for each coil there was a recorded value for all the process variables described above. The dependent variable is the bottom scale count from the Parsytec system – denoted as y in Table 1. In total the data set consisted of some 1530 data points. Table 1 lists these process variables together with the means and standard deviations over the whole sample of data. These process variables are designated as x_j.

5. Fuzzy logic, neural networks and principal components

The Adaptive Neuro Fuzzy Inference System (ANFIS), first put forward by Jang (1993), integrates neural network and fuzzy logic principles. The advantage of such a combination is that the resulting predictive model is not just of a black box because it also provides guidance to operators of manufacturing systems as to how to achieve such predictions –through the appropriate setting of the levels for the process variables. The basic philosophy behind the original ANFIS model was to split the data up into subsets and then to fit a linear model to the data in each subset. This part of the system is therefore similar to the characterisation and regression tree methodology first put forward by Breiman et. al. (1984). The main difference is that instead of using an average in each subset of data a linear regression is used instead. In the use of such linear models, this part of an ANFIS is similar to the multivariate adaptive regression splines (MARS) put forward by Friedman (1991). The fundamental difference between an ANFIS and these more traditional data analysis techniques is that the boundaries defining each sub set of data are blurred.

5.1 Principal component analysis

Principal component analysis (PCA), see for example Jolliffe (2002) for a good review of this technique, is an extensively used technique of multivariate linear data analysis. The main objectives of PCA are to reduce the dimensionality of the data whilst still

maintaining the original variability and to remove the correlation amongst all the process variables. Such principal components (PC) can be calculated as follows. The p process variables (x_1 to x_p) are first standardised to have a zero mean and unit variance to avoid the problems associated with units of measurement. (For the hot mill data this is achieved by subtracting from each observation on x_j the mean value and dividing the result by the standard deviation (both of which are shown in Table 1). Then p linear combinations of these variables can be formed as follows

$$\begin{aligned} PC_1 &= a_{11}z_1 + a_{12}z_2 + \text{.....\_} a_{1p}z_p \\ PC_2 &= a_{21}z_1 + a_{22}z_2 + \text{.....\_} a_{2p}z_p \\ &\vdots \\ PC_p &= a_{p1}z_1 + a_{p2}z_2 + \text{.....\_} a_{pp}z_p \end{aligned} \qquad \text{(1a)}$$

where the z variables are the standardised values of the x process variables. The so called loadings, a_{11} to $a_{1p,}$ in the first principal component, are then chosen so as to maximise the variance of PC_1 subject to the normalising condition

$$a_{11}^2 + a_{12}^2 + \text{.....} + a_{1p}^2 = 1 \qquad \text{(1b)}$$

PC_1 is then said to be the first principal component and is a linear function of the z's (and thus the process variables) that has the highest variance. PC_2 then has the next highest variance and so on until all the variation in the z's is picked up by the p principal components. That is

variance(PC_1) + variance(PC_2) + ... + variance(PC_p) = variance (z_1) + variance(z_2) + ... + variance(z_p) = p

with

variance(PC_1) > variance(PC_2) > ... >variance(PC_p)

Various different algorithms can be used to find values for these loadings, ranging from the simple summation method put forward by Burt (1945) – (of which a good review can be found in Childs (1970)) - to algorithms that calculate a spectral decomposition of the correlation matrix amongst all the x variables (see for example the text by Mardia et al. (1979)). The Eigen values from this decomposition measure the variation in all the process variables explained by each principal component. Thus the first Eigen values from this decomposition measures the variation in all the process variables explained by the first principal component. The Eigen Vector from this decomposition contains the required loadings for each principal component. By construction these principal components are also all orthogonal and hence completely uncorrelated with each other.

Principal components can also be given a best fit line interpretation as well. This is best illustrated using just two of the process variables from the hot mill – the maximum rougher mill temperature and the average rougher mill temperatures, x_{13} and x_{15} respectively. Let **V** be the variance - covariance matrix for these two variables

$$\mathbf{V} = \begin{bmatrix} 1 & 0.9064 \\ 0.9064 & 1 \end{bmatrix} \tag{2a}$$

Down one of the diagonals of **V** are the variances for the standardised values for x_{13} and $x_{15,}$ which because of this standardisation, are by definition equal to unity. The other diagonal contains the covariance between x_{13} and x_{15}. The eigenvalues and eigenvectors of the symmetric matrix **V** are calculated to be

$$eigenvectors = \begin{bmatrix} a_{11} = 0.7071 & a_{21} = -0.7071 \\ a_{12} = 0.7071 & a_{22} = 0.7071 \end{bmatrix} \quad eigenvalues = \begin{bmatrix} 1.906 \\ 0.094 \end{bmatrix} \tag{2b}$$

In Eq. (2b) the matrix called *eigenvectors* is in fact made of up two eigenvectors, with the first column of the matrix being the first Eigen vector and the second column the second Eigen vector. If z_{13} and z_{15} are the standardised values for x_{13} and x_{15}, the total variation in z_{13} and z_{15} is two and the principal component

$$PC_1 = a_{11}z_{13}+a_{12}z_{15} = 0.7071z_{13} + 0.7071z_{15} \tag{2c}$$

picks up 1.906/2 = 95.3% of this total variation. The final principal component

$$PC_2 = a_{21}z_{13}+a_{22}z_{15} = -0.7071z_{13}+ 0.7071z_{15} \tag{2d}$$

picks up the remaining 0.094/2 = 4.7% of this total variation. Also PC_1 and PC_2 are completely uncorrelated as shown in Fig. 3a.

Further, Cattell (1952) has suggested that only those principal components having an Eigen value greater than 1 should be considered as essential and therefore retained in the analysis. This simple rule allows for a substantial reduction in the number of variables to be included in any model of the hot mill.

Looking at the principal components in a less mathematical way allows a comparison with the technique of least squares to be made. In Fig. 3b, the standardised values for the average and maximum rougher mill temperatures are plotted against each other and the slope of the solid best fit line shown in the figure was chosen so as to minimise the total squared residual defined as

$$\sum_{k=1}^{n} e_k^2 = \left[z_{13,k} - \beta z_{15,k}\right]^2 \tag{3a}$$

where β is the slope of the best fit line. This best fit line and the variation around it can therefore be written as

$$z_{13} = \beta z_{15} + e = 0.9064 z_{15} + e \tag{3b}$$

where e represents the residual. As the mean value for e is by the definition of a best fit line zero, minimising Eq. (3a) is equivalent to minimising the variance of the residuals. One of

these residuals is highlighted in Fig.3b, and so the least squares technique positions the best fit line so as to minimise all the (squared) **vertical** distances between the data points and the line. The principal components constructed in Eqs. (2c,d) can also be given a residual interpretation. For example, take the principal component that has the **smallest** variance, i.e. that given by Eq. (2d) and write it as,

$$z_{13} = (a_{21}/ a_{21})z_{15} + (PC_2 / a_{21}) = 1.0z_{15} + (PC_2 / 0.7071) \quad (3c)$$

Notice that Eqs. (3b,c) have a similar form, and so (PC_2 / 0.7071) can be given a residual interpretation similar to e, whilst $z_{13} = 1.0z_{15}$ can be given a best fit line interpretation. This is because PC_2 has the smallest (i.e. minimised) variance of the two principal components. This best fit line is shown as the dashed line in Fig. 3b. However, (PC_2/0.7071) is not the same residual as e. In fact (PC_2/0.7071) is the distance along a line joining a data point to the best fit line given by Eq. (3c) - which is perpendicular to that best fit line. Thus principal component analysis positions a best fit line so as to minimise all the (squared) **perpendicular** distances between the data points and the best fit line.

5.2 Binary partitioning and regression trees

Fig. 4 describes a binary partition of a hot mill process using just the first two principal components of the hot mill process variables. The variable y in this figure is the scale count on the bottom of the processed sheet just before final coiling at the hot mill. The number of principal components determines the number of layers present in this partition and consequently the number of sub divisions of the data. In the illustration of Fig. 4 there are two components and so two levels and four sub divisions. The decision tree partitions the input space into a number of non-overlapping rectangular regions depending on the values for PC_1 and PC_2. The tree identifies four simple if – then rules. The parameter *a*, which has the same units as variable PC_1, defines where the first split in the data occurs such that all values for PC_1 that are less than or equal to *a* form one subset (together with the values for PC_2 that were recorded at these values for PC_1) and the remaining values for PC_1 form the other sub set. These two subsets of data are then further split depending upon whether the values for PC_2 are above or below the parameter *b*, where *b* is in the same units as variable PC_2. Parameters *a* and *b* are referred to as decision rule parameters. This is a neat way to describe the running conditions at the hot mill and will provide information on how to control the amount of scale formation.

In each sub division the response of interest y (i.e. scale count) is modelled using separate functions, label μ_i. The subscript i in μ_i defines the (four) sub divisions of the data. Each μ_i is the mean (or possibly the median) scale count associated with each sub division:

$$\mu_i = \frac{\sum_{k=1}^{n_i} y_{ik}}{n} \quad (4a)$$

where y_{ik} is the scale count made on the kth coil processed under all the different process conditions making up sub division i. n_i is the sample size of sub division i (number of coils process under the hot mill conditions corresponding to this sub division). Essentially, this binary regression model states that under certain process conditions there is a mean or typical scale count and in addition to this there will be a frequency of occurrence of other scale count

values around this mean. Traditionally, this frequency of occurrence of count data is modelled using the Poisson distribution

$$P[y_i = v] = \frac{e^{-\mu_i}\mu_i^v}{v!} \tag{4b}$$

$P[y_i = v]$ reads the proportion of all coils belonging to sub group i (that are measured for scale) that have a scale count equal to v. The main limitation of the Poisson distribution is that the mean and the variance of scale count in each sub group must be equal. It is clear from Table 1 that before splitting the sample up this is not the case – with the variance being much larger than the mean (but it may be true for sub groups). The equality of mean and variance can be relaxed through the use of the negative binomial distribution. The Negin Q version of this distribution, as put forward by Cameron and Trivedi (1998), takes the form

$$P[y_i = v] = \frac{\Gamma(\theta_i\mu_i^Q + v)}{\Gamma(1+v)\Gamma(\theta_i\mu_i^Q)}\left(\frac{\mu_i}{\theta_i\mu_i^Q + v}\right)^v\left(\frac{\theta_i\mu_i^Q}{\theta_i\mu_i^Q + v}\right)^{\theta_i\mu_i^Q} \tag{4c}$$

where Q and θ_i are additional parameters defining this negative binomial distribution. The variance in the scale count for sub group i is then given by

$$Var[y_i] = \mu_i\{1 + 1/\theta_i\ \mu_i^{1-Q}\} \tag{4d}$$

Thus when Q = 1 and as $\theta_i \rightarrow \infty$ the variance becomes equal to the mean scale count in sub group i, and so the Poisson distribution corresponds to this restriction. The role of θ_i is therefore to allow the mean and variance to differ. Also, when Q = 1 (and with θ_i not equal to infinity) the Negin 1 variant of the negative binomial distribution is obtained and when Q = 0 the Negin 2 variant of the negative binomial distribution is obtained.

This simple visualisation of the hot mill process can be easily generalised. For example, each μ_i could be made a function of the principal components PC_1 and PC_2 and for each region there would be a separate function.

$$\mu_i = \exp(\beta_{i,0} + \beta_{i,j}\sum_{j=1}^{2} PC_j) \tag{4e}$$

Eq. (4e) is an empirical specification, but if the PC_j variables are mainly temperature dependent this specification would have some theoretical grounding to it as well. Further generalisations could include the use of the squares of the process variables in Eq. (4e). Eq. (4e) essentially states that over all the hot mill operating conditions associated with sub division i, scale counts follows a distribution whose means shifts with these conditions.

However as more and more sub groups are created, by using more and more principal components, so the hot mill operating conditions associated with sub division get narrower, it is to be expected that the β_{ij} in Eq. (4e) become negligible so the mean of the scale count distribution does not shift over such sub regions.

5.3 Fuzzy Logic

The problem with the approach outlined in the sub section above is that the resulting modelled response surface is highly discontinuous in that it changes abruptly at the decision rules – *a* and *b*. This problem is overcome by fuzzyfying the decision rules. For example, the crisp decision rule associated with the left most branch of the binary tree in Fig. 4 is

If $PC_1 \leq a$ and $PC_2 \leq b$ then $y = \mu_1$

It is well known that fuzzy expert systems use membership functions to quantify possibilities (Jang et. al. (1997)), but the context is very different from possibility as defined by statisticians. A possibility is a fuzzy measure indicating the degree of evidence or belief that a certain value for say PC_1 belongs to a subset, say set $PC_1 \leq a$. A membership function has a value between 0 and 1 such that PC_1 values further and further below *a*, have membership values closer and closer to one. Similarly, PC_2 values further and further below *b*, have membership values closer and closer to one. A common functional form used for these membership functions is the sigmoidal function

$$\gamma_{PC1 \leq a} = \frac{1}{1+exp[\rho_1(PC_1 - a)]} \qquad \gamma_{PC2 \leq b} = \frac{1}{1+exp[\rho_2(PC_2 - b)]} \tag{5}$$

where ρ_1 and ρ_2 are parameters requiring estimation. The values for ρ determine the steepness of the membership functions about *a* and *b*. So the further PC_1 is below *a*, the greater will be the value for $\gamma_{PC1} \leq a$, indicating a stronger belief that this value for PC_1 belongs to the set $PC_1 \leq a$. $\gamma_{PC1 \leq a}$ varies over the range 0 to 1, with 1 indicating the strongest possible belief and this will occur the further PC_1 is below *a* (when $PC_1 = a$, the membership function takes on a value 0.5). The further PC_2 is below *b*, the greater will be the value for $\gamma_{PC1 \leq b}$, indicating a stronger belief that this value for PC_2 belongs to the set $PC_2 \leq b$. $\gamma_{PC2 \leq b}$ varies over the range 0 to 1, with 1 indicating the strongest possible belief and this will occur the further PC_2 is below *b.* The theory of fuzzy logic suggests various ways to combine membership functions and this is discussed in the next sub section.

5.4 Combining neural networks, binary classification, fuzzy logic and principal components

Eqs. (4a,4c,5) contains a number of unknown parameters – ρ_1, ρ_2, a, *b*, μ_i and θ_i. A neural network is one possible framework within which the values for these unknown parameters can be estimated. The Adaptive Network –based Fuzzy Interference System (or ANFIS for short) of Tsoukalas and Uhrig (1997) is one such neural network. There are various ANFIS architectures, but one using a first order Sugeno fuzzy model put forward by Takagi and Sugeno (1995) is the most common. The ANFIS proposed by Tsoukalas and Uhrig requires modification to incorporate principal components and a negative binomial distribution for the scale count data. This modified architecture is shown in Fig. 5 for the decision rules shown in Fig.4:

Rule 1: If $x_1 \leq a$ and $x_2 \leq b$, then $y = \mu_1$
Rule 2: If $x_1 \leq a$ and $x_2 > b$, then $y = \mu_2$
Rule 3: If $x_1 > a$ and $x_2 \leq b$ then $y = \mu_3$
Rule 4: If $x_1 > a$ and $x_2 > b$, then $y = \mu_4$

In the first layer of the ANFIS shown in Fig. 5, each of the k values for PC_1 and PC_2 are given membership quantities using the following sigmoidal functions

$$\gamma_{PC1\le a} = \frac{1}{1+\exp[\rho_1(PC_1 - a)]} \text{ with inverse (INV) } \gamma_{PC1>a} = 1 - \frac{1}{1+\exp[\rho_1(PC_1 - a)]} \tag{6a}$$

$$\gamma_{PC2\le b} = \frac{1}{1+\exp[\rho_2(PC_2 - b)]} \text{ with inverse (INV) } \gamma_{PC2>b} = 1 - \frac{1}{1+\exp[\rho_2(PC_2 - b)]} \tag{6b}$$

where ρ_1 to ρ_2 and *a* and *b* are parameters requiring estimation. In layer 2 weights are determined that represent the possibility that each pairing for the k values of PC_1 and PC_2 belong to one of the four sets given by the decision rules above. These weights are given by

$$w_1 = (\gamma_{PC1} \le a)(\gamma_{PC2} \le b) \tag{7a}$$

$$w_2 = (\gamma_{PC1} \le a)(\gamma_{PC2} > b) \tag{7b}$$

$$w_3 = (\gamma_{PC1} > a)(\gamma_{PC2} \le c) \tag{7c}$$

$$w_4 = (\gamma_{PC1} > a)(\gamma_{PC2} > c) \tag{7d}$$

These are examples of a T-norm operator for working out the possibility, for example, that PC_1 is less than or equal to *a* AND PC_2 is less than or equal to *b*. In Fig. 5, Π stands for the use of this T-Norm.

In layer 3 the likelihood of observing the scale count equal to y_{ik} is calculated by substituting $y_{ik} = v$ into Eq. (4c). These likelihoods are then multiplied by the w_i values so that more emphasis is placed on likelihoods corresponding to the rule most likely to describe the PC_1 and PC_2 pairing. Finally, in layer 4 these weighted likelihoods are added up to give the weighted average likelihood. Conjugate gradient methods are then used to optimise the values for the parameters *a*, *b*, ρ_1 and ρ_2, μ_i and θ_i (as well as the value for Q). This non-linear optimisation algorithm chooses values for all these parameters so as to maximise this weighted average likelihood.

The outputs of this model are predicted mean scale counts for each sub group of hot mill operating conditions, together with a predicted frequency of occurrence for scale counts in each sub group. The above structure could be generalised to account for all of the principal components created from all the hot mill variables. In this paper the resulting ANFIS with principal components, or ANFISPC for short, is built up sequentially. First a principal component analysis is carried out using all p process variables at the hot mill. Second, all those principal components that have eigenvalues in excess of unity are constructed using the loadings from the Eigen vectors associated with these Eigen values. Next the principal component with the highest Eigen value, PC_1, is used to split the data into two parts with the value for *a* in the optimised ANFISPC determining this split point. Optimising this simple ANFISPC also involves estimating values for μ_1, μ_2, θ_1 and θ_2 that maximise the weighted average likelihood. Fourthly, the principal component with the next highest Eigen value, PC_2, is added to the above ANFISPC so that the data is split into 4 parts using PC_1 then PC_2. The weighted average likelihood for this larger ANFISPC is then maximised. Additional principal components are added to the ANFISPC until there is little reduction in this weighted average

likelihood. It should be pointed out that this may not be the optimal search procedure in that it may be best to split first by a principal component other than PC_1 – say PC_5. Different search procedure could form an interesting topic for future research.

6. Results and discussion

Table 2 shows the results of applying the spectral decomposition method to the twenty two process variables measured at the hot mill. It can be seen that there are six principal components that have Eigen values noticeably above 1. These components in turn account for just under 60% of the total variation in all the process variables. Table 3 shows the loadings associated with these six principal components. By looking at the absolute values of these loadings, it appears that the variation in the various temperatures at the hot mill are picked up by the first principal component (PC_1), whilst the variation in the different chemistries seems to be picked up by the second principal component, PC_2. The variation in the minimum rougher mill temperature is not being picked up at all by these six components. Using Cattells rule, these six components can be used in the ANFISPC instead of the original twenty two process variables.

PC_1 is now used to split the scale count data set into two sub components. The ANFISPC estimates that the best split occurs above and below $PC_1 = a = -0.098$. The estimate for ρ_1 in Eq. (6a) is 5 and the resulting membership function is shown in Fig. 6. The membership function is relatively step around a so that the value for the membership function ranges from 0.1 to 0.9 over a very limited range of values for PC_1. Thus PC_1 values less than -0.52 have memberships functions in excess of 0.9 and so can strongly be considered to belong to the subset $PC_1 \leq -0.098$. W a fairly step membership function the range of PC_1 values that could belong to either subset ($PC_1 > -0.098$ as well) is quite narrow. The degree of fuzzyfication is therefore quite limited.

The ANFISPC estimates that the mean scale counts for each sub group defined above, are $\mu_1 = 161$ and $\mu_2 = 354$. Thus the model predicts that if the hot mill is operated under those conditions for which $PC_1 \leq -0.098$, the average scale count on all coils produced under such conditions will be 161. Then if the hot mill is operated under those conditions for which $PC_1 > -0.098$, the average scale count on all coils produced under such conditions will be much higher at 354. This however says very little about the spread to be expected about these averages. A feel for this can be obtained by looking at the ANFISPC estimates of θ_i. These estimates are $\theta_1 = 1.59$ and $\theta_2 = 0.43$ and with $Q = 1.1$. These values, together with Eq. (4d), suggest that in both sub groups the variance in the scale count is more than the mean, but that this process variability is much greater in the sub group corresponding to $PC_1 > -0.098$. The model therefore predicts that those hot mill conditions leading to a higher mean scale count also leads to greater variability about this mean count.

In Figs. 7 the actual proportion of coils having scale counts of various values is plotted together with the predicted proportion from the ANFISPC model using the negative binomial distribution for each sub group. It can be seen from Fig.7a that when the hot mill is run under conditions corresponding to $PC_1 \leq -0.098$, just over 75% of all coils actually manufactured had a bottom scale count of 100 or less with over 85% of such coils having a scale count of 200 or less. The operating conditions corresponding to $PC_1 > -0.098$ (Fig.7b) are much less favourable with only 42% of all coils manufactured having a bottom scale

count of 100 or less and with only 54% of coils having a scale count of 200 or less. The ANFISPC model predicts these proportions reasonably well in each operating sub group.

In terms of actual process conditions, recall that in Table 3 the size of the loadings in PC_1 were such that this component was mainly determined by the temperature variables and hardly at all by the alloying elements. Using the loading values in Table 3 for the first principal component it is straight forward to calculate what a particular hot mill temperature must be to ensure $PC_1 \leq -0.098$ under various hot mill running scenarios. As an illustration this paper considers two such scenarios but many others can easily be considered. The first scenario assumes the hot mill is running at the average values for all the process variables shown in Table 1, except the one temperature corresponding to each column heading in Table 4. Table 4 therefore shows that to ensure $PC_1 \leq -0.098$, the maximum rougher mill temperature must be no higher than 1117^0C when all other process variables are at their average values. As another example, under scenario 1, the average finishing mill temperature must be no higher than 880^0C to ensure that $PC_1 \leq -0.098$.

These types of scenario may be an unrealistic representation of actual hot mill practice. Often the hot mill accepts slabs of a given chemistry and then tries to control scale through temperature variations. Scenario 2 illustrates how this model copes with this type of situation. Consider a slab entering the hot mill with the chemistry shown in the first row of Table 5. The second scenario assumes the hot mill is running at the average values for all the temperature variables shown in Table 1, except the one temperature corresponding to each column heading in Table 4 when this slab chemistry enters the hot mill. Table 4 now shows that to ensure $PC_1 \leq -0.098$, the maximum rougher mill temperature must be no higher than 1111^0C (at this chemistry and when all the other temperatures are at their average values). As another example, under scenario 2, the average finishing mil temperature must be no higher than 830^0C to ensure that $PC_1 \leq -0.098$.

To find out what the best alloying contents are, the second principal component needs to be added to the ANFISPC model, because as in Table 3 show the size of the loading in PC_2 were such that this component was mainly determined by the alloying variables and hardly at all by the temperature variables. The second principal component was therefore added to this simple ANFISPC. The scale count data is now therefore split into four sub sections using cut off values for PC_1 and PC_2. The estimated value for *a* in Eq. (6a) is unchanged at -0.098. Thus the scale count data set described above was first split into two sub sections using $PC_1 = -0.098$ as the split point. The estimated value for *b* in Eq. (6b) was 0.805. Thus each of these two sub groups are split further into another two sub groups using $PC_2 = 0.805$ as the split point. The estimate for ρ_2 in Eq. (6b) was the same as that for ρ_1, i.e. 5.

The ANFISPC estimates that the mean scale count for each of these four sub groups was $\mu_1 = 139$, $\mu_2 = 220$, $\mu_3 = 350$ and $\mu_4 = 365$. Thus the model predicts that if the hot mill is operated under those conditions for which $PC_1 \leq -0.098$ and $PC_2 \leq 0.805$, the average scale count on all coils produced under such conditions will be at its lowest - 139. This is a big improvement on the average scale count associated with the best sub group of the previous simpler ANFISPC. Then if the hot mill is operated under those conditions for which $PC_1 \leq -0.098$ and $PC_2 > 0.805$, the average scale count on all coils produced under these conditions will be much higher at 220. If the hot mill is operated under those conditions for which $PC_1 > -0.098$ and $PC_2 \leq 0.805$, the average scale count on all coils produced under these conditions will be higher again at 350. Finally, if the hot mill is operated under those conditions for

which $PC_1 > -0.098$ and $PC_2 > 0.805$, the average scale count on all coils produced under these conditions will at its highest of 365.

This says very little about the spread to be expected about these averages. A feel for this can be obtained by looking at the ANFISPC estimates of θ_i. These estimates are $\theta_1 = 3.43$, $\theta_2 = 0.81$, $\theta_3 = 0.54$ and $\theta_4 = 0.36$ and with Q = 1.04. These values, together with Eq. (4d), suggest that in all sub groups the variance in the scale count is more than the mean, but that this process variability is much greater in the sub group corresponding to higher mean scale counts.

In Figs. 8 the actual proportion of coils having scale counts of various values is plotted together with the predicted proportion from the ANFISPC model using the negative binomial distribution for each sub group. It can be seen from Fig.8a that when the hot mill is run under conditions corresponding to $PC_1 \leq -0.098$ and $PC_2 \leq 0.805$, nearly 81% of all coils actually manufactured had a bottom scale count of 100 or less with over 90% of such coils having a scale count of 200 or less. The corresponding percentages in the best sub group from the previous simplified model were around 75% and 85% respectively, so this newly identified set of operating conditions associated with new sub group produces coils with consistently lower scale counts than the conditions associated with the best sub group of the simpler ANFISPC above. The least favourable operating conditions corresponding to $PC_1 > -0.098$ and $PC_2 > 0.805$ where the mean scale count is 365 and where only 46% of all coils manufactured have a bottom scale count of 100 or less and where only 55% of such coils have a scale count of 200 or less. The ANFISPC model predicts these proportions reasonably well in each operating sub group. In the same way as above, the actual chemistries and temperatures leading to hot mill operations where $PC_1 \leq -0.098$ and $PC_2 \leq 0.805$ is readily found.

Next the third principal component is added to the ANFISPC so that the data set is then split up into eight sub groups. The estimated value for *a* and *b* in Eqs. (6a,6b) are unchanged at -0.098 and 0.805 respectively. Thus the scale count data set described above was first split into two sub sections using $PC_1 = -0.098$ as the split point and then into four sub sections using $PC_2 = 0.805$ as the split point. The estimated value for *c* in

$$\gamma_{PC3 \leq c} = \frac{1}{1 + \exp[\rho_3 (PC_3 - c)]}$$

was -0.482. Thus each of the four sub groups just described were split into two further sub groups using $PC_3 = -0.482$ as the split point. The estimate for ρ_3 in in the equation above was the same as that for ρ_1, i.e. 5.

The resulting ANFISPC estimated that the mean scale count for each of these eight sub groups was $\mu_1 = 146$, $\mu_2 = 133$, $\mu_3 = 306$, $\mu_4 = 185$, $\mu_5 = 331$, $\mu_6 = 362$, $\mu_7 = 464$ and $\mu_8 = 289$. Thus the model predicts that if the hot mill is operated under those conditions for which $PC_1 <= -0.098$ and $PC_2 <= 0.8058$ and $PC_3 > -0.482$, the average scale count on all coils produced under these conditions will be at its lowest at 133. This is only a small improvement on the average scale count associated with the best sub group of the previous simpler ANFISPC (where the average scale count was 139). This suggests there would be little to gain from expanding this ANFISPC through the introduction of the 4th principal component and the construction of 16 sub group operating conditions.

If the hot mill is operated under those conditions for which $PC_1 > -0.098$ and $PC_2 > 0.8058$ and $PC_3 <= -0.482$, the average scale count on all coils produced under these conditions will be at its highest of 464. These estimates for θ_i were $\theta_1 = 1.99$, $\theta_2 = 6.91$, $\theta_3 = 0.34$, $\theta_4 = 1.65$, $\theta_5 = 0.48$, $\theta_6 = 0.57$, $\theta_7 = 0.24$ and $\theta_8 = 0.60$ and with $Q = 1.06$. These values, together with Eq. (4d), suggest that in all sub groups the variance in the scale count is more than the mean, but that this process variability is much greater in the sub group corresponding to higher mean scale counts. In fact in the sub group with the lowest mean, the variance is substantially lower than in any other group so that operating conditions associated with this sub group produces coils that consistently have the lowers scale counts.

In Figs. 9 the actual proportion of coils having scale counts of various values is plotted together with the predicted proportion from the ANFISPC model using the negative binomial distribution for each sub group. It can be seen from Fig.9b that when the hot mill is run under conditions corresponding to $PC_1 <= -0.098$ and $PC_2 <= 0.8058$ and $PC_3 > -0.482$, nearly 83% of all coils actually manufactured had a bottom scale count of 100 or less with over 90% of such coils having a scale count of 200 or less. The least favourable operating conditions corresponding to $PC_1 > -0.098$ and $PC_2 > 0.8058$ and $PC_3 <= -0.482$ where the mean scale count was 464 and where only 35% of all coils manufactured had a bottom scale count of 100 or less and where only 44% of such coils had a scale count of 200 or less (Fig. 9g). The ANFISPC model predicts these proportions reasonably well in each operating sub group.

In the same way as above, the actual chemistries and temperatures leading to hot mill operations where $PC_1 <= -0.098$ and $PC_2 <= 0.8058$ and $PC_3 > -0.482$ is readily found. In Fig.10a, the three principal components are plotted against the average rougher mill temperature. The values for the principal components are obtained using the average values for all the process variables together with the average rougher mill temperatures shown on the horizontal axis of this figure. It is clear from this graph that all the average rougher mill temperatures lead to a PC_2 value below 0.8058. However, an average rougher mill temperature below about 1140^{o}C is needed to ensure that $PC_3 > -0.482$. Further, an average rougher mill temperature below about 1100^{o}C is need to ensure that $PC_1 <= -0.098$. Thus to achieve a consistently low scale count, the average rougher mill temperature must be kept below 1100^{o}C when the hot mill is running at the average values for all the other process variables.

In the Fig.10b, the three principal components are again plotted against the average rougher mill temperature but this time when the incoming slab had the chemistry shown in the last row of Table 5 and when all the other temperatures are set at their average values. It is clear from this figure that all the average rougher mill temperatures lead to a PC_2 value below 0.8058. However, an average rougher mill temperature below about 1095^{o}C is needed to ensure that $PC_1 <= -0.098$. Further, an average rougher mill temperature below about 1030^{o}C is need to ensure that $PC_3 > -0.482$. Thus to achieve a consistently low scale count the average rougher mill temperature must be kept below 1030^{o}C when the hot mill is running at the average values for all the other temperatures and when the slab chemistry is as shown above.

7. Conclusions

The scale count data from the hot mill contained a number of complexities that required the novel amalgamation of a number of well know data mining techniques. This was achieved in this paper by combining the analytical techniques of principal components, regression trees, fuzzy logic and neural networks into a single framework for estimation purposes. The resulting generalisation of the ANFIS has the potential to further improve past studies undertaken on count data. The practical implications of this addition to the state of the art in data mining techniques was the ability to accurately identify for the first time the operating conditions leading to consistently low scale counts at the Port Talbot hot mill.

More precisely, it was found that the variability present in the twenty two hot mill variables could be captured adequately using just six principal components and that an accurate ANFISPC could be obtained using just the first three of these. This model identified eight unique hot mill operating conditions, with the condition $PC_1 \leq -0.098$ and $PC_2 \leq -0.098$ and $PC_3 > -0.4828$ leading to the smallest mean scale count (133), the smallest variability in scale count about this mean (a standard deviation of 51) and the greatest proportion of coils with scale counts below 100 (83%). A few scenarios were looked at to give a feel for the hot mill temperatures required to achieve these range of PC values. For example, the average rougher mill temperature must be kept below 1030^{o}C when the hot mill is running at the average values for all the other temperatures and when the slab chemistry is as shown in the last row of Table 5.

One of the short comings of this paper, is the search rule used for introducing each principal component to the ANFISPC. It may be the case that the first split in the data that best predicts the scale count may not be the one with the highest Eigen value. So one area for future research would be to develop a more comprehensive algorithm for choosing how the create each sub division of the data. Another fruitful area for future research would be to used components from partial least squares as inputs to the ANFIS. This might actually simplify the search algorithm, because the first component from partial least squares is the one that picks up the most variation in the scale count.

References

Bolt, H., (June 2003). Understanding the properties of oxide scale on hot rolled strip. In: METEC Congress 03 / 3rd European Rolling Conference (ERC), Dusseldorf.

Yang, Y., Yang, C., Lin, S., Chen, C., Tsai, W., (2008). Effects of Si and its content on the scale formation on hot-rolled steel strips, *Materials Chemistry and Physics,* 112: 566-571.

Taniguchi, S., Yamamoto, K., Megumi, D., Shibata, T., (2001). Characteristics of scale/substrate interface area of Si-containing low-carbon steels at high temperatures, *Materials Science and Engineering A*, 308:250-257.

Munther, P., & Lenard, J., (1999). The effect of scaling on interfacial friction in hot rolling of steels, *Journal of Materials Processing Technology,* 88:105-113.

Sun W., Tieu, A. K., Jiang, Z., Zhu, H., Lu, C., (2004). Oxide scales growth of low-carbon steel at high temperatures, *Journal of Materials Processing Technology,* 155-156: 1300-1306.

Bolt, H., (2000). The properties of oxide scales on hot rolled steels: A literature review, Corus Research, Development & Technology, Ijmuiden Technology Centre.

Tsoukalas, L. H., & Uhrig, R. E., (1997). *Fuzzy and Neural Approaches in Engineering*. John Wiley & Sons Inc, New York, Chapter 2.

Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J., (1984). *Classification and Regression Trees*. Wadsworth International Group, Belmont.

Friedman, J. H., (1991). Multivariate Adaptive Regression Splines, *Annals of Statistics*, 19 (1):1–67.

Jolliffe, I. T., (2002). *Principal Component Analysis*. Springer-Verlag, New York, 2nd Edition.

Burt, C. L., (1945). *How the mind works*. Allen & Unwin, London.

Childs, D., (1970). *The Essentials of Factor Analysis*. Holt, Rinehart and Winston, London.

Mardia, K. V., Kent, J. T., Bibby, M., (1979). *Multivariate Analysis. Academic Press*. London.

Cattell, R. B., (1952). *Factor Analysis*. Harper, New York.

Cameron, C., Trevedi, P., (1998). *Regression Analysis of Count Data*. Cambridge University Press, New York.

Jang, J. S. R., Sun, C. T., Mizutani, E., (1997). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice Hall, Upper Saddle River.

Jang, J. S. R., (1993). ANFIS: Adaptive-Network-based Fuzzy Inference Systems, *IEEE Transactions on Systems, Man, and Cybernetics*, 23(03):665-685.

Takagi, T., Sugeno, M., (1995). *Fuzzy Identification of Systems and its Applications to Modelling and Control*. In: IEEE Transactions on Systems, Man and Cybernetics 15:116-132.

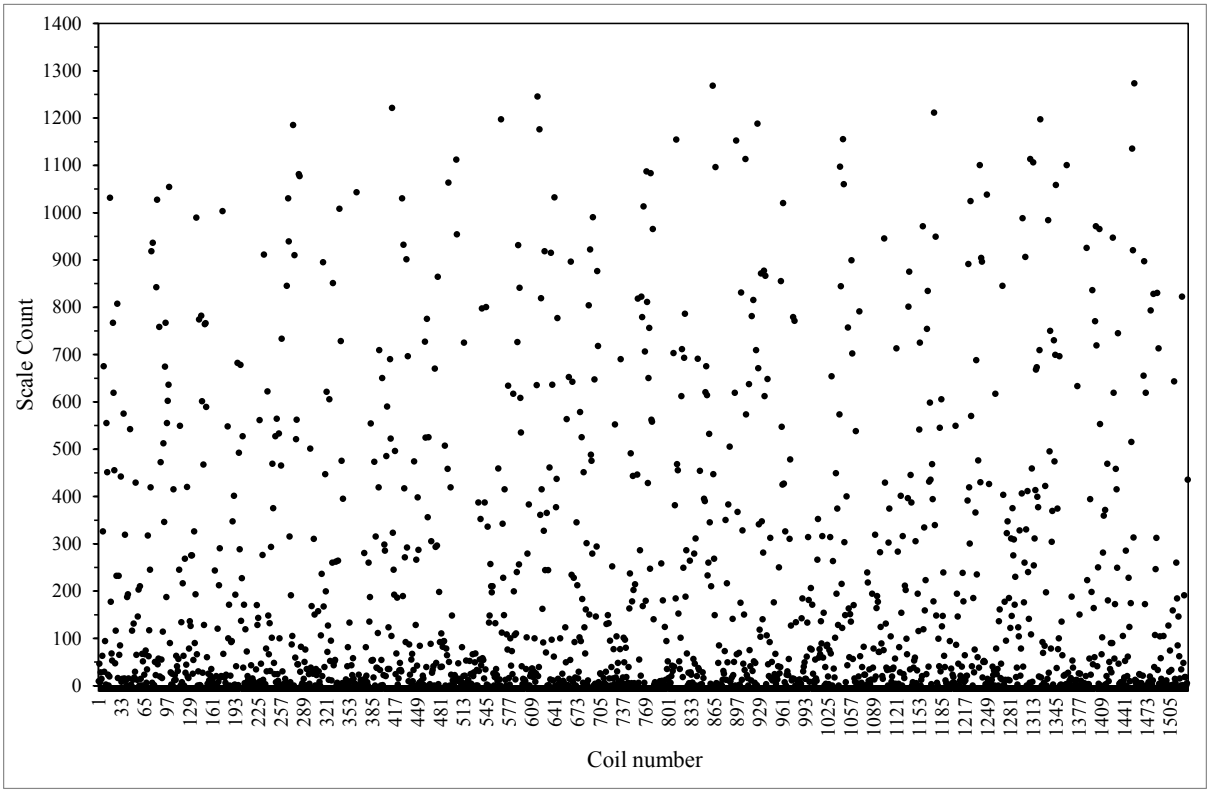


Fig. 1 The scale count data collected at the Port Talbot hot mill over 3 months under different processing conditions.

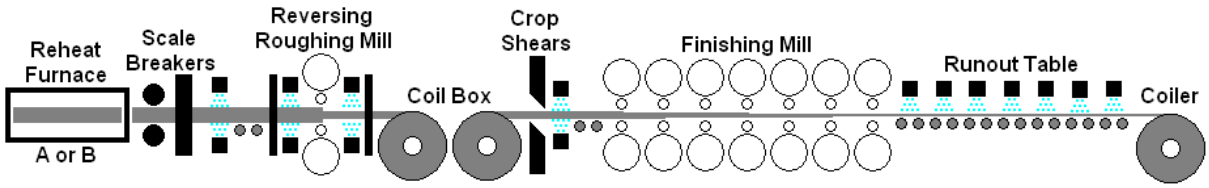


Fig. 2 Port Talbot hot mill layout.

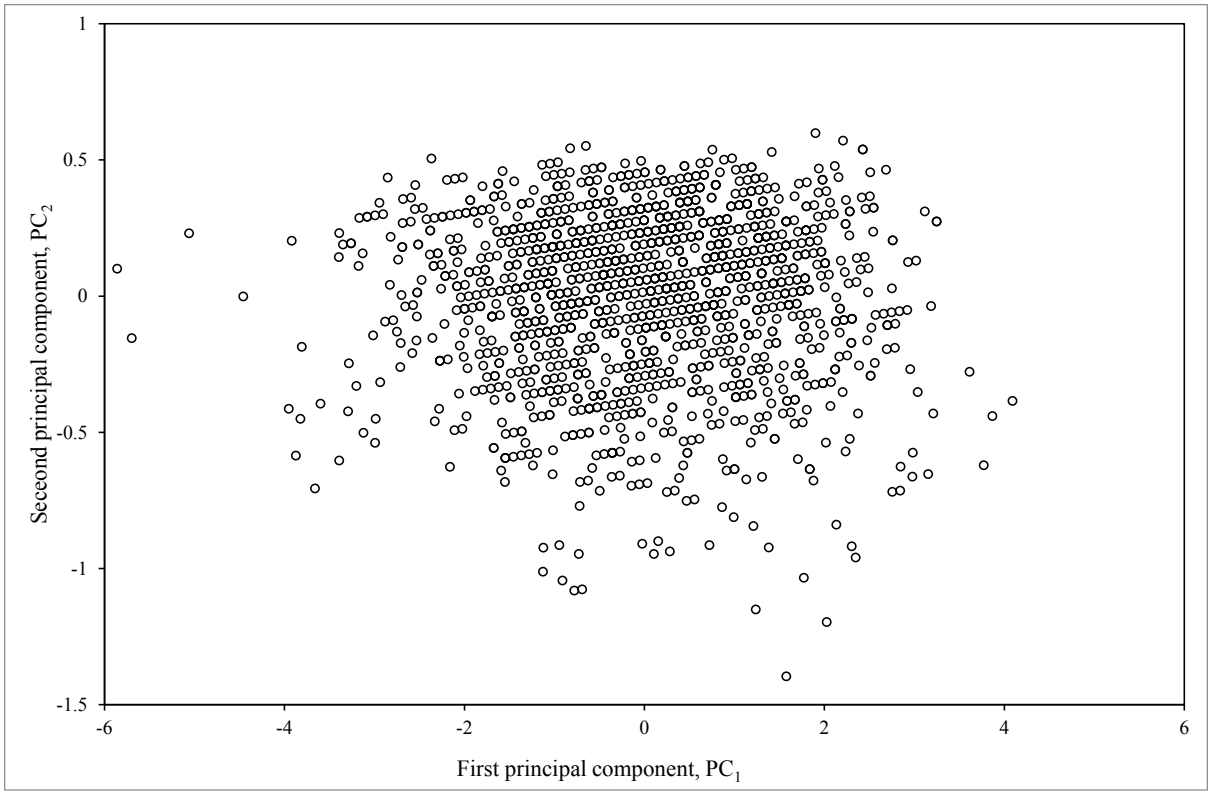


Fig. 3a A cross plot of PC_1 and PC_2 as calculated using. Eqs. (2c,d).

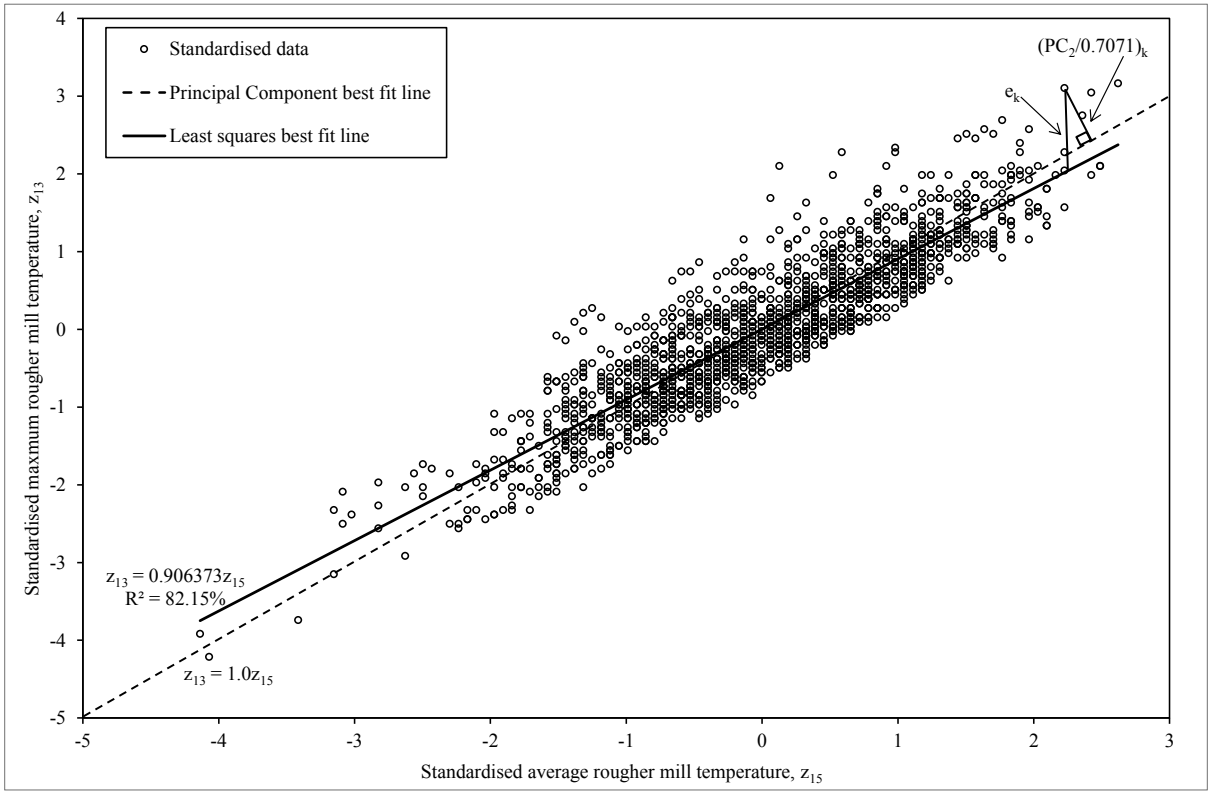


Fig. 3b A best fit line interpretation of principal components.

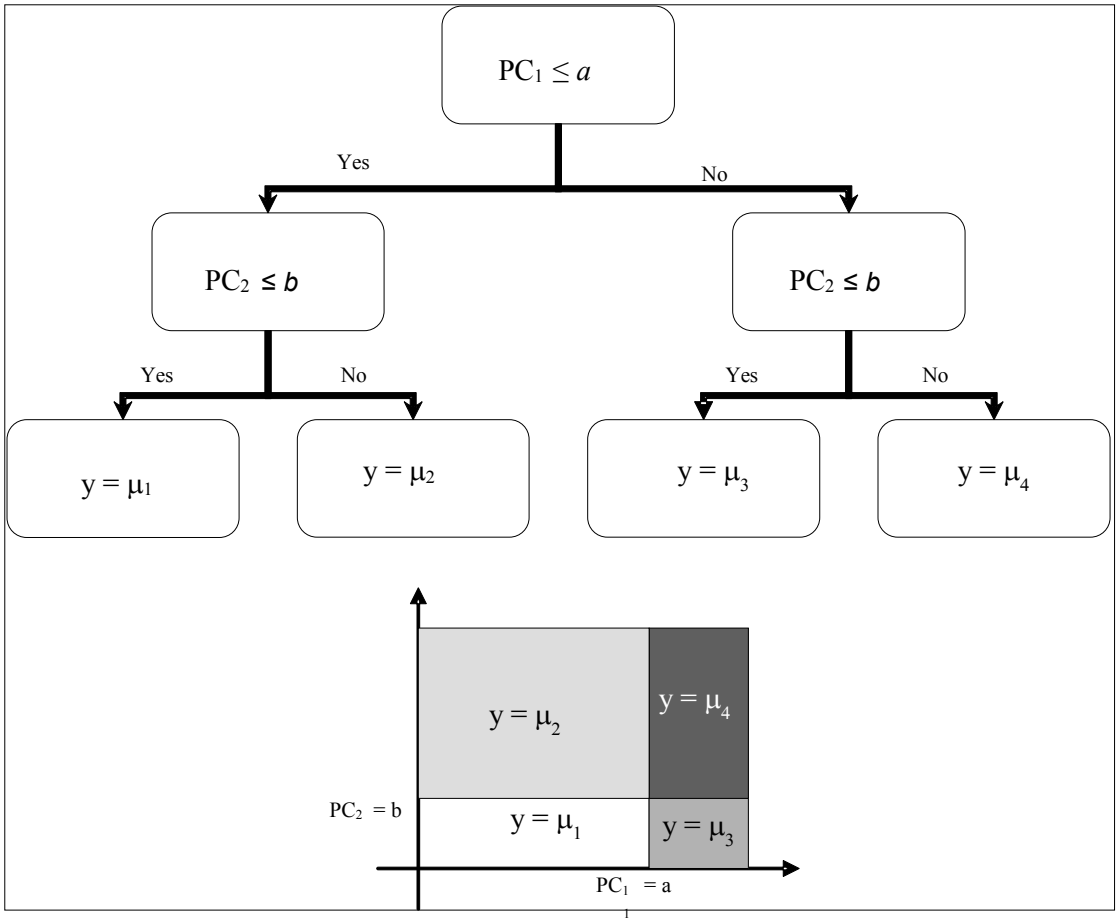


Fig. 4 A typical binary regression tree with two principal components (PC_1 and PC_2) and one output y.

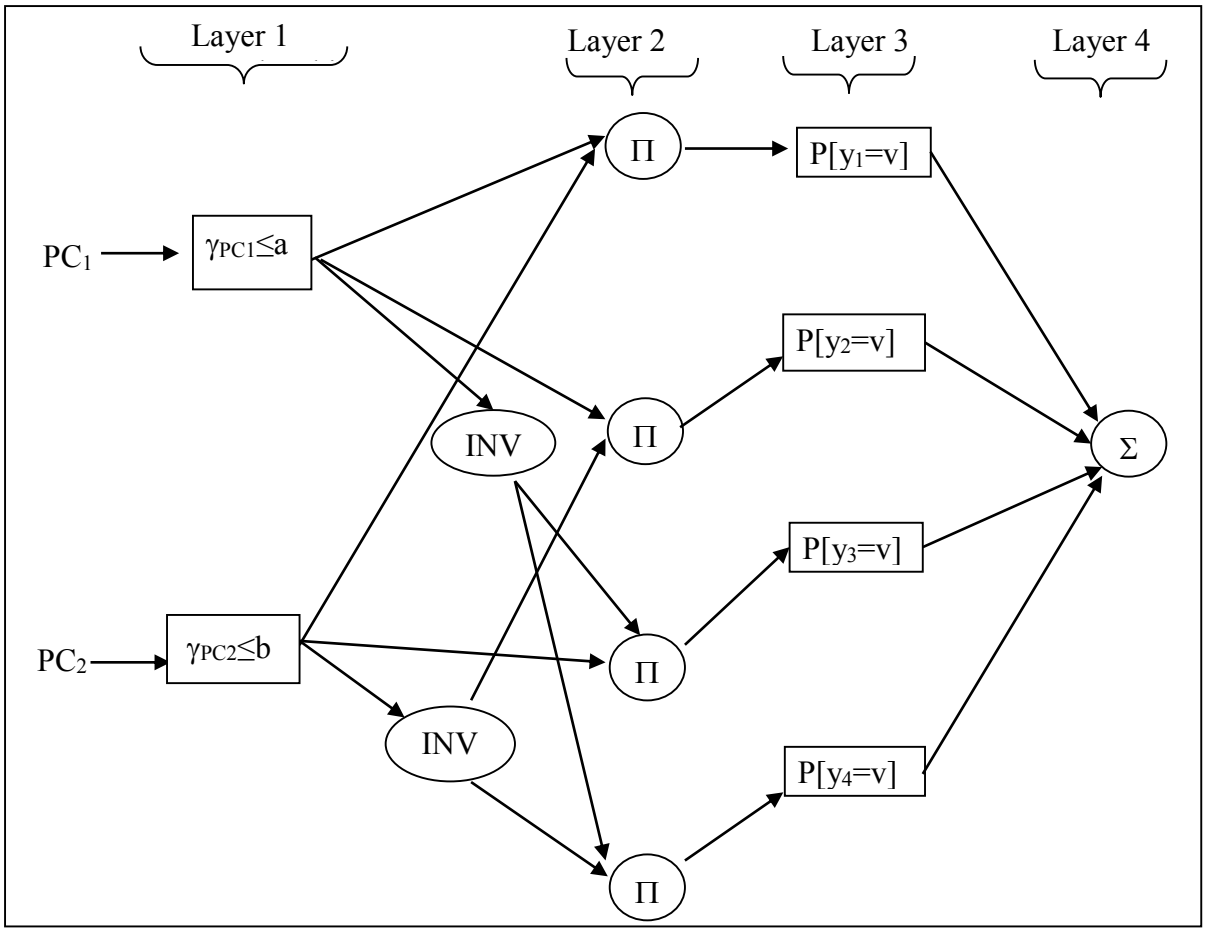


Fig. 5. ANFIS architecture corresponding to the representation shown in Fig. 4.

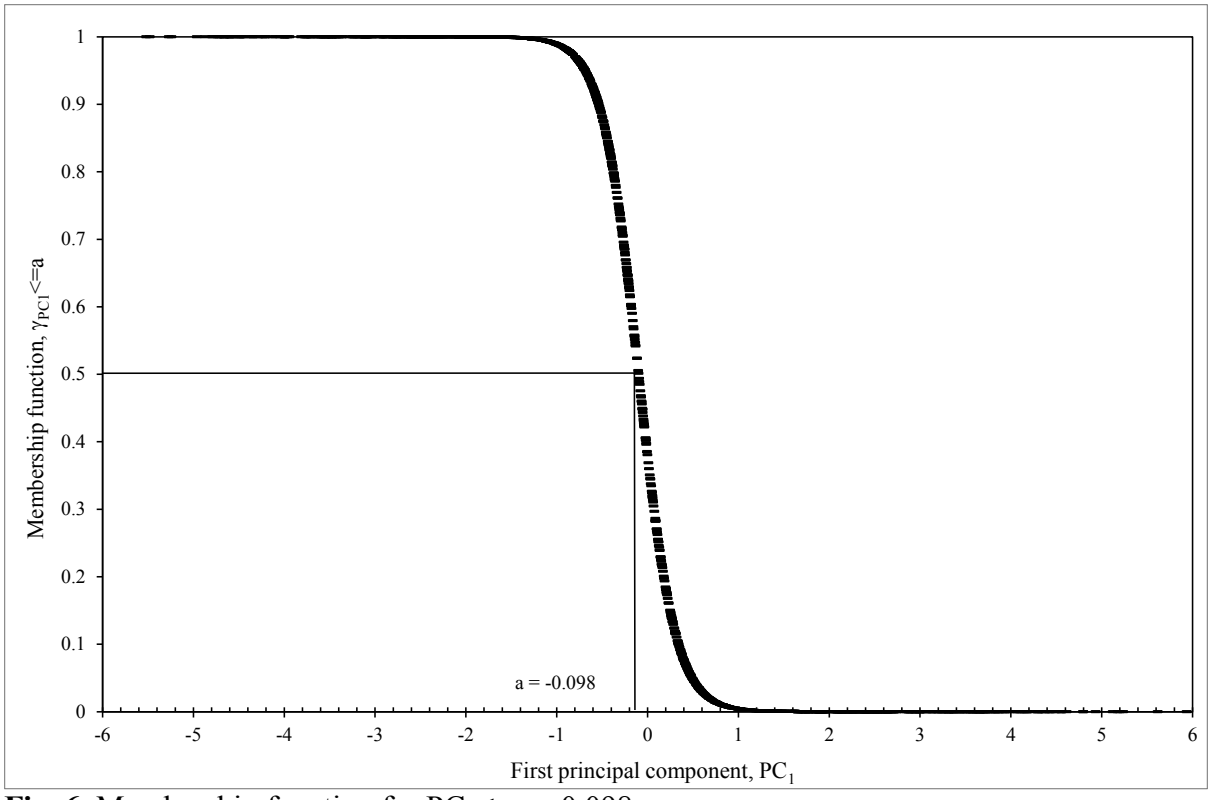
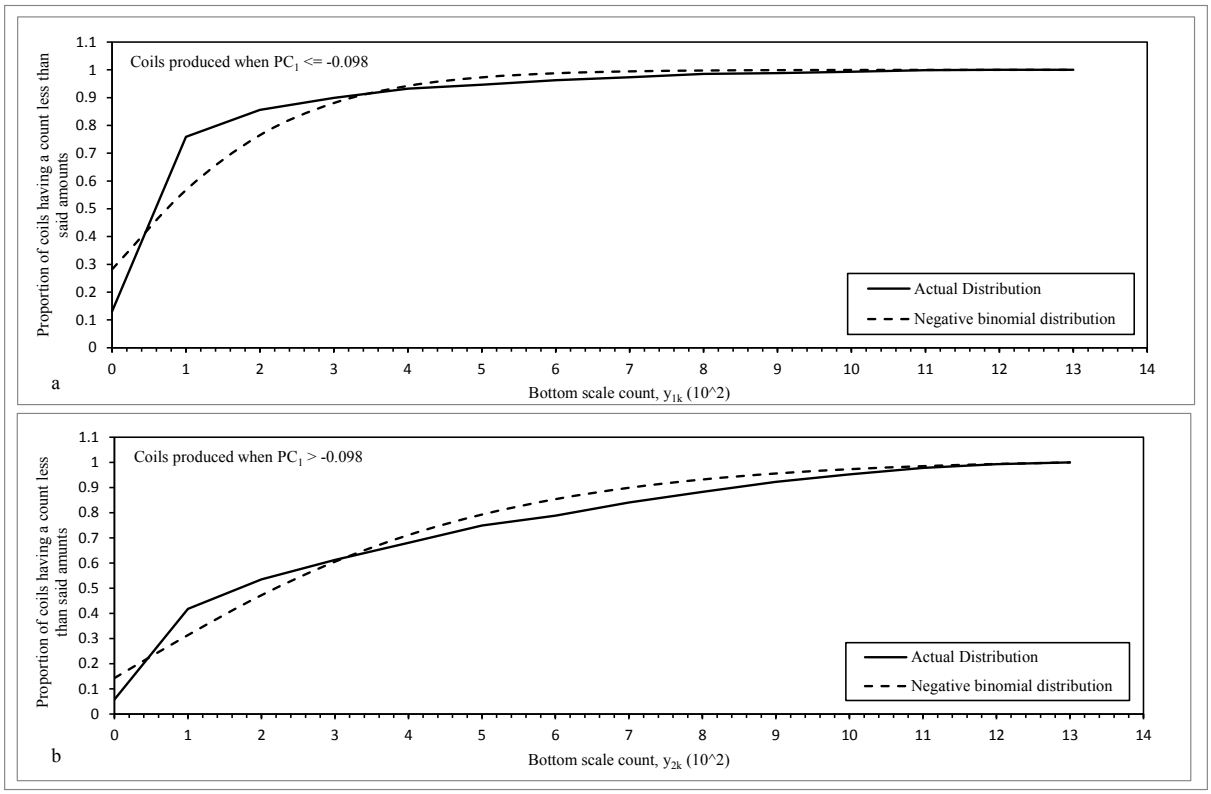
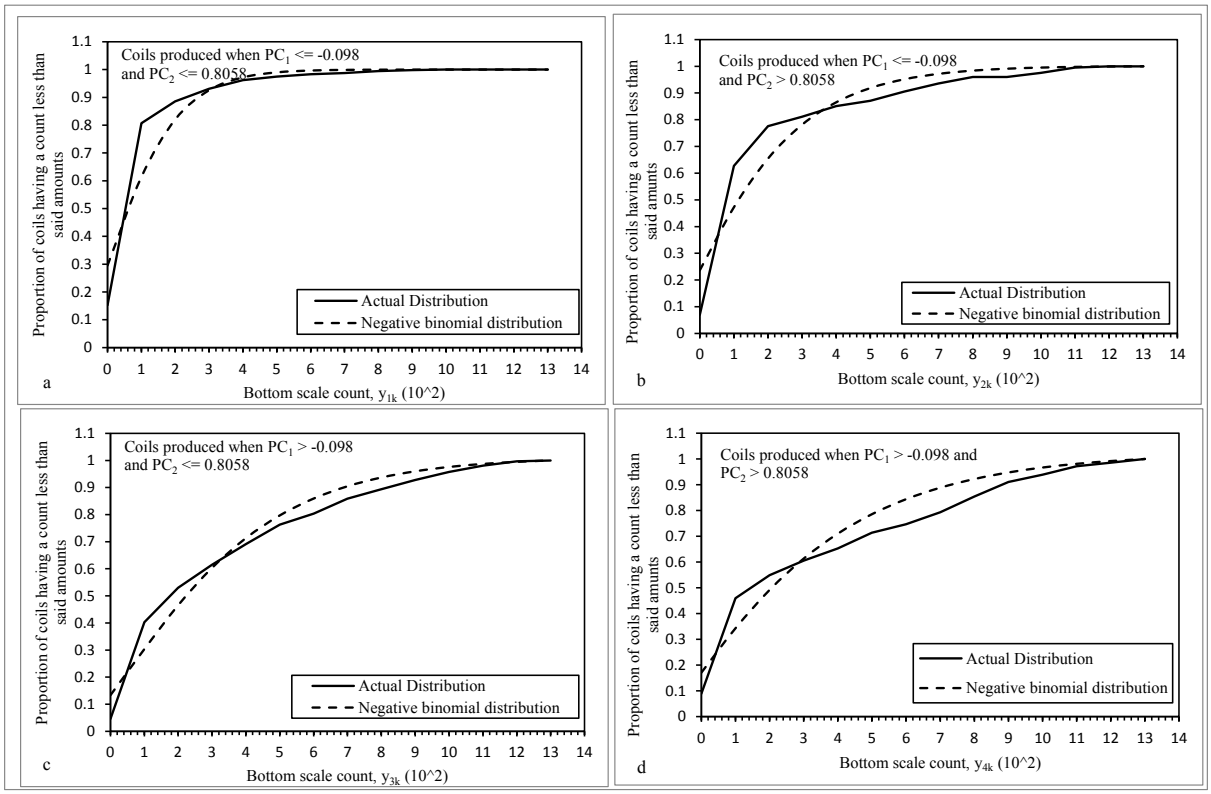


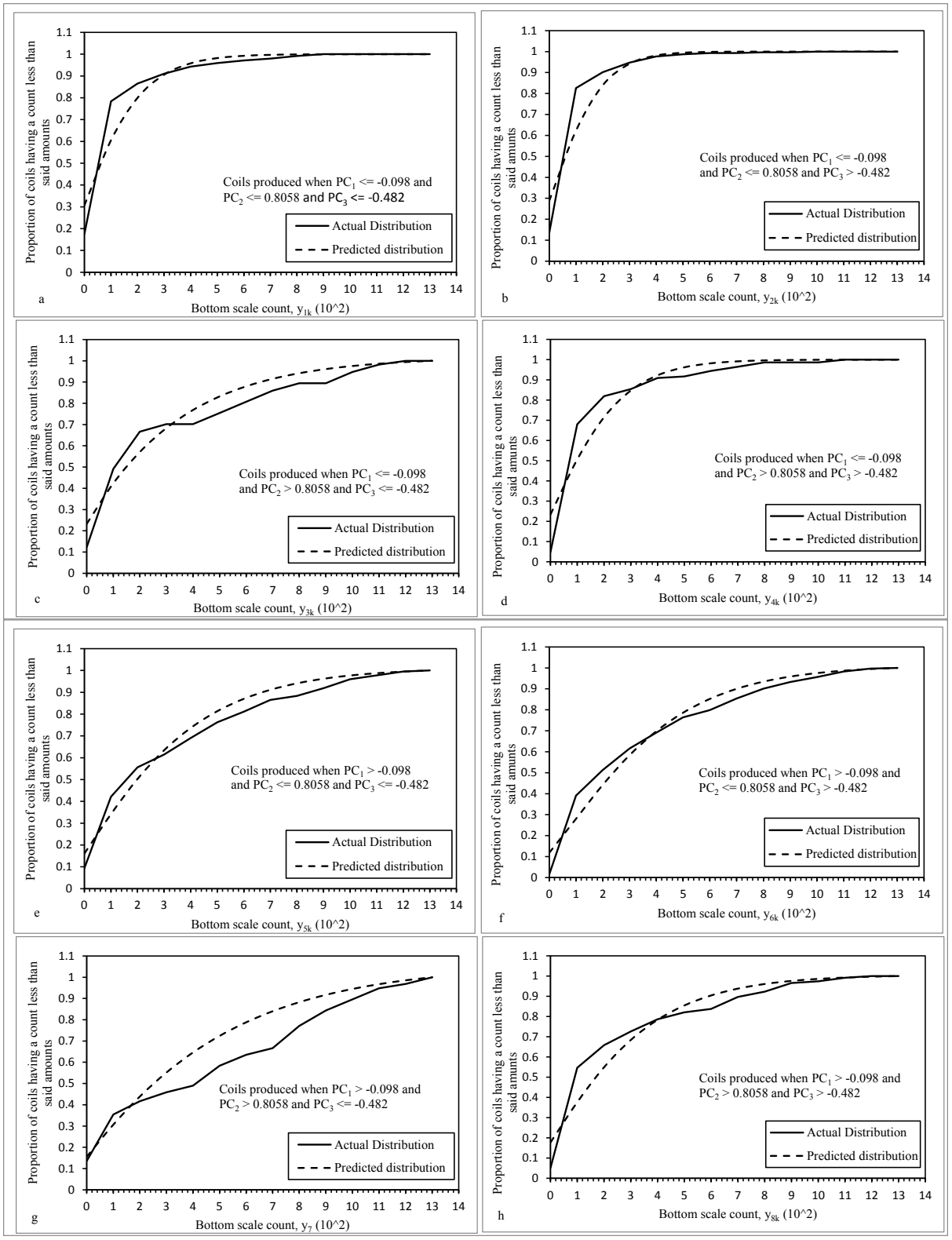
Fig. 6 Membership function for $PC_1 \leq a = -0.098$.



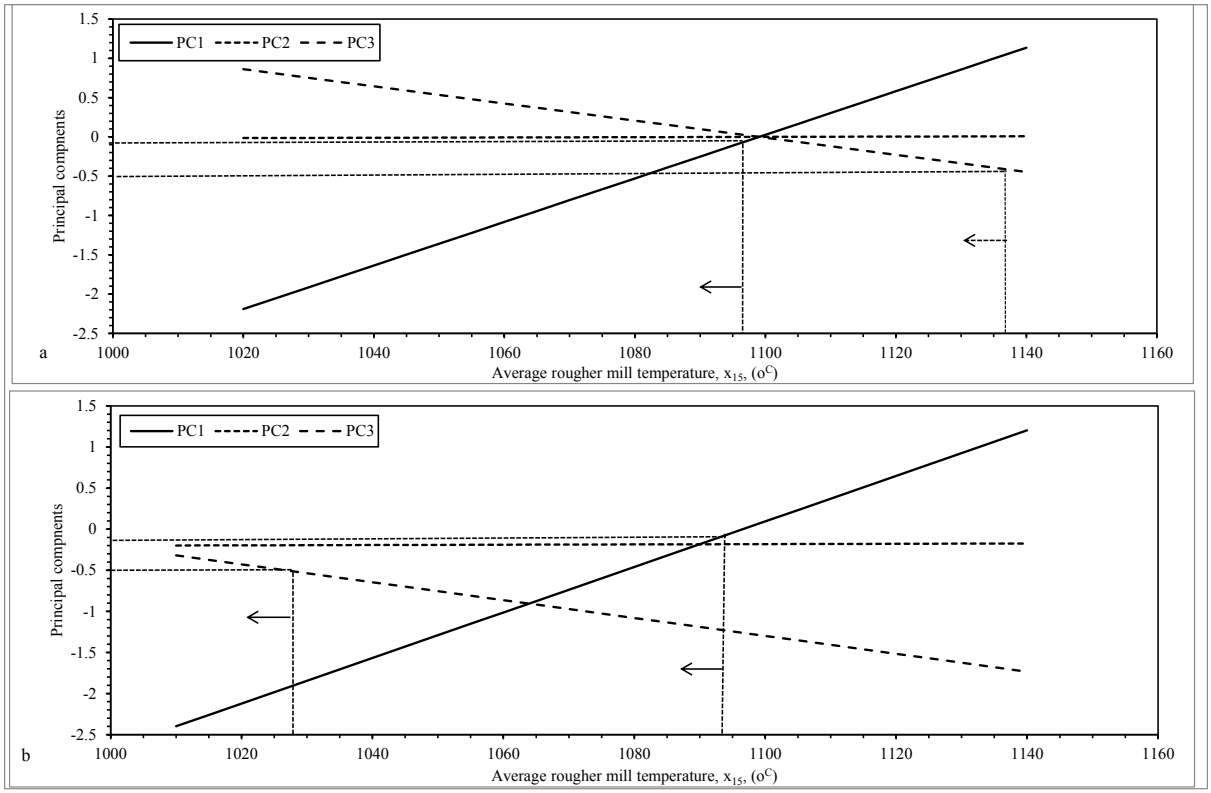
Figs. 7 The actual and negative binomial distribution for coils produced when a. $PC_1 \leq -0.098$ and b. $PC_1 > -0.098$.



Figs. 8 The actual and negative binomial distribution for coils produced when a. $PC_1 \leq -0.098$ with $PC_2 \leq -0.098$, b. $PC_1 \leq -0.098$ with $PC_2 > -0.098$, c. $PC_1 > -0.098$ with $PC_2 \leq -0.098$, d. $PC_1 > -0.098$ with $PC_2 > -0.098$.



Figs. 9 The actual and negative binomial distribution for coils produced when a. $PC_1 \leq -0.098$ with $PC_2 \leq -0.098$ and $PC_3 \leq -0.4828$, b. $PC_1 \leq -0.098$ with $PC_2 \leq -0.098$ and $PC_3 > -0.4828$, c. $PC_1 \leq -0.098$ with $PC_2 > -0.098$ and $PC_3 \leq -0.4828$, d. $PC_1 \leq -0.098$ with $PC_2 > -0.098$ and $PC_3 > -0.4828$, e. $PC_1 > -0.098$ with $PC_2 \leq -0.098$ and $PC_3 \leq -0.4828$, f. $PC_1 > -0.098$ with $PC_2 \leq -0.098$ and $PC_3 > -0.4828$, g. $PC_1 > -0.098$ with $PC_2 > -0.098$ and $PC_3 \leq -0.4828$, h. $PC_1 > -0.098$ with $PC_2 > -0.098$ and $PC_3 > -0.4828$.



Figs. 10 Variation in the three principal components with the average rougher mill temperature.

Table 1 Process variables and their sample means and standard deviations

Process Variable	x_j	Mean	Standard Deviation
% Carbon	x_1	0.069	0.0042
% Manganese	x_2	0.485	0.0226
% Phosphorus	x_3	0.014	0.0031
% Sulphur	x_4	0.014	0.0030
% Silicon	x_5	0.005	0.0032
% Copper	x_6	0.016	0.0059
% Nickel	x_7	0.010	0.0030
%Chrome	x_8	0.018	0.0040
% Aluminium	x_9	0.034	0.0047
%Tin	x_{10}	0.003	0.0018
% Nitrogen	x_{11}	0.012	0.0009
% solAl	x_{12}	0.031	0.0043
Maximum RM Temperature, ^{0}C	x_{13}	1121.75	17.87
Minimum RM Temperature, ^{0}C	x_{14}	918.34	11.97
Average RM Temperature, ^{0}C	x_{15}	1099.43	16.18
Maximum CS Temperature, ^{0}C	x_{16}	1099.88	19.89
Minimum CS Temperature, ^{0}C	x_{17}	951.69	48.59
Average CS Temperature, ^{0}C	x_{18}	1069.55	20.80
Maximum FM Temperature, ^{0}C	x_{19}	907.16	7.18
Minimum FM Temperature, ^{0}C	x_{20}	829.87	28.96
Average FM Temperature, ^{0}C	x_{21}	881.05	5.75
ROT Temperature, ^{0}C	x_{22}	731.35	25.13
Scale Count	y	203.72	286.78

RM = Rougher Mill, CS = Crop Sheer, FM = Finishing Mill and ROT = Run out Table.

Table 2 Principal component analysis for the hotmill process variables

Components	Eigenvalues	Variation, %	Cumulative Variation, %
PC_1	4.141	18.82	18.82
PC_2	2.445	11.11	29.93
PC_3	1.896	8.62	38.55
PC_4	1.366	6.21	44.76
PC_5	1.287	5.85	50.61
PC_6	1.259	5.72	56.33
PC_7	1.032	4.69	61.02
PC_8	1.029	4.68	65.7
PC_9	0.981	4.46	70.16
PC_{10}	0.926	4.21	74.37
PC_{11}	0.874	3.97	78.34
PC_{12}	0.853	3.88	82.22
PC_{13}	0.716	3.25	85.47
PC_{14}	0.671	3.05	88.52
PC_{14}	0.65	2.95	91.47
PC_{16}	0.444	2.02	93.49
PC_{17}	0.444	2.02	95.51
PC_{18}	0.355	1.61	97.12
PC_{19}	0.294	1.34	98.46
PC_{20}	0.24	1.09	99.55
PC_{21}	0.094	0.43	99.98
PC_{22}	0.005	0.02	100

Table 3 Loadings associated with the principal components having the six largest Eigen values

Loadings	Variable	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6
a_1	x_1	0.0069	0.0827	-0.0446	0.035	-0.0001	-0.7143
a_2	x_2	-0.035	0.2776	0.2336	-0.0691	0.4536	0.0242
a_3	x_3	-0.0379	0.2736	-0.1254	0.5209	0.2891	-0.0835
a_4	x_4	-0.0073	0.0807	0.3311	-0.0376	0.5229	0.0281
a_5	x_5	-0.0379	0.2736	-0.0263	0.3677	-0.321	0.0793
a_6	x_6	0.0162	0.1703	0.4816	0.1202	-0.3054	0.1402
a_7	x_7	-0.0399	0.1703	0.4816	0.1202	-0.3045	-0.2007
a_8	x_8	-0.0279	0.2023	-0.0385	0.4672	0.012	0.0434
a_9	x_9	-0.0768	0.5539	-0.1678	-0.3388	-0.0323	-0.002
a_{10}	x_{10}	-0.0072	-0.0458	0.3278	-0.3285	-0.1857	-0.0157
a_{11}	x_{11}	-0.0427	0.0869	-0.2851	-0.0095	-0.3335	-0.0063
a_{12}	x_{12}	-0.0768	0.5539	-0.1678	-0.3388	-0.0323	-0.002
a_{13}	x_{13}	0.4223	0.0029	-0.1659	-0.0036	0.0027	-0.0121
a_{14}	x_{14}	0.4223	0.0029	-0.1659	-0.0036	0.0027	-0.0121
a_{14}	x_{15}	0	0	0	0	0	0
a_{16}	x_{16}	0.4487	0.0485	0.0214	-0.0033	-0.0172	0.0033
a_{17}	x_{17}	0.3593	0.1263	0.0755	-0.0029	-0.003	-0.0092
a_{18}	x_{18}	0.2107	0.0274	-0.01	-0.0017	-0.0018	0
a_{19}	x_{19}	0.349	0.0778	0.0499	-0.0028	-0.0029	0.0193
a_{20}	x_{20}	0.3299	0.0738	0.1513	-0.0027	-0.0028	0.0184
a_{21}	x_{21}	0.1315	0.0171	0.1352	-0.0011	-0.0011	-0.0017
a_{22}	x_{22}	-0.0069	0.0827	-0.0446	0.035	0.0001	0.6418

Table 4 Maximum temperatures required to ensure $PC_1 \leq -0.098$ under different scenarios

	x_{13}	x_{14}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}
Scenario 1	1117	1096	1095	1064	929	905	880	808	1090
Scenario 2	1111	1090	1088	1055	892	902	830	773	1670

All shown temperatures are in degrees Celsius. Scenario 1: All process variables are set at their average values except the one temperature corresponding to each column above. Scenario 2: A slab enters the hot mill with the following chemistry:
%C:0.074, %Mn 0.492, %P:0.014, %S:0.014, %Si:0.005, %Cu:0.015, %Ni:0.008, %Cr:0.019, %Al : 0.029, %Sn:0.003, %N :0.013, and %solAl:0.027. Also the temperature variables are set at their average values except the one temperature corresponding to each column above.

Table 5. Illustrative slab chemistries - %

C	Mn	P	S	Si	Cu	N	Cr	Al	Sn	N	solAl
0.074	0.492	0.014	0.014	0.005	0.015	0.008	0.019	0.029	0.003	0.013	0.027
0.073	0.483	0.013	0.014	0.004	0.013	0.007	0.017	0.035	0.001	0.012	0.033