

Overall model of the dynamic behaviour of the steel strip in an annealing heating furnace on a hot-dip galvanizing line^(*)

F. J. Martínez-de-Pisón*, A. Pernía*, E. Jiménez-Macías** and R. Fernández*

Abstract

Predicting the temperature of the steel strip in the annealing process in a hot-dip galvanizing line (HDGL) is important to ensure the physical properties of the processed material. The development of an accurate model that is capable of predicting the temperature the strip will reach according to the furnace's variations in temperature and speed, its dimensions and the steel's chemical properties, is a requirement that is being increasingly called for by industrial plants of this nature. This paper presents a comparative study made between several types of algorithms of Data Mining and Artificial Intelligence for the design of an efficient and overall prediction model that will allow determining the strip's variation in temperature according to the physico-chemical specifications of the coils to be processed, and fluctuations in temperature and speed that are recorded within the annealing process. The ultimate goal is to find a model that is effectively applicable to coils of new types of steel or sizes that are being processed for the first time. This model renders it possible to fine-tune the control model in order to standardise the treatment in areas of the strip in which there is a transition between coils of different sizes or types of steel.

Keywords

Hot-dip galvanizing line; Data mining; Artificial intelligence; Modelling.

Modelo global del comportamiento dinámico de la banda de acero en un horno de recocido de una línea de galvanizado

Resumen

La predicción de la temperatura de la banda de acero dentro del proceso de recocido de una planta de galvanizado continuo en caliente es importante para garantizar las propiedades físicas del material procesado. El desarrollo de un modelo preciso que sea capaz de predecir la temperatura que va a alcanzar la banda según las variaciones de temperaturas y velocidades del horno, y sus dimensiones y propiedades químicas del acero, es una necesidad cada vez más demandada por este tipo de plantas industriales. En el presente estudio se muestra una comparativa realizada entre diversos tipos de algoritmos de Minería de Datos e Inteligencia Artificial para el desarrollo de un modelo de predicción eficiente y global que permita determinar la variación de temperatura de la banda según las características físico-químicas de las bobinas a procesar y las fluctuaciones de temperaturas y velocidades que aparezcan dentro del proceso de recocido. El objetivo final es la búsqueda de un modelo que sea eficiente ante bobinas con nuevos tipos de acero o dimensiones que no hayan sido procesadas anteriormente. Con este modelo es posible optimizar los modelos de control para poder conseguir homogeneizar el tratamiento en zonas de la banda donde existe la transición entre bobinas con diferentes dimensiones o tipos de acero.

Palabras clave

Línea de galvanizado continuo en caliente; Minería de datos; Inteligencia artificial; Modelado.

1. INTRODUCTION

The commissioning of new production plants, the processing of new types of products or the readjustment of the original production conditions tend to require a large amount of human effort and a lot of time and money. In these cases, having robust

models that are capable of responding correctly to the requirements not only of the products that have already been processed but also of new ones is a need that is being increasingly called for in today's industry.

Modern techniques in Data Mining (DM) and Artificial Intelligence (AI) allow designing prediction models based on historical information on the

^(*) Trabajo recibido el día 22 de Septiembre de 2009 y aceptado en su forma final el día 15 de Abril de 2010.

* EDMANS Group (<http://www.mineriadatos.com>), Department of Mechanical Engineering, Universidad de La Rioja, Spain. Tel.: + 34 941 299232; fax: + 34 941 299794, fjmartin@unirioja.es.

** IDG Group (<http://www.mineriadatos.com>), Department of Mechanical Engineering, Universidad de La Rioja, Spain.

industrial process stored in databases. The challenge lies in designing overall models that learn from the past yet which are capable of still dealing efficiently with any new operating conditions that may arise in the future.

Hot-dip galvanizing line (HDGL) plants process coils of different sizes, thicknesses and types of steel. This means that the parameters for the annealing furnaces need to be recalculated for each one of the products to be galvanized. This is the point when use is made of the control models that help to determine the best parameters for the furnace according to the physico-chemical specifications of each one of the coils to be processed.

This paper presents a comparative study of multiple DM and AI techniques and their practical application to the design of an overall dynamic model that allows predicting the temperature that the steel strip is going to reach when it leaves the heating zone of an HDGL furnace at a time $t+1$ based on the present conditions of the process (time t), the variation that is expected to be recorded in the same and the physico-chemical properties of the steel strip at that moment. The ultimate goal is to design an effective model that is capable of explaining the behaviour of the strip for different types of steels and sizes (width and thickness) in order to be used for the development of ever more efficient and effective control models.

The process of creating the model is undertaken in three stages:

- First, a database is created with the variables that have the greatest influence on the strip

heating process. This database uses historical data from the industrial process. This process involves the development of a stratified sampling that allows standardising existing cases in order to increase the degree of reliability of the models created.

- Subsequently, validation is made of a battery of different techniques arising from Data Mining (DM) and Artificial Intelligence (AI) with a view to identifying which of them generate better predictive models.
- Finally, the models created are tested with new types of steel coils to identify the degree of generalisation of the models created.

Section II in this paper describes the problem to be resolved. This is followed by Section III, which presents the stages in the development of the Data Mining process: the capture and selection of variables, the design of the model, the pre-processing of the information and the search for the best DM and AI techniques for obtaining the best regression models. Section IV presents the results obtained and, finally, Section V reports the final conclusions.

2. DESCRIPTION OF THE PROBLEM

A continuous hot-dip galvanizing line is composed of several stages (Fig. 1). The initial material is the steel coil from the cold-rolling with the required thickness. The steel is unwound and run through a series of vertical loops within the furnace. The

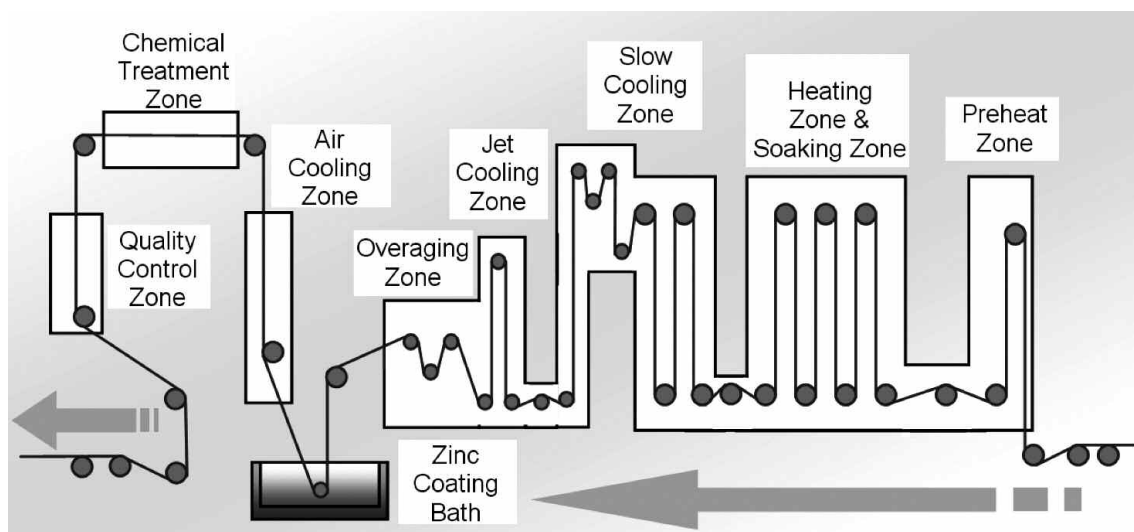


Figure 1. Basic scheme of a Hot Dip Galvanizing Line (HDGL).

Figura 1. Esquema básico de una Línea de Galvanizado en Caliente (LGC).

temperature and cooling rates are controlled to obtain the desired mechanical properties for each steel type. Figure 2 represents one example of thermal treatment that each steel coil has to undergo in the annealing furnace. TMPP2CNG is the final target heating temperature of the strip.

One of the most important stages in the continuous hot-dip galvanizing line (HDGL) is the thermal treatment of the steel strip before zinc immersion. An efficient control of this heat treatment is fundamental both for the process of coating and for improving the properties of the steel on the coil, as well as for reducing energy costs.

The steel strip then runs through a molten-zinc-coating bath followed by an air stream “wipe” that controls the thickness of the zinc finish. Finally, the strip passes through a series of auxiliary processes, winding the product back into a coil.

2.1. Control of the annealing process for the steel strip

There are numerous control techniques for the annealing process that use mathematical models that try to explain the complex mechanisms of heat transfer due to radiation or convection phenomena^[1-4]. These types of phenomena occur inside the furnace and between it and the steel strip itself.

In recent times, importance is being given to the modelling of the behaviour of the steel strip in order to improve the control of the annealing process in an HDGL. Thus, in Prieto *et al.*^[5] a stepwise mathematical model is reported that allows determining the

temperature of the strip based on both its and the furnace’s mathematical characterisation, and which considers the phenomena of conduction, convection and radiation existing in the furnace and also present between it and the steel strip.

However, over the past several years research has been directed more towards the use of neural networks to control the modelling and fine-tuning of steel manufacturing processes. This is due primarily to the fact that these processes and sub-processes are repetitive, highly automated, and have a large number of well-known variables that define them^[6-12].

Most of the papers published that report the use of neural networks for enhancing the annealing process in HDGL focus on the design of models for predicting the set temperatures for the furnace according to the size of the strip and the process conditions^[13-15].

However, in Martínez-de-Pisón *et al.*^[16] we report the use of a dynamic model of temperatures for the steel strip whereby genetic algorithms are used to fine-tune the set speeds and temperatures in the furnace in which there are transitions, within the steel strip, between coils of different sizes. Accordingly, two multi-layer perceptron (MLP) models are developed: the first is used to determine the parameters for the furnace in stationary regime and the second is used to predict the dynamic behaviour of the strip when there are fluctuations in speed or temperature in the furnace. With these two models, and largely with the second one, it is possible to simulate the behaviour of the strip when there are sudden changes between coils of different sizes and, based on that, find the best fitting straight-line for the set signals in order to obtain a heat treatment that is as uniform as possible in that area.

In addition, Bloch *et al.*^[17] develop an RBF network model that seeks to model the energy delivered to the steel strip based on the size and speed of the same. The control system uses that model to determine the furnace’s set temperatures.

The implementation of these models in an industrial plant requires the creation of a different model for each one of the types of steel existing in the database. The problem is that a lot of time and effort is required for generating and validating the different models for each one of the products existing in the company. Furthermore, it may happen that certain coils whose chemical composition differs slightly from the others are incorrectly processed by the control model.

Due to this, it is much more interesting to develop an overall model that can be used not only for the products that already exist in the historical data

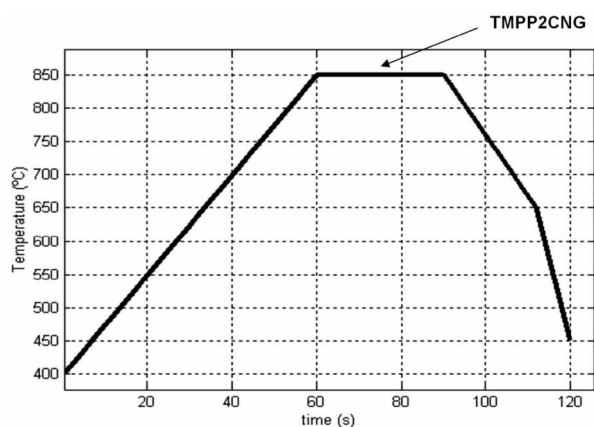


Figure 2. Example of thermal treatment curve in the annealing phase.

Figura 2. Ejemplo de curva de tratamiento térmico en la fase de recocido.

but also for coils with new types of steel or with different sizes to those processed beforehand.

In order to achieve this goal, in addition to the sizes of each coil and the process conditions, the model needs to take into account the chemical composition of the steel in the same.

There follows a description of the steps taken to create this model. This involves the use of several DM techniques in order to determine whether any of the current techniques is an improvement on MLP modelling.

3. DATA MINING PROCESS

3.1. Attributes selection

Data acquisition is obtained from the computer processing area based on the historical data continuously generated during the galvanizing process. The variables are selected according to their relevance to the furnace Heating Zone.

The database consists of 53,910 records obtained from a galvanizing process involving 1,950 coils in 511 castings. The selected variables are:

- WIDTHCOIL: coil width (mm).
- THICKCOIL: coil thickness (mm).
- TMPP1: strip temperature upon entering the Heating Zone (°C).

- TMPP2: strip temperature upon leaving the Heating Zone (°C)
- TMPP2CNG: strip set point temperature at the output of the Heating Zone (°C)
- VELMED: speed of the strip (m/min).
- C, Mn, Si, S, P, Al, Cu, Ni, Cr, Nb, V, Ti, B, N: chemical composition of the steel (in %).
- The temperature of three zones in the furnace heating zone: THC1=initial zone, THC3=intermediate zone, THC5=final zone (°C).

All variables are measured every 100m along the strip. The strip velocity is measured in the centre of the furnace, and it is reasonable to assume that the strip maintains the same velocity throughout the Heating Zone.

The relevant variables and their abbreviations can be found in table I.

Figure 3 shows one example of the data from the historical database process.

3.2. Designing the regression model

The design of the regression model is shown in figure 4. The purpose of this model is to predict the temperature of the strip upon leaving the heating zone at time $t+1$ ($TMPP2(t+1)$) according to:

- The chemical composition of the steel at that moment (C, Mn, Si, S, P, Al, Cu, Ni, Cr, Nb, V, Ti, B, N).

Table I. Relevant variables and their abbreviations.

Tabla I. Variables relevantes y sus abreviaturas.

VARIABLES	
Abbreviation	Meaning
VELMED	Strip velocity inside the Furnace (m/min)
THICKCOIL	Strip thickness at the input of the Furnace (mm)
WIDTHCOIL	Strip width at the input of the Furnace (mm)
TMPP2	Strip temperature at the output of the Heating Zone (°C)
TMPP2CNG	Strip set point temperature at the output of the Heating Zone (°C)
TMPP1	Strip temperature at the input of the Heating Zone (°C)
C, Mn, Si, S, P, Al, Cu, Ni, Cr, Nb, V, Ti, B, N	Chemical composition of steel (in percentage of weight) (%)
THC1	Zone 1 set point temperature (initial Heating Zone) (°C)
THC3	Zone 3 set point temperature (intermediate Heating Zone) (°C)
THC5	Zone 5 set point temperature (final Heating Zone) (°C)

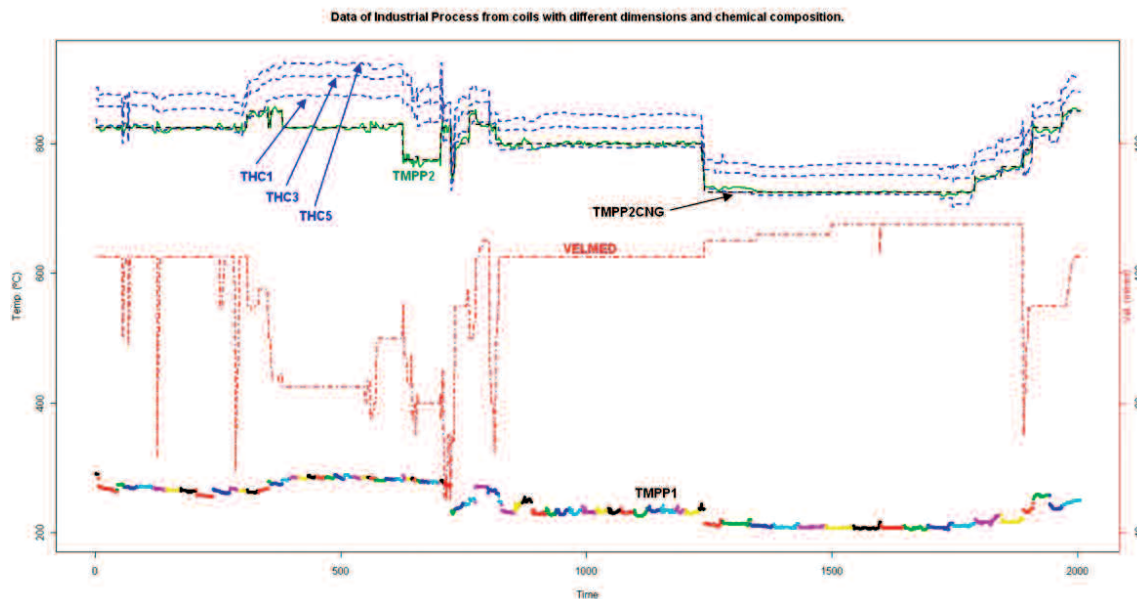


Figure 3. Example of process data extracted from the historical database.

Figura 3. Ejemplo de datos del proceso extraídos de la base de datos de históricos.

- The size of the strip at time t ($THICKCOIL(t)$ and $WIDTHCOIL(t)$).
- The input and output speeds and temperatures of the strip at time t ($VELMED(t)$, $TMPP1(t)$ and $TMPP2(t)$).
- The furnace temperatures in the heating zone at that moment ($THC1(t)$, $THC3(t)$ and $THC5(t)$).
- The difference in the input temperature of the strip between time t and $t+1$:

$$DIFFTMPP1(t) = TMPP1(t+1) - TMPP1(t) \quad (1)$$

- The difference in the speed of the strip between time t and $t+1$:

$$DIFFVELMED(t) = VELMED(t+1) - VELMED(t) \quad (2)$$

- The difference in temperatures in each one of the zones in the furnace between time t and $t+1$:

$$\begin{aligned} DIFFTHC1(t) &= THC1(t+1) - TCH1(t) \\ DIFFTHC3(t) &= THC3(t+1) - TCH3(t) \\ DIFFTHC5(t) &= THC5(t+1) - TCH5(t) \end{aligned} \quad (3)$$

Given that the model has too many input variables, 23, above all due to the high number of elements, 14, in the chemical composition of the steel, use is made of Principal Components Analysis (PCA) to reduce the high dimensionality and eliminate the high dependence between them.

Accordingly, three PCAs are made grouping the variables corresponding to chemical composition, temperature and temperature differences.

Table II presents the results of the PCA corresponding to the chemical composition of the steels. The aim is to include these variables so that in the training process the model can try to learn, in an approximate manner, the complex non-linear relations that may exist between the chemical composition of the steel and the steel's heat transfer and thermal emissivity coefficients.

In order to improve the prediction capacity, each variable pertaining to the chemical composition of the steel is previously multiplied by a weighting coefficient (w_i), established beforehand by agreement of the plant's experts according to the approximate degree of influence it is estimated to have on these processes.

Figure 5 shows the PCA projection of the coils using the first two principal components obtained from the 14 standardised and weighted values of the chemical composition of the coils. There is a large group of coils of one specific type of steel and several smaller groups of coils with different chemical compositions.

From the PCA obtained, a selection is made of the first 4 main axes that explain 87.86 % of the original variance (Table II). This reduces the 14 input variables to 4 variables that are independent of each other.

Figure 6 shows the scatter plot of the temperature variables of the furnace and steel strip. The upper triangular matrix shows very high correlations

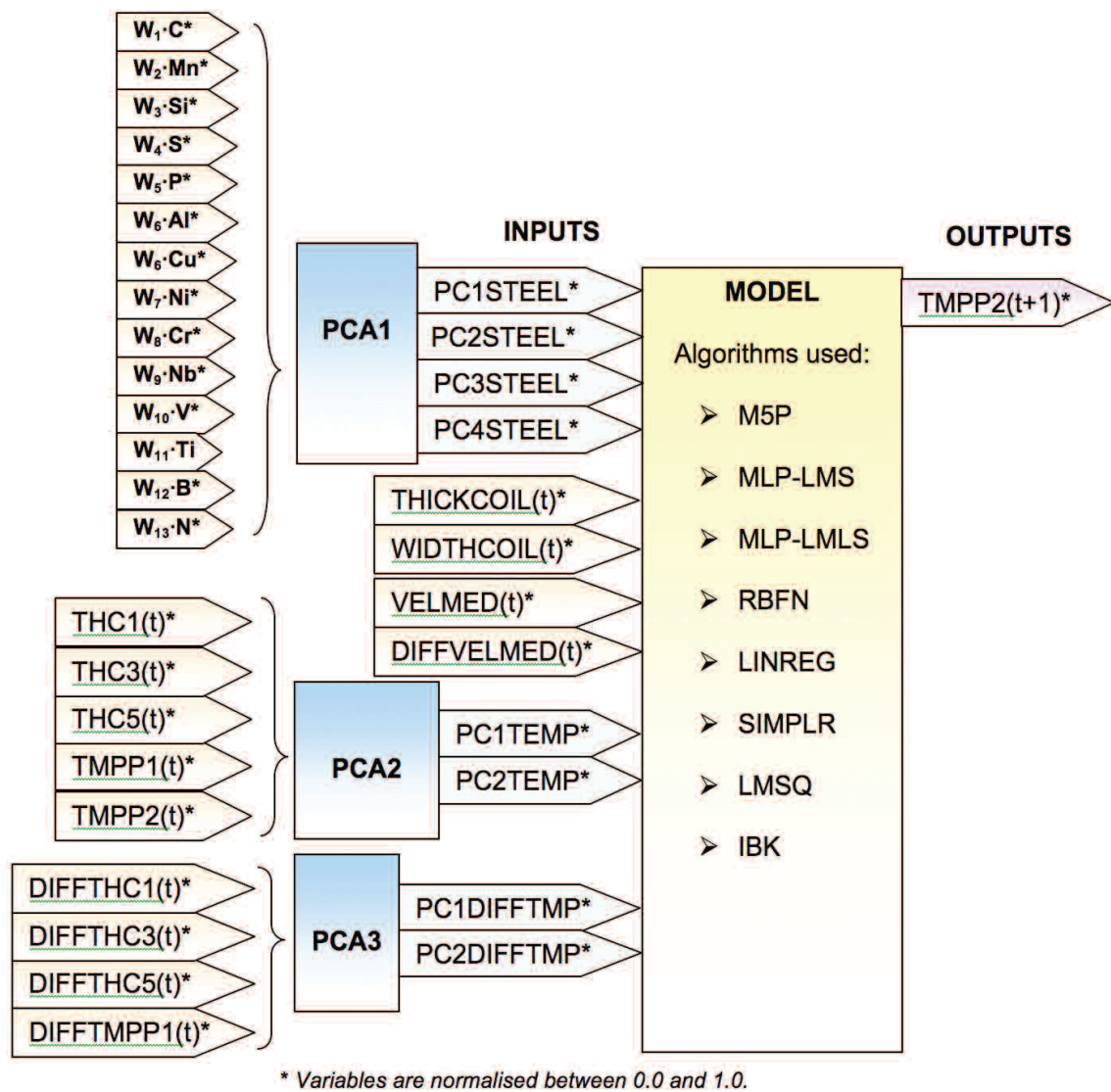


Figure 4. Design of the regression model.

Figura 4. Diseño del modelo de regresión.

Table II. Results of principal component analysis (PCA1) for steel chemical composition.

Tabla II. Resultados del análisis de componentes principales (PCA1) de la composición química del acero.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
Standard deviation	0.64	0.46	0.19	0.16	0.15	0.12	0.12	0.11	0.09	0.08	0.07	0.07	0.06	0.04
Proportion of variance	0.53	0.27	0.05	0.03	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.00
Cumulative proportion	0.53	0.80	0.85	0.88	0.91	0.93	0.94	0.96	0.97	0.98	0.99	0.99	1.00	1.00

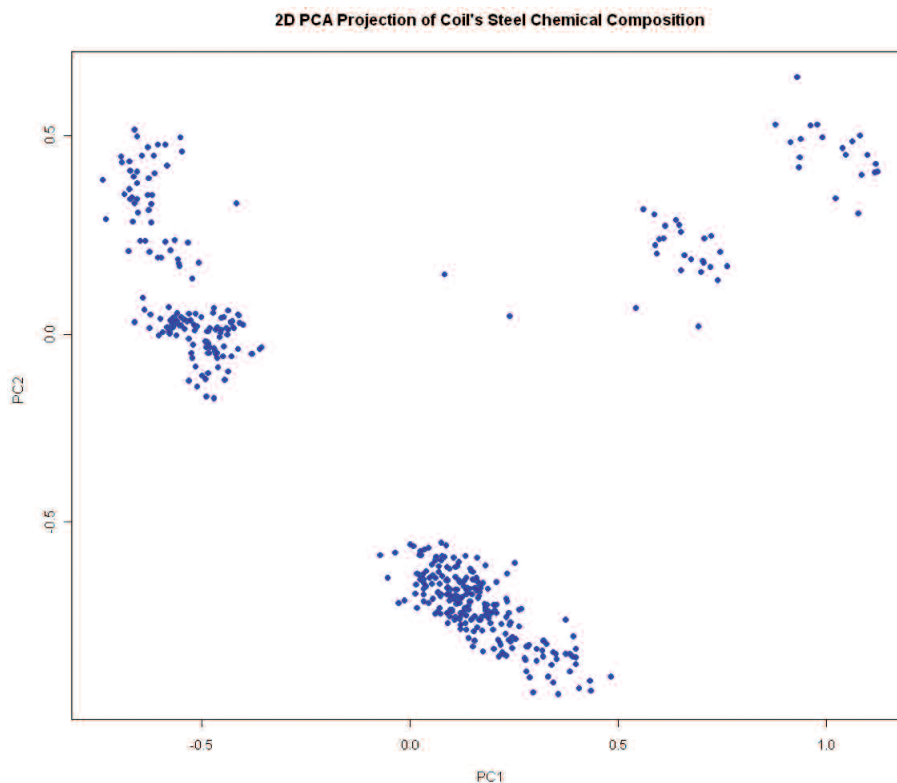


Figure 5. PCA Projection of coils according to the steel chemical composition using the two principal components (PC1 and PC2).

Figura 5. Proyección PCA de las bobinas según la composición química del acero usando los dos componentes principales (PC1 y PC2).

between them. The PCA provides new variables that are a function of the previous ones yet independent of each other. Furthermore, the number of variables required is reduced.

Table III shows the results of the PCA applied to the temperature variables. The two principal axes manage to explain 97 % of the existing variance. This reduces the number of variables from 5 to 2.

As in the previous case, a PCA is made with the variables corresponding to the difference in temperatures between time t and $t+1$. Figure 7 shows the scatter plot of these variables. In this case, the selection of the two principal axes explains 89 % of the total variance (Table IV).

Finally, the model for predicting the temperature of the strip consists of 12 input variables and one output.

3.3. Final pre-processing

In order to improve the prediction capacity of the models and avoid the model learning better from the

more widely used coils and worse from the less used ones, a prior stratified sampling is made with replacement that standardises the number of cases in the database. Accordingly, a hierarchical clustering is performed (Fig. 8) using the 14 variables of the steel's chemical composition, obtaining 4 large groups or clusters. Finally, a sampling is made with replacement of 10,000 records using each one of the clusters to create a uniform database of 40,000 cases.

A random selection is made from the final database of 434 steels to generate the training database and of another 77 steels for the test database. Special care is taken to ensure that the steels in the test database are distributed throughout the entire range of instances (Fig. 9). The aim is to have greater guarantees of success when analysing the degree of generalisation of each one of the trained models.

Subsequently, all the variables are normalised between 0 and 1 to improve the degree of convergence of certain algorithms.

This finally provides a training database consisting of 32,729 records for 1,636 coils of 434 different types of steel, and a test database consisting of 7,271 records for 253 coils of 77 different steels.

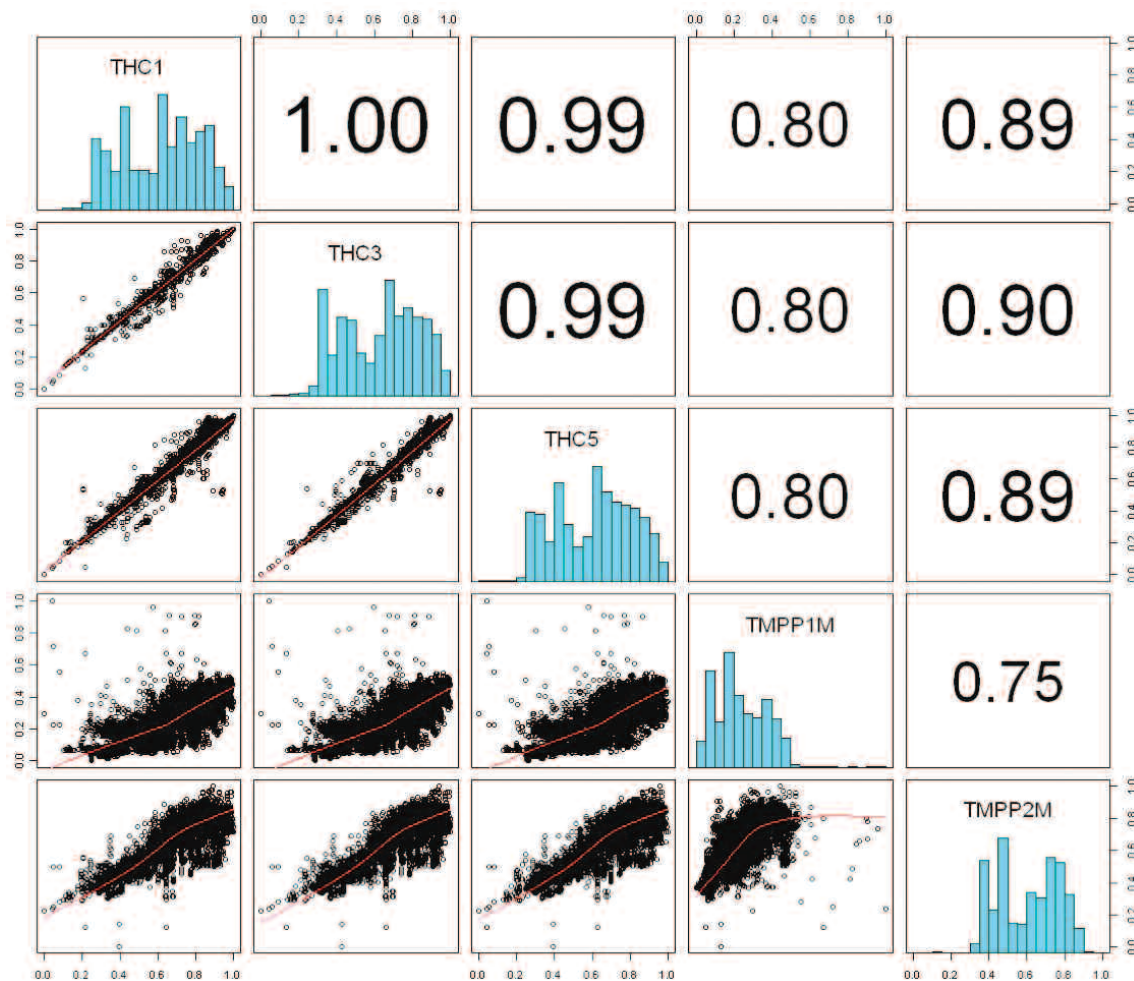


Figure 6. Scatter plot of temperatures.

Figura 6. Diagrama de dispersión por pares de las temperaturas.

Table III. Results of principal component analysis (PCA2) for temperatures of heating zone and strip.

Tabla III. Resultados del análisis de componentes principales (PCA2) de las temperaturas de la zona de calentamiento y de la banda.

	PC1	PC2	PC3	PC4	PC5
Standard deviation	0.396	0.074	0.067	0.022	0.011
Proportion of Variance	0.937	0.033	0.027	0.003	0.001
Cumulative Proportion	0.937	0.970	0.996	0.999	1.000

3.4. Selecting the best data mining techniques

In order to find models that generate a low prediction error, a battery of algorithms are used:

- *M5P algorithm (M5P)*: Implements base routines for generating M5Model trees. A decision list for regression problems is generated using separate-and-conquer. In each iteration, it builds a model tree using M5 and makes the “best” leaf into a rule. Quinlan’s M5P can learn such piece-wise linear models. M5P also generates a decision tree that indicates when to use which linear model^[18].
- *Multilayer Perceptron (MLP)*: A classifier and predictor that uses backpropagation to classify instances. All nodes in this network are sigmoid, except when the class is numeric. In the latter case, the output nodes become

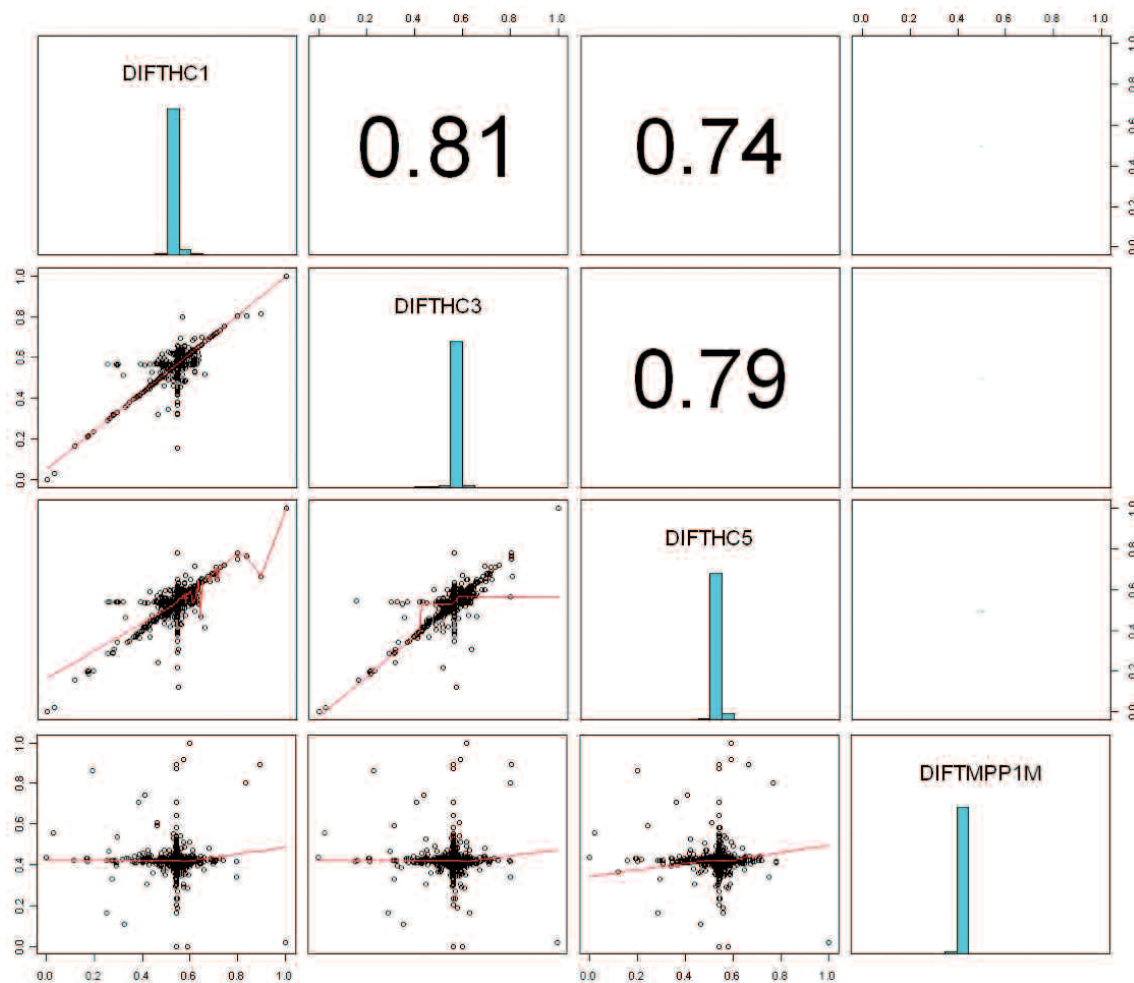


Figure 7. Scatter plot of temperatures differences.

Figura 7. Diagrama de dispersión por pares de las diferencias de temperaturas.

Table IV. Results of principal component analysis (PCA3) for the difference of temperatures of heating zone and strip.

Tabla IV. Resultados del análisis de componentes principales (PCA3) de las diferencias de temperaturas de la zona de calentamiento y de la banda.

	PC1	PC2	PC3	PC4
Standard deviation	0.029	0.018	0.009	0.008
Proportion of Variance	0.638	0.252	0.063	0.046
Cumulative Proportion	0.638	0.890	0.954	1.000

unthresholded linear units^[19 and 20]. Training is performed with networks that have between 1 and 30 neurons in the hidden layer.

- *RBF Network (RBFN)*: Implements a normalized Gaussian radial basis function network. It uses the k-means clustering algorithm to provide the basis functions and learns either a logistic regression (discrete class problems) or a linear regression (numeric class problems). In addition, a symmetric multivariate Gaussian distribution is fitted to the data from each cluster. If the class is nominal, it uses the given number of clusters per class. It standardizes all numeric attributes on a zero mean and unit variance^[19].
- *Linear Regression (LINREG)*: A class for using linear regression for prediction. It uses the Akaike criterion for variable selection and is able to deal with weighted instances^[21].

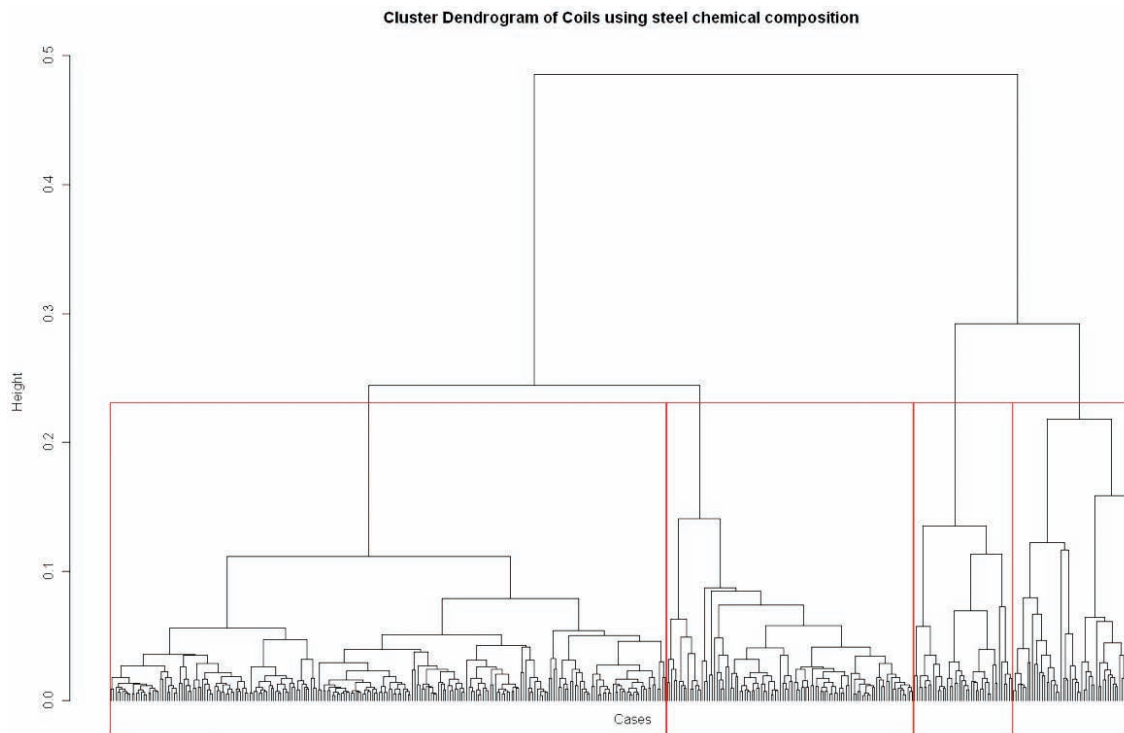


Figure 8. Dendrogram used to obtain homogeneous training cases from four clusters.

Figura 8. Dendrograma utilizado para obtener casos de entrenamiento homogéneos a partir de cuatro grupos.

- *Simple Linear Regression (SIMPLR)*: Uses only the best attribute to obtain the model. It is useful for comparing with other algorithms.
- *LeastMedSq (LMSQ)*: Implements a least median squared linear regression to make predictions. Least squared regression functions are generated from random sub-samples of the data. The least squared regression that has the lowest median squared error is chosen as the final model^[22].
- *IBk (IBk)*: A version of the k-nearest neighbour algorithm. K is the number of neighbours to be used. It also permits the use of distance weighting. As it is a lazy algorithm, there is no training time^[23].

WEKA^[24] suite and AMORE^[25] library of R^[26] software are used to develop the different models.

24 different configurations of these 8 algorithms are trained: 10 MLPs with different numbers of neurons in the hidden layer (1, 2, 3, 4, 5, 7, 10, 15, 20 & 30), 3 of the IBK algorithms with different numbers of k neighbours (1, 2 & 3), 6 RBFNs with different numbers of clusters (3, 5, 10, 15, 20 & 30), 2 M5Ps with the minimum number of individuals per

leaf (4 & 8) and one each for the LINREG, LSQ and SIMPLR algorithms.

To obtain the best precision, ten models of each type of algorithm configuration are trained with 70 % of the data from the training database and the remaining data (30 %) are used to validate each model. By generating 10 models of each algorithm configuration, the influence of local minima is reduced and much more realistic errors are obtained.

The purpose of this work is to determine the algorithm configuration that provide the best prediction or, in other words, the algorithm configuration that yields the lowest Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for other different coils not used for model construction. These errors are:

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n |y(k) - \hat{y}(k)|^2} \quad (4)$$

and

$$MAE = \frac{1}{n} \sum_{k=1}^n |y(k) - \hat{y}(k)| \quad (5)$$

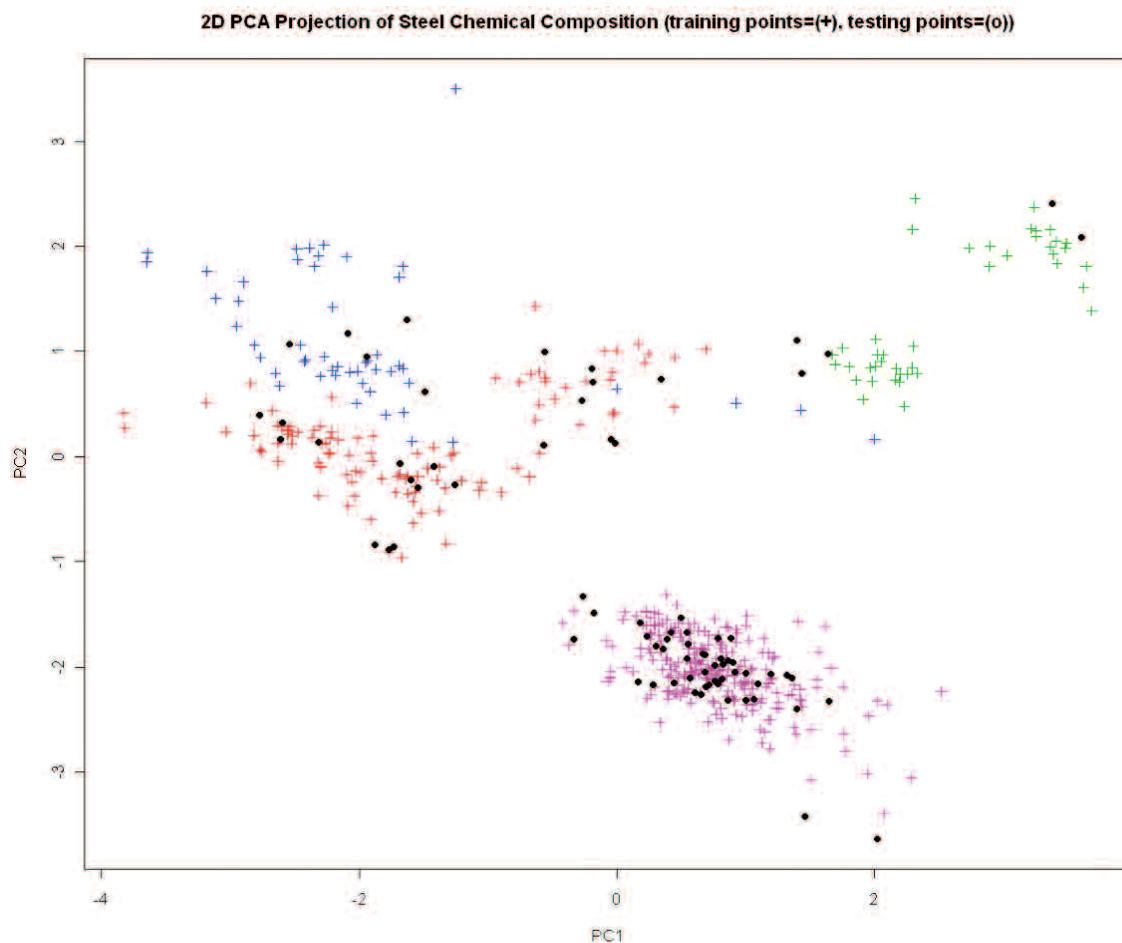


Figure 9. Two-dimension PCA projection of the training and testing database. Coloured training cases (crosses) and testing cases (black dots).

Figura 9. Proyección PCA de dos dimensiones de la base de datos de entrenamiento y testeo. Casos coloreados (cruces) y casos para el testeo (puntos negros).

where, and are, respectively, the measured and predicted outputs and is the number of points in the database used to validate the models.

4. RESULTS

The result of the training and validation process is shown in table V. This table provides a summary of the validation errors arranged by the RMSE corresponding to ten trained models for the 24 algorithm configurations. This table presents the mean (MEAN), maximum (MAX), minimum (MIN) and standard deviation (SD) of RMSE and MAE validation of ten models of each type of algorithm configuration.

The last column shows the time spent creating the ten models and obtaining the validation errors. Obviously, the models that have required the most

training time are the MLP networks with a high number of neurons in the hidden layer.

It can be seen that the validation RMSE, performed with 30 % of cases not used in the creation of the models, is close to 1.0 % for the models based on K-nearest-neighbours (IBk).

The following best models correspond to MLP networks with a high number of neurons in the hidden layer (20 or 30).

The linear models (LINREQ and LMSQ) have approximately 1.4 % more RMSE than the best MLP network.

This is followed by M5P regression trees with 4 or 8 cases per leaf, with an RMSE of 5 %. It has a very low MAE (0.96 %), but the high RMSE indicates that they have a significant number of high residuals.

Finally, the radial basis function networks (RBFNs) are the ones that record the worst performance.

Table V. Results of training and validating process. Validation errors for each model's configuration (ordered by the mean of the root mean squared error (RMSEMEAN)).

Tabla V. Resultados del proceso de entrenamiento y validación. Errores de validación para cada tipo de configuración de los modelos (ordenados según la media de la raíz del error cuadrático medio (RMSEMEAN)).

Algorithm	RMSEMEAN	RMSEMAX	RMSEMIN	RMSESD	MAEMEAN	MAEMAX	MAEMIN	MAESD	TIME(s)
IBk(K=1)	0.0103	0.0116	0.0092	0.0009	0.0023	0.0025	0.0022	0.0001	0.08
IBk(K=2)	0.0123	0.0135	0.0108	0.0008	0.0036	0.0039	0.0034	0.0001	0.08
IBk(K=3)	0.0135	0.0143	0.0129	0.0005	0.0046	0.0049	0.0045	0.0001	0.08
MLP(30)	0.0264	0.0293	0.0239	0.0019	0.0168	0.0195	0.0147	0.0015	18043.05
MLP(20)	0.0271	0.0304	0.0241	0.0022	0.0172	0.0195	0.0154	0.0013	9364.69
MLP(15)	0.0271	0.0304	0.0251	0.0016	0.0179	0.0211	0.0163	0.0014	8651.93
MLP(10)	0.0275	0.0302	0.0248	0.0016	0.0180	0.0200	0.0160	0.0012	4804.70
MLP(07)	0.0336	0.0362	0.0303	0.0020	0.0210	0.0234	0.0188	0.0015	2945.15
MLP(05)	0.0339	0.0363	0.0307	0.0018	0.0213	0.0235	0.0190	0.0013	1466.93
MLP(04)	0.0341	0.0363	0.0315	0.0016	0.0216	0.0239	0.0198	0.0012	522.27
MLP(03)	0.0347	0.0378	0.0317	0.0018	0.0219	0.0251	0.0198	0.0016	430.55
MLP(02)	0.0371	0.0439	0.0335	0.0029	0.0237	0.0309	0.0208	0.0029	1334.58
MLP(01)	0.0401	0.0449	0.0366	0.0026	0.0261	0.0312	0.0236	0.0027	697.59
LINREG	0.0406	0.0413	0.0388	0.0009	0.0262	0.0267	0.0259	0.0002	2.85
LMSQ	0.0453	0.0504	0.0419	0.0023	0.0276	0.0293	0.0263	0.0008	998.37
M5P (4)	0.0507	0.3369	0.0162	0.1006	0.0096	0.0123	0.0088	0.0010	138.19
M5P (8)	0.0508	0.3369	0.0166	0.1006	0.0097	0.0122	0.0090	0.0009	136.11
RBFN(30)	0.0611	0.0655	0.0558	0.0032	0.0385	0.0417	0.0349	0.0024	196.22
SIMPLR	0.0634	0.0640	0.0624	0.0005	0.0445	0.0448	0.0438	0.0003	0.81
RBFN(20)	0.0683	0.0713	0.0635	0.0027	0.0446	0.0471	0.0411	0.0023	142.65
RBFN(15)	0.0728	0.0790	0.0667	0.0036	0.0489	0.0540	0.0437	0.0034	118.77
RBFN(10)	0.0760	0.0813	0.0727	0.0026	0.0512	0.0552	0.0481	0.0025	86.24
RBFN(05)	0.0899	0.0946	0.0866	0.0027	0.0653	0.0708	0.0615	0.0035	39.37
RBFN(03)	0.1041	0.1467	0.0940	0.0155	0.0796	0.1261	0.0711	0.0165	40.87

When observing these results, performed with simple validation, a researcher might be tempted to use the models based on K-nearest-neighbours, due to the excellent results they record (1 % RMSE and 0.23 % MAE).

The problem is that these types of techniques select, for a new case to be predicted and according to a criterion of distance, the nearest K cases and return the mean of the same as the result. As the data correspond to the means of industrial processes, it is highly likely that for a new case of the validation database (30 %) there are several cases repeated within the training database (70 %). For this reason, algorithms of this kind generate very good results in the training phase when there are many repeated cases in the databases. But when we use this model for new cases of coils with different types of steel or sizes, the results worsen considerably.

Table VI shows the errors of previous models with a new test database made up of coils with different

types of chemical compositions and sizes for the steel. These coils are different to the ones used for generating the training and validation databases.

When comparing the models with data from new coils, it can be seen that the algorithms based on K-nearest-neighbours (4.5 % of RMSE) respond much worse than the MLP networks with a mean number of neurons (15 or 20) with an RMSE of 2.63 %.

It can be clearly seen that MLP networks generate overall models that can more reliably predict other types of steel, even those that have not previously been introduced into the training database.

Figure 10 shows the analysis of residuals of the best model corresponding to the MLP network with 20 neurons in its hidden layer. This graph shows the normal distribution of the residuals revealing the absence of structures not explained by the model.

To use the final model and simulate the behaviour of the strip, consideration is given to the strip that would be measured by the pyrometers, at time t , at the furnace

Table VI. Test errors for each model's configuration (ordered by root mean squared error (RMSETEST)).

Tabla VI. Errores de testeo para cada tipo de configuración de los modelos (ordenados según la raíz del error cuadrático medio (RMSETEST)).

Algorithm	RMSETEST	MAETEST	TIME(s)
MLP(20)	0.0263	0.0176	536.24
MLP(15)	0.0291	0.0193	140.43
MLP(07)	0.0300	0.0205	1577.49
MLP(03)	0.0308	0.0215	738.72
MLP(04)	0.0310	0.0215	125.05
MLP(10)	0.0311	0.0196	2187.74
MLP(02)	0.0312	0.0215	29.68
MLP(05)	0.0313	0.0211	594.93
MLP(30)	0.0317	0.0187	6285.85
MLP(01)	0.0355	0.0243	1.68
LINREG	0.0356	0.0246	0.39
LMSQ	0.0425	0.0298	146.30
IBk(K=3)	0.0452	0.0281	0.01
IBk(K=2)	0.0458	0.0287	0.01
IBk(K=1)	0.0465	0.0291	0.01
M5P (8)	0.0466	0.0284	19.29
M5P (4)	0.0503	0.0287	20.04
SIMPLR	0.0584	0.0446	0.10
RBFN(30)	0.0656	0.0424	30.37
RBFN(20)	0.0772	0.0548	20.82
RBFN(10)	0.0774	0.0559	9.82
RBFN(15)	0.0798	0.0570	17.63
RBFN(05)	0.0920	0.0651	8.05
RBFN(03)	0.0979	0.0723	3.84

input ($TMPP1M(t)$) and output ($TMPP2M(t)$), the set temperatures for the furnace ($THCx(t)$), the set speed for the strip ($VELMED(t)$), and their differences ($DIFFTHCx(t)$, $DIFFVELMED(t)$), which would be supplied by the control model; and the information corresponding to the coil being processed at that moment (chemical composition of the steel, width ($WIDTHCOIL(t)$) and thickness of the strip ($THICKCOIL(t)$)). These data provide the projections of the axes of the selected PCAs ($PCxSTEEL(t)$, $PCxTEMP(t)$ and $PCxDIFFTEMP(t)$). Finally, the preceding variables give the temperature for the strip at time $t+1$ ($TMPP2(t+1)$).

Figures 11 and 12 show the results of using the model for simulating the temperature of the strip (line of points) compared to its true temperature (thick black line). The historical information used

corresponds to other dates of the annealing process with coils that do not appear in either the training or the test databases.

As can be seen, the model's behaviour is fairly consistent with the steel strip's dynamic performance. Table VII presents the final results of this simulation process with the new database consisting of 59 coils with 25 different steels, widths and thicknesses. The mean error is 4.18°C and the maximum does not exceed 25.43°C .

A wide range of steels were used for training and testing: steels for cold rolling or drawing, structural steels, high yield-strength, low alloy steels, TRIP steels, multiphase steels, dual phase steels, etc.

For the testing database, some of the coils selected were of steel types which were not already on record in the database. Others were of the same type as coils on record in the training database but their actual chemical composition differed. Special care was also taken to select coils with dimensions other than those used in the training database. In short, coils with a range of dimensions and steel types markedly different from those in the training database were used to check the degree to which the model obtained could be generalised.

5. CONCLUSIONS

This paper shows that the use of classic techniques of simple or cross validation for determining the best model based on historical data on the annealing process can lead us to choose models that closely fit products that have already been processed but which are less efficient when used for predicting new ones.

In order to obtain overall prediction models that are capable of predicting the strip's dynamic performance in the event of temperature and speed fluctuations and which take into account the size and type of steel on the coil being processed, it has been shown that MLP neural networks continue to be some of the more promising techniques for the design of overall prediction models and outperform other Data Mining techniques currently being used.

The final model has proven to be efficient at dealing with new types of coils and process conditions. Its use can help to improve control systems and conveniently design the parameters in the transition zones between coils in order to achieve a more uniform treatment in this area.

It should be pointed out that the models developed are based always on data from cold-rolled coils. This model would not be suitable for predicting the behaviour of strips of hot pickled coils because their surface conditions are substantially different

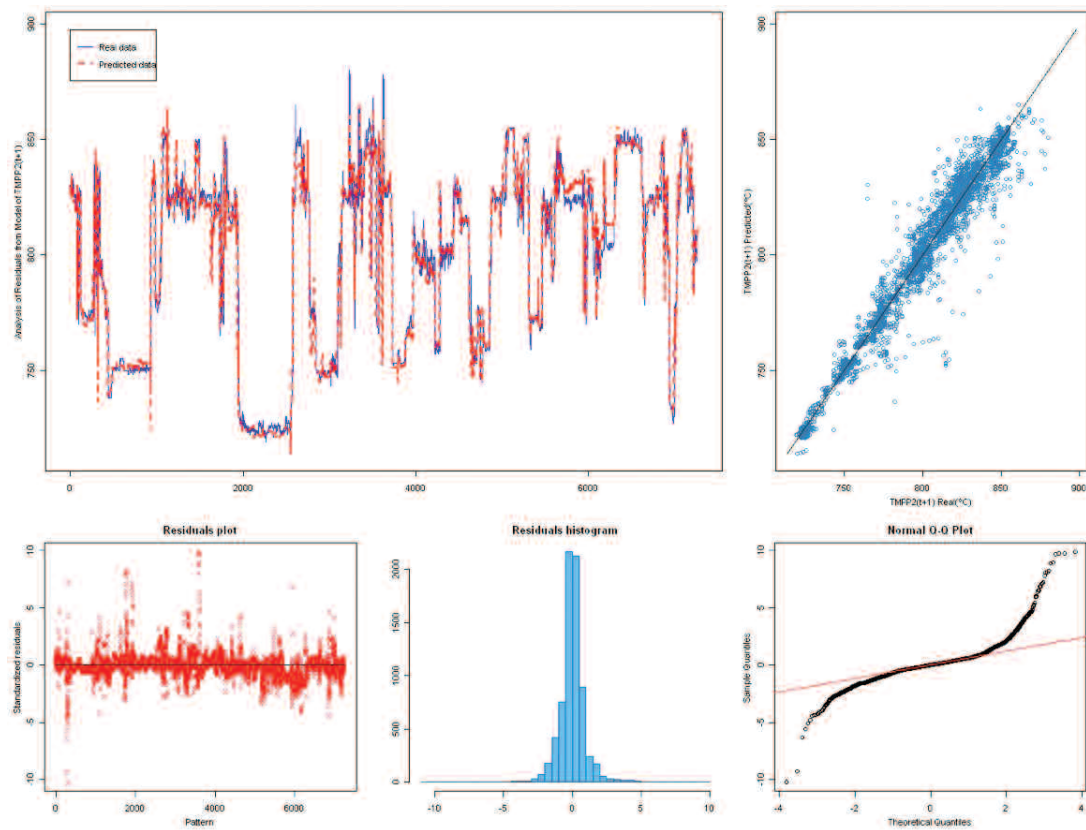


Figure 10. Analysis of residuals of the best model obtained from the test database.

Figura 10. Análisis de los residuos del mejor modelo obtenidos de la base de datos de testeo.

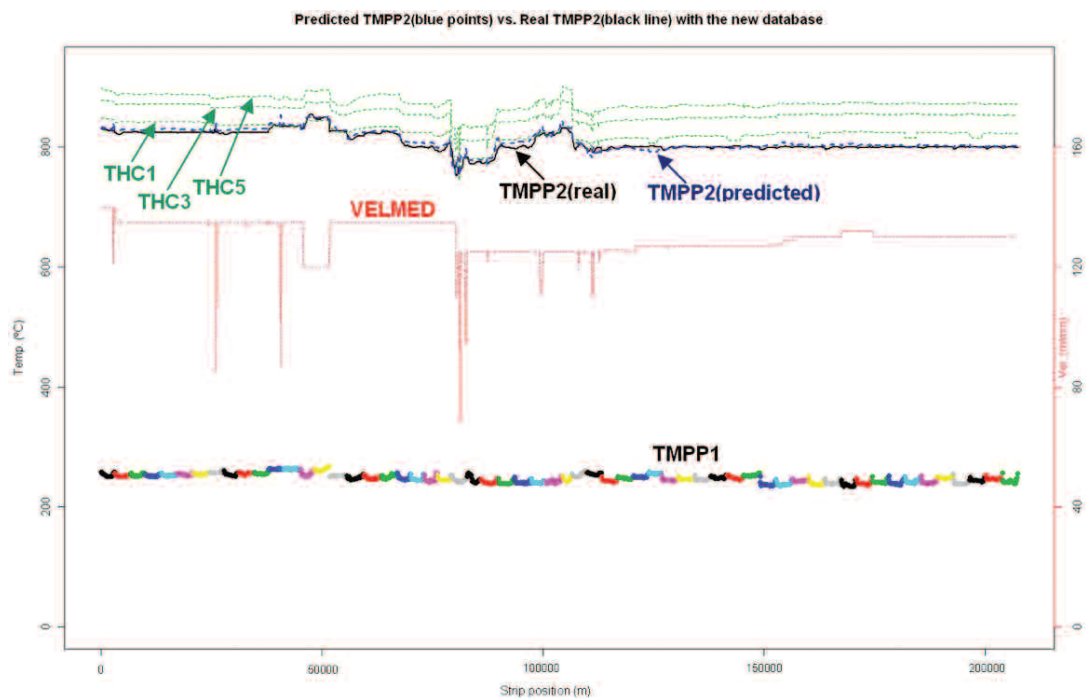


Figure 11. Predictions results of TMPP2 with the new database.

Figura 11. Resultados de la predicción de TMPP2 con la nueva base de datos.

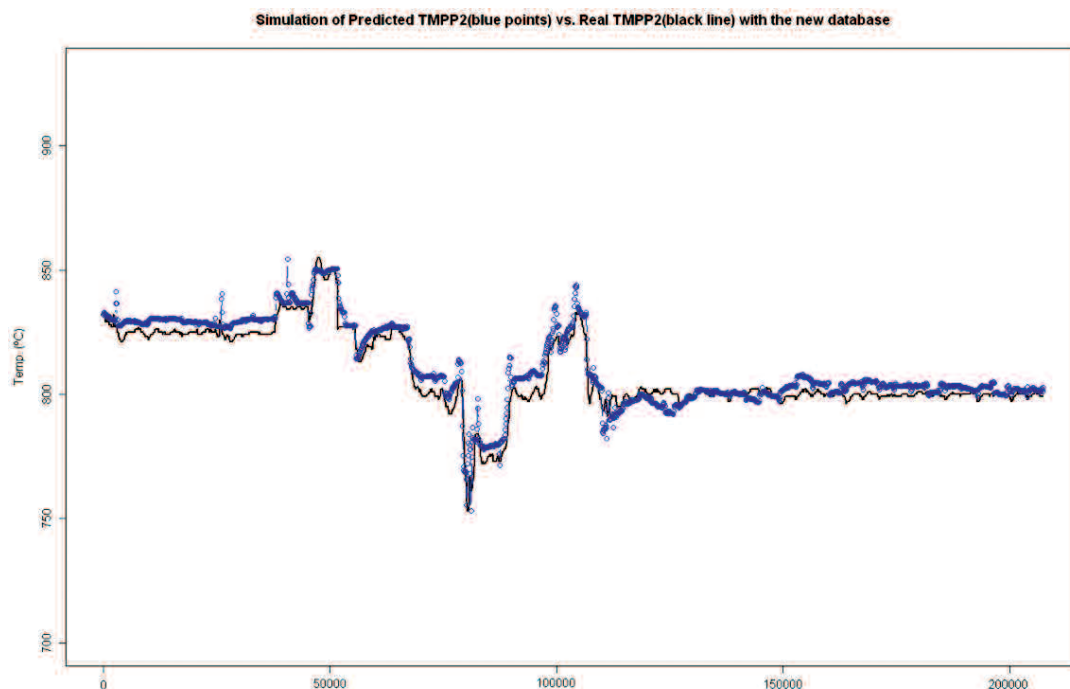


Figure 12. Enlargement of the prediction results in Fig. 11.

Figura 12. Ampliación de los resultados de la predicción de la Fig. 11.

Table VII. Final results of best model with the new database.

Tabla VII. Resultados del mejor modelo con la nueva base de datos.

Description	Value
Number of coils of database	59
Different type of steels into database	25
MAE	4.18 °C (1.67 %)
RMSE	5.13 °C (2.05 %)
Abs(error) (range)	0.0-25.43 °C
THICKCOIL (range)	0.601-0.775 mm.
WIDTHCOIL (range)	805-1180 mm.

from those of cold-rolled steel. For instance, pickled steel coils may contain scaly residues, may be rougher, may have more peaks per square centimetre, etc. All these factors have a considerable influence on the emissivity of steel and therefore on its final temperature. For products of such types to be included in the model, further variables would have to be added to take roughness, the percentage of scale, etc. into account.

Acknowledgments

The authors thank the Dirección General de Investigación of the Spanish Ministry of Science and Innovation for the financial support of the projects DPI2006-03060, DPI2006-14784, DPI-2006-02454 and DPI2007-61090; and the European Union for the project RFS-PR-06035.

Finally, the authors also thank the Autonomous Government of La Rioja for its support through the 3th Plan Riojano de I+D+i.

REFERENCES

- [1] R. Mehta and S.S. Sahay, *J. Mat. Eng. Perform* 18 (2009) 8-15.
- [2] S.S. Sahay and K. Krishnan, *Ironmaking Steelmaking* 34 (2007) 89-94.
- [3] W. Wenfei, Y. Fan, Z. Xinxin and Z. Yi, *J. Therm. Sci.* 11 (2002) 2-134.
- [4] X. Zhang, F. Yu, W. Wu and Y. Zuo, *Int. J. Thermophys.* 24 (2003) 1395-1405.
- [5] M.M. Prieto, F.J. Fernández and J.L. Rendueles, *Ironmaking Steelmaking* 32 (2005) 165-170.
- [6] Y. Lu and S.W. Markward, *IEEE Trans. on Neural Networks* 6 (1997) 1328-1337.

- [7] J. Tenner, D.A. Linkens, P.F. Morris and T.J. Bailey, *Ironmaking Steelmaking* 28 (2001) 15-22.
- [8] A.A. Gorni, *The modelling of hot rolling processes using neural networks*, http://www.gorni.eng.br/e/neural_1998.html, 1998.
- [9] M. Schlang, B. Lang, T. Poppe, T. Runkler and K. Weinzierl, *Control Eng. Pract* 9 (2001) 975-986.
- [10] A. González-Marcos, J.B. Ordieres-Meré, A.V. Pernía-Espinoza and V. Torre-Suárez, *Rev. Metal. Madrid* 44 (2008) 29-38.
- [11] F. J. Martínez-de-Pisón, J. Ordieres, A. Pernía, F. Alba and V. Torre, *Rev. Metal. Madrid* 43 (2007) 325-336.
- [12] J. B. Ordieres-Meré, A. González-Marcos, J.A. González and V. Rubio, *Ironmaking Steelmaking* 31 (2004) 43-50.
- [13] A. V. Pernía-Espinoza, M. Castejón-Limas, A. González-Marcos and V. Lobato-Rubio, *Ironmaking Steelmaking* 32 (2005) 1-9.
- [14] C. Schiefer, H.P. Jörgl, F.X. Rubenzucker and H. Aberl, *Proc. 14th IFAC World Congress*, 1999, pp. 61-66.
- [15] Y.I. Kim, K. Cheol Moon, B. Sam Kang, C. Han and K. Soo Chang, *Control Eng. Practice* 6 (1998) 1009-1014.
- [16] F.J. Martínez-de-Pisón, F. Alba, M. Castejón and J.A. González, *Ironmaking Steelmaking* 33 (2006) 344-352.
- [17] G. Bloch, F. Sirou, V. Eustache and P. Fatrez, *IEEE T. Neural Networks* 8 (1997) 910-918.
- [18] J.R. Quinlan, *Proc. Australian Joint Conf. on Art. Int. World Scientific*, 1992, pp. 343-348.
- [19] S. Haykin, *Neural networks, a comprehensive foundation* (2nd ed.), Prentice Hall, New Jersey, EE.UU, 1999.
- [20] A.V. Pernía-Espinoza, J.B. Ordieres-Meré, F.J. Martínez-de-Pisón and A. González-Marcos, *Neural Networks* 18 (2005) 191-204.
- [21] G.N. Wilkinson and C.E. Rogers, *Appl. Statistics* 22 (1973) 392-399.
- [22] S. Portnoy and R. Koenke, *Statistical Sci.* 12 (1997) 279-300.
- [23] D. Aha and D. Kibler, *Mach. Learning* 6 (1991) 37-66.
- [24] I.H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*, 2nd Edition, Morgan Kaufmann, San Francisco, EE.UU., 2005.
- [25] M. Castejón-Limas, J.B. Ordieres Meré, E.P. Vergara, F.J. Martínez-de-Pisón, A.V. Pernía and F. Alba, *The AMORE package: A MORE flexible neural network package*, CRAN Repository, <http://wiki.r-project.org/rwiki/doku.php?id=packages:cran:amore>, 2009.
- [26] R Development Core Team, *R: A language and environment for statistical computing*, <http://www.R-project.org>, R Foundation for Statistical Computing, Vienna, Austria, 2008, URL.