

Document downloaded from:

<http://hdl.handle.net/10251/60940>

This paper must be cited as:

Debón Aucejo, AM.; García-Díaz, JC. (2012). Fault diagnosis and comparing risk for the steel coil manufacturing process using statistical models for binary data. *Reliability Engineering and System Safety*. 100:102-114. doi:10.1016/j.ress.2011.12.022.



The final publication is available at

<http://dx.doi.org/10.1016/j.ress.2011.12.022>

Copyright Elsevier

Additional Information

# Fault diagnosis and comparing risk for steel coil manufacturing process using statistical models for binary data. <sup>\*</sup>

A. Debón <sup>a,\*</sup> J. Carlos Garcia-Díaz <sup>a</sup>

<sup>a</sup>*Centro de Gestión de la Calidad y del Cambio  
Universitat Politècnica de València.  
E-46022. Valencia. Spain.*

---

## Abstract

Use of advanced statistical models can help industries to design more economical and rational investment plans. Fault detection and diagnosis is an important problem in continuous hot dip galvanizing. Increasingly stringent quality requirements in the automotive industry also require ongoing efforts in process control to make processes more robust. Robust methods for estimating the quality of galvanized steel coils are a tool for the comprehensive monitoring of the performance of the manufacturing process. This study applies different statistical regression models: generalized linear models, generalized additive models and classification trees to estimate the quality of galvanized steel coils on the basis of short time histories. The data, consisting of 48 galvanized steel coils, was divided into sets of conforming and nonconforming coils. Five variables were selected for monitoring the process: steel strip velocity and four bath temperatures.

The present paper reports a comparative evaluation of statistical models for binary data using Receiver Operating Characteristic (ROC) curves. A ROC curve is a graph or a technique for visualizing, organizing and selecting classifiers based on their performance. The purpose of this paper is to use them in research to obtain the best model to predict defective steel coil probability. In relation to the work of other authors who only propose goodness of fit statistics, **we should highlight one distinctive feature of the methodology presented here, which is the possibility of comparing the different models with ROC graphs which are based on model classification performance.** Finally, the results are validated by bootstrap procedures.

*Key words:* Generalized Linear Models, Generalized Additive Models, ROC Curve, Bootstrapping.

---

## 1 Introduction

According to Himmelbau (1978), the term fault is generally defined as a departure from an acceptable range of an observed variable or a calculated parameter associated with a process. A fault is defined as abnormal process behaviour whether associated with equipment failure, sensor degradation, set point change or process disturbances. Fault detection is the task of determining whether a fault has occurred. Fault diagnosis is the task of determining which fault has occurred. Fault detection and diagnosis is an important problem in process engineering. A diagnosis problem is characterised by a set of observations to be explained. When these observations are different from the expected behaviour, we define a faulty state.

Galvanized steel is a value added product, furnishing effective performance by combining the corrosion resistance of zinc with the strength and formability of steel. Fault detection and diagnosis is an important problem in continuous hot dip galvanizing and the increasingly stringent quality requirements in the automotive industry have also demanded ongoing efforts in process control to make the process more robust.

**Statistical Process Control (SPC) is a tool oriented to improving quality and productivity. Control charts are a mature tool widely employed in industry, from the simplest Shewhart control chart to the more elaborate and powerful Cumulative Sums (CUSUM) or Exponentially Weighted Moving-Average (EWMA) charts (Montgomery, 2009). Several researchers have investigated applications of neural networks (Pacella et al., 2004), and genetic algorithms (Aparisi and García-Díaz, 2007) for manufacturing quality control. Many authors have proposed various statistical methods to quality control in the steel industry. Some of the widely used traditional statistical tools applied for prediction and fault diagnosis in steel coils manufacturing are Principal Components Analysis (García-Díaz, 2009a), and Logistic Regression (García-Díaz, 2009b). Non-linear statistic and non-parametric strategies as neural networks (Zheng et al., 2007; Pernía-Espinoza et al., 2005), have been used for control quality in hot-dip galvanizing of coils manufacturing. Modern techniques in Data Mining and Artificial Intelligence have been used to design dynamic model for predicting the temperature of the steel strip in the annealing process in a hot-dip galvanizing line (Martínez-De-Pisón et al., 2010). Ordieres-Meré et al. (2010)**

---

\* This work was supported by a grant from PAID-06-08 (Programa de Apoyo a la Investigación y Desarrollo) of the Universidad Politécnica de Valencia

\* Corresponding author. Tel: +34 963877007 (Ext. 74961); Fax: +34 963877499; E-mail: andeau@eio.upv.es

**presents a comparative assessment of the different techniques currently being used for estimate the mechanical properties of steel coils before being processed on the hot-dip galvanizing line.**

The problem of determining which factors influence the correct manufacture of parts has been considered from different perspectives and statistical methodologies. The investigation of this issue has made frequent use of traditional statistical methods. Discriminant analysis is one of the most widely used techniques for classification, but the requirement of normality and homoscedasticity are not met. As a result it is often necessary to use different techniques, such as logistic regression. Other classification techniques based on decision trees have also been considered more recently. One is “Classification And Regression Trees” (CART) from machine learning starting in the 1970s (see Qinlan (1993) for an overview), but it was Breiman et al. (1984) who introduced these models into statistical science. The study of statistical models to determine which factors influence the correct manufacture of parts is essential to influence the aspects that guarantee the successful implementation of quality control strategies. This study applies different statistical regression models: generalized linear models, generalized additive models and classification trees for estimating the quality of galvanized steel coils on the basis of short time histories. However, there is a research gap in the comparison of models based on their results or discrimination capacity, given that most of the decisions about the adjustment of one model over another are based on measures of goodness of fit. To this end, this study also introduces ROC curves, which can help us by allowing us to compare models.

Firstly, this study determine which steel coil manufacturing process characteristics affect defective risk and which models fit the data best. In addition to proper diagnosis of the model after a detailed study of the basic assumptions from their residuals.

Secondly, the present paper reports a comparative evaluation of statistical models for binary data using Receiver Operating Characteristic (ROC) curves following the methodology used for a pipe failure database by Debón et al. (2010). A ROC curve is a technique for visualizing, organizing and selecting classifiers based on their performance. ROC graphs are commonly used in medical decision making, and in recent years have been increasingly used in machine learning and data mining research. The purpose of this paper is to use them in research to obtain the best model to predict defective steel coil probability.

Finally, bootstrap techniques are used to validate the results of the ROC curves.

The contents of this article are structured as follows. In the second section, we

briefly present the three models for binary data: Generalized Linear Models, Generalize Additive Models and Classification And Regression Trees. We also introduce a Receiver Operating Characteristics (ROC) as a way to evaluate concordance between models and real data. The third section starts by describing defective steel coil data. In this third section, we apply the three different models to these data, commenting on their advantages and disadvantages, as well as on their suitability in the failure risk analysis in question. The section ends with a comparison of two different fits by means of the ROC curves. Finally, we draw some conclusions in which we try to assess their validity not only for this case but also, to some extent, for any other data.

## 2 Statistical Models

Statistical models of differing complexity have been suggested in the literature for predicting fault diagnosis in the steel coil manufacturing process, from Logistic regression to Generalised Linear Models (GLM). A review of these statistical models applied in quality control can be found in **El Aroui and Lavergne (1996)**; **Skinner et al. (2004)**; **Hosmer and Lemeshow (2000)**. The most recent contributions have also been compiled by Myers et al. (2010) with special emphasis on applications in engineering and the usefulness of these models by means of goodness of fit statistics. These models were designed to show the impact of each predictive variable for the risk of failure of a individual product.

GLMs have several areas of application ranging from biological and biomedical fields to economics, occurrence of natural phenomena including earthquakes and hurricanes (Liu et al., 2005, 2008) quality control and reliability engineering (Hamada and Nelder, 1997), but its use in engineering disciplines is very recent. GLMs are also applied to statistical quality control because there are many cases of nonnormal responses in industrial experiments (Lewis et al., 2001). Another possibility is a binary response variable. Studies where the response variable is either “success” or “failure” (i.e., 0 or 1) are fairly common in nearly areas of science and engineering. The GLMs has been used in the manufacturing process field (Myers et al., 2010; Roshan and Adya, 2001). The application of classification and regression tree models in quality control problems has also been examined in recent years (Guh, 2005). Steel coils manufacturing is extremely competitive and the financial margins that differentiate between success and failure are very tight, with most established steel companies needing to compete, produce and sell at global level. The current operating philosophy is try to maintain the strip velocity at the maximum value or as high as possible, in order to keep

to the productivity goal. It is of practical interest to locate some operating conditions that achieve an optimum within the region of strict compatibility rules based on coil characteristics. The idea is to find the classifier that is capable of predicting the quality measure of product based on its manufacturing parameters. CART replaces equation of usual regression type model with set of rules. This future is an important aid for the process engineer as it allows him to change the operating parameters as the line speed in order to produce high-quality galvanized steel economically.

## 2.1 Generalised Linear Models (GLM)

GLM are generalized linear regressions to allow for non-normal, count data and non-linear transformations. They link the mean response of a specified conditional distribution to a predictor function. They are based on an assumed probability mass function (PMF) for discrete (count) data and a link function that connects the parameters of this PMF to the available covariates. Then a GLM provides a method for estimating a function of the average of the response variable as a linear combination of the set of predictive variables, that is

$$l(E(Y|X)) = l(m) = \eta(x) = \beta_0 + \sum_{i=1}^p \beta_i x_i. \quad (1)$$

The function of the response average,  $l(m)$ , is called the *link* function, and is considered to be the same as a linear function of the predictors,  $\eta(x)$ , which is called a *linear predictor*. **where  $E(Y/X) = m$  is the expected value of  $Y$ ;  $\sum_{i=1}^p \beta_i x_i$  is the linear predictor, a linear combination of unknown parameters,  $\beta$  and independent variables  $x_i$ .** Each component  $y_i$  of  $Y$  has a Binomial, Poisson Gamma, **Negative Binomial or Conway-Maxwell Poisson** distribution.

The GLM comprehensive reference is McCullagh and Nelder (1989). GLM with qualitative predictors are well described in Agresti (2002).

Logistic regression is a specific type of GLM. This model predicts the probability of a discrete outcome, such as group membership (e.g. galvanized steel coils conforming and nonconforming), from a set of explanatory variables that may be discrete, continuous, or dichotomous or a combination of any of these. The dependent variable in a binary logistic regression model takes the value 1 (nonconforming galvanized steel coils) with a probability of  $p$ , or the value 0 (conforming galvanized steel coils) with a probability of  $(1 - p)$ . The relationship between the dependent and independent variables is assumed to be

non-linear. The estimated regression model may be used to predict the probability of coils conforming given the values of independent variables. Using the sample data, we obtain the model of estimation of probability function of  $p$ .

The application of GLM as tool in coils steel manufacturing quality control may be found in García-Díaz (2009b). The author proposed a logistic regression model to predict the probability of coils conforming given the values of independent process variables: steel strip velocity, and four bath temperatures. The best-fit model results show that steel strip velocity and one bath temperature variables are statistically significant. The results from using the logistic GLM suggest that it can provide a sound basis for estimating the probability of conforming coils and can be useful for quality control in manufacturing process. This type of model explicitly models zero counts, and, in many cases, a utility would be more interested in the probability of having at least nonconforming coil than in predicting the precise number of nonconforming coils.

## 2.2 Generalised Additive Models (GAM)

Generalised additive models (GAM) are a natural extension of GLM in the sense that they adjust nonparametric functions to study the relationship between predictive variables and the answer. Non-parametric relationships between response and predictor variables were expressed in terms of smooth functions. In effect, the predictor  $\eta$  no longer has to be a linear function of predictive variables, but is

$$\eta = \alpha + \sum_{j=1}^p f_j(x_j) + \epsilon \quad (2)$$

where  $f_j$  is a smooth function (splines, locally-weighted regression, ...).  $\alpha + \sum_{j=1}^p f_j(x_j) + \epsilon$  **is the predictor, a linear combination of unknown parameter  $\alpha$  and smoothing  $f_j$ , of independent variables  $x_i$ .**

The GAM was developed as a prognostic tool for the investigation of logistic regression. GAMs replace the linear component of logistic regression (GLM) with a flexible additive function.

These models are a semi-parametric approach, an improvement with respect to the nonparametric techniques as they allow us to use, in a similar way to generalised linear models, the exact distribution that corresponds to the count data, the Binomial distribution. The importance these models have in quality control in manufacturing process is that they allow for an improvement in

the formulation of some of the nonparametric methods already described in engineering theory. For example, cubic spline graduation uses Normal approximate distribution. However, this is not necessary with the GAM approach as these models are able to use an exact distribution. Comprehensive research on GAM can be found in Hastie and Tibshirani (1990) and Hastie et al. (2001). GAM have not been used in the forecast of defective steel coil probabilities.

GAM and GLM can be applied in similar situations, but they serve different analytic purposes. GLM emphasize estimation and inference for the parameters of the model, while generalized additive models focus on exploring data nonparametrically. **GAM are more suitable for exploring the data set and visualizing the relationship between the dependent variable and the independent variables.**

### *2.3 Classification And Regression Trees (CART)*

Quality control research is frequently faced with handling complex data that include a large number of variables, which are necessary to obtain information, find patterns and identify trends. For this purpose Sonquist et al. (1971) proposed the AID program (Automatic Interaction Detection), which represents one of the earliest methods for the adjustment of data based on classification tree models. Kass (1980) proposes a recursive algorithm called non-binary classification CHAID (Chi Square Automatic Interaction Detection). Other newer methods are: FIRM (Formal Inference-based Recursive Modeling) proposed by Hawkins (1997) And MARS (Multivariate Adaptive Regression Splines) proposed by Hawkins (1997). This section focuses on the methodology which CART uses for the construction of regression trees and classification, using an algorithm recursive binary partition on each node.

The analysis of classification and regression trees (CART) is generally carried out in three steps Timofeev (2004):

- (1) construction of the maximum tree,
- (2) pruning the tree and,
- (3) optimal tree selection by cross-validation procedure

The tree is built by the following process: first the single variable is found which best splits the data into two groups, which are as homogeneous as possible, observations between groups need to be as different as possible. The “best” is defined by choosing the partition that minimizes the residual sum of squares (RSS) using a partitioning regression algorithm,

$$RSS(partition) = RSS(part_1) + RSS(part_2).$$



The data is separated, and then this process is applied separately to each sub-group, and so on recursively until the subgroups either reach a minimum size or until no improvement can be made. The resultant tree model is too complex and it must be simplified by pruning as it is over-fitting, and will not be representative for new samples. The procedure must only remove the nodes that improve the tree's precision very little. The last stage of the procedure consists of using cross-validation to select the best tree. The main differences between the algorithms for the construction of decision trees are centered on the pruning strategies and the rules adopted to remove the nodes.

A detailed description of CART can be found in Hastie et al. (2001). As in the case of the GAM, CART have not been used in the forecast of defective steel coil probabilities. The building of a decision tree starts with a description of a problem which should specify the variables, actions and logical sequence for decision-making. In a decision tree, a process leads to one or more conditions that can influence an action or other conditions, until all conditions determine a particular action, once built you have a graphical view of decision-making.

### 3 Receiver Operating Characteristic curves (ROC)

There are two objectives for the models described above:

- (1) find the factors that have predictive capacity, and
- (2) generate a predictive classification.

The first objective has already been covered by selecting a set of available variables in each one of the models. Regarding the second objective, since the response is the occurrence of an event, prognostic models may be called and must assess the predictive ability of each model. **Discrimination is a way of assessing the fit of the model through its ability to distinguish between subjects with an event and those which do not have one.** **COMENTARIO 28 DEL REFEREE DICE QUE NO SE ENTIENDE**

ROC curves are particularly useful for comparing the classification power of different fitted models. ROC curves are well described in Fawcett (2006). We consider problems where the items can only belong to two classes and some classification models (or classifiers) that produce a continuous output (e.g., an estimate of failure probabilities) to which different thresholds may be applied to predict class. For each individual we have both the model prediction and the actual class. Given a classifier and a threshold, there are four possible outcomes, Table 1 showing the possibilities

Table 1  
Results in table form

		True class	
		A	B
Predicted class	A	True positives	False positives
	B	False negatives	True negatives
		Total positives	Total negatives

The true positive rate (tp rate) of a classifier is estimated as

$$\text{tp rate} = \frac{\text{True positives}}{\text{Total positives}}.$$

The false positive (fp rate) rate of a classifier is estimated as

$$\text{fp rate} = \frac{\text{False positives}}{\text{Total negatives}}.$$

Additional terms associated with ROC curves are

$$\text{sensitivity} = \text{tp rate} = S$$

and

$$\text{specificity} = 1 - \text{fp rate} = \frac{\text{True negatives}}{\text{Total negatives}} = E.$$

ROC graphs are two-dimensional graphs in which the tp rate is plotted on the Y axis and the fp rate is plotted on the X axis. A ROC graph depicts relative tradeoffs between benefits (true positives) and cost (false positives).

If the cost of a false positive is similar to a false negative, the best model is the one with the ROC curve closest to the left and top of the chart (model without misclassification,  $S=E=1$ ). The worst is that which shows the ROC curve closest to the main diagonal of the graph (model with no predictive value). **Resulta obvio que lo ideal sería trabajar con pruebas diagnósticas de alta sensibilidad y especificidad, pero esto no siempre es posible pues dependen de la tecnología y opciones que se disponen. Then, to assess the classification power of the model, it was first verified whether the specificity and sensitivity have acceptable levels. Para las técnicas de screening o también llamadas de tamizaje se prefiere a las técnicas con gran sensibilidad y de bajo costo ya que permite estudiar un gran número de casos con menores costos y minimizando**

los riesgos. Las técnicas con alta especificidad se las prefiere como pruebas confirmatorias ya que permiten saber definitivamente si un caso es favorable o no (si es positivo con gran seguridad es positivo).

The assessment of the predictive ability of the model is defined by the area under the ROC curve (**AUC**) constructed for all possible cutoff points to classify as positive or negative events. Since the AUC is a portion of the area of the unit square, its value will always be between 0 and 1, the random guessing procedure having an area of 0.5. The greater the AUC, the better the classifier.

There are other tools that allow classification of the defective and non-defective steel coils such as discriminant analysis. However, ROC curves have the advantage of being able to decide about the breakage point or risk for the classification.

The strategies for the validation of the results of the ROC curves can be threefold:

- (1) to test the model in another different sample,
- (2) develop the model with 75 % of the sample and calculate the predictive power with the remaining 25%, or
- (3) use the same sample, but calculating predictive indicators using bootstrap techniques.

In this paper we use the last one **because we have only one sample and its size is small**. Starting from the observations, we simulate  $N$  bootstrap samples,  $n = 1, 2, \dots, N$ , where each sample is selected randomly, with replacement and with the same size as the original sample. For each bootstrap sample, the AUC is estimated, with fitted probabilities from the model adjusted with the original data. This yields  $N$  realizations of AUC for each model. Both samples were compared using a unilateral t-test for paired samples.

## 4 Application to real data

### 4.1 Industrial process under study

An overview of the industrial process is shown in Figure 1. The strip heated in the annealing furnace goes into a zinc pot which contains molten zinc. Air knives above the zinc pot blow compressed air. The strip is cooled in a cooling tower and passes a coating gauge, skin-pass mill and exit looper. In a continuous hot dip galvanizing line, the steel strip is coated by passing it

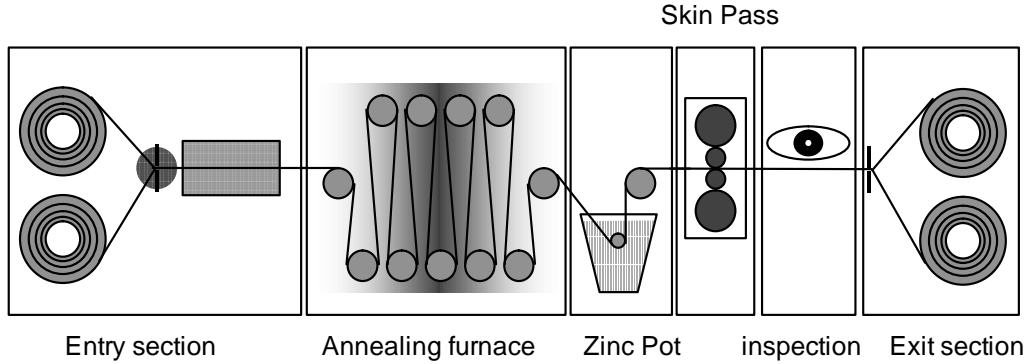


Fig. 1. Simplified model of the hot dip galvanizing line processing plant: each coil is welded in a welding machine and goes to an annealing furnace; the strip heated goes a zinc pot which contains molten zinc (galvanizing bath operation), the strip is cooled and passes a coating gauge, skin-pass mill, quality control inspection, and exit looper.

through a pot of molten zinc normally between 450 and 480 °C (average bath temperature of 460°C) and withdrawing it from the pot through a pair of air-wiping jets, to remove the excess liquid zinc, as shown schematically in Figure 2. Galvannealing is an in-line process during which the zinc layer of the steel strip is transformed into a iron-zinc alloy (ZnFe) layer by diffusion. Depending on the annealing temperature, annealing time, steel grade, aluminium content in the Zn bath, and other parameters, different intermetallic ZnFe phases are formed in the coating, which influence the formability of the galvanized material (Tang, 1999). **The material consumption is compensated by the ingots of Zn-Al alloys which are added to the bath to maintain a constant bath Zn level and Al content. Local decreases in the temperature near the melting ingot surface reduce the solubility of iron and aluminum, and are considered to be the primary source of dross particles. Control of bath temperature is therefore of primary importance in order to prevent this precipitation phenomena. The presence of an ingot significantly changes the temperature distribution and also results in important variations in the local aluminum concentration, since the makeup ingot has a higher aluminum concentration. It was shown that precipitates form near the ingot surface because this region is surrounded by a solution with a temperature which is lower than the average bath temperature of 460 °C (Ilinca et al., 2007). The temperature of the molten zinc coating bath must also be carefully controlled to avoid excessive temperature variations of steel strips in the inlet of the bath which could cause excess alloy layer formation in the bath on the heavier zinc coated side.**

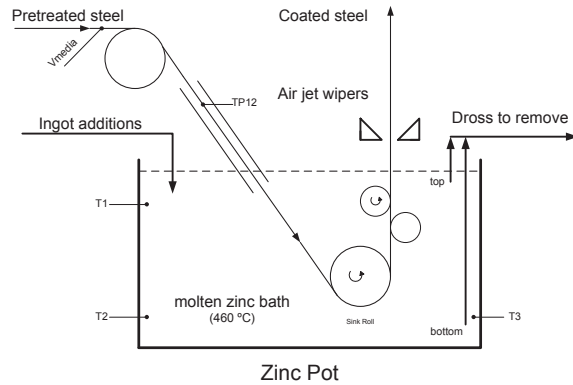


Fig. 2. Galvannealing section (zinc pot) of a hot dip galvanizing line. Schema of galvanizing bath operation and variables location used to quality and process control: the input strip temperature of the zinc pot zone (TP12), the input strip velocity ( $V_{media}$ ), and temperature of the zinc bath (T1, T2 and T3).

#### 4.2 Data

Data for 48 batches (steel coils) were available. Five variables were selected for monitoring the process: the steel strip velocity ( $V_{media}$ ), the input strip temperature of the zinc pot zone (TP12) and three bath temperatures (T1, T2, T3). The data were analyzed to determine whether or not a fault occurred in the process. The entire data consisting of 48 galvanized steel coils was divided into sets. The first training data set was 25 conforming coils and the second data set was 23 nonconforming coils.

#### 4.3 Generalized linear models

As is common with dichotomous data, the estimation of the covariable effects are analyzed on the probability associated with a defective steel coil. We begin with the Generalised linear Model analysis which allows us to analyze the data on the assumption that the number of defective coils is Binomial. Its mathematical expression is in equation (1). Our model is formally a Binomial Generalized linear Model with several link functions, it was fitted in R (R Devel-

opment Core Team, 2009). The most commonly used GLM regressions can be obtained using R command `glm` except negative binomial regression which can be obtained using `glm.nb` into **MASS** library. Detailed arguments and estimations are available in Venables and Ripley (2002). But the recently GLM with Conway-Maxwell Poisson distribution can be obtained using library **COMpoissonReg** (Sellers and Lotze, 2011).

The results of the GLM Regression are shown in Table 2, which includes the value of the Deviance for each of the Binomial model with corresponding links. There are many commonly used link functions, and their choice can be somewhat arbitrary. The Binomial family links are `logit`, `probit`, `cauchit`, (corresponding to logistic, normal and Cauchy CDFs respectively) `log` and `cloglog` (complementary log-log), *Deviance* is a quality of fit statistic for a model, it is calculated 2 times the log-likelihood of the full model minus 2 times the log-likelihood of the model and  $\nabla Deviance$  es el incremento de Deviance de un modelo respecto del anterior. To explore the changes in the number of terms in a model we consider the Mallows statistic  $C_p$  which penalizes the complexity of models because it increases with the number of parameters. Then, we must choose the model with the lowest Deviance and  $C_p$  (see formulas and details in McCullagh and Nelder (1989) and Venables and Ripley (2002)). From these results we can conclude that although there are no major differences in link, the best results are for the clog-log link fit. Therefore this link was used for the regression.

Table 2  
GLM Regression

Link	Deviance	$\nabla Deviance$	$C_p$
cauchit	50.803		62.803
logit	49.956	0.847	61.956
probit	49.705	0.252	61.705
cloglog	48.946	0.759	60.946

The results of the GLM Regression with cloglog link are shown in Table 3, which include the value of the coefficients  $\beta$  for each of the covariates, the standard error  $se(\beta)$  and the p-values or significance associated for each of the coefficients.

**According with their p-values ( $\leq 0.1$ ) only Vmedia and TP12 are significant variables.** The meaningfulness of the positive coefficient  $\beta$  which corresponds to the Vmedia can be interpreted thus: the failure probability increases with the speed. On the contrary, the meaningfulness of the negative coefficient which corresponds to TP12 can be interpreted thus: the failure

Table 3  
GLM Regression

Covariables	$\beta$	$se(\beta)$	t-value	p-value
Intercept	-150.87	-992.70	-0.152	0.8792
Vmedia	0.1154	0.0409	2.822	0.0048
TP12	-0.6929	0.2928	-2.366	0.0180
T1	0.8749	0.7067	1.238	0.2157
T2	-1.3195	0.8663	-1.523	0.1277
T3	1.4189	2.3570	0.602	0.5472

probability decreases when the temperature increases.

The next step is to select the most relevant covariables for adjustment. The R command `step` will be used to select a formula-based model by the Akaike information criterion (AIC is a measure of the relative goodness of fit of a statistical model and it provides a means for comparison among models, the preferred model is the one with the minimum AIC value). This command takes all the possible selection of variables into account and chooses the minimum number of significant variables that fit best.

Then, we can check model assumptions and fit using its residuals. The graphs of the residuals may indicate a problem in the choice of link function, a wrong scale of one of the linear predictors, omission of a quadratic term predictor or heteroscedasticity. In addition, high levels of residuals reported outliers. In this case, it is necessary to eliminate some of these observations and reformulate the model. Deviance residuals are usually used for these purposes, and graphs are to represent residuals versus fitted values and against the predicted values. The first standard graphs are shown in Figure 3 using R command `plot`.

Figure 3 shows in the upper left panel the plot of the deviance residuals versus the predicted values where there are two groups, the variance does not remain more or less constant with the average, i.e. the dispersion of residuals fluctuates around zero. So you have two different groups for the fitted values corresponding to failures and non failures. In the top right panel there is a Q-Q plot of standardized deviance residuals. If these residuals are adjusted perfectly to the diagonal, their distribution is exactly normal. Especially at the ends, our residuals are separated from the line, which means that the distribution of these residuals has thicker tails than the Normal distribution. The graph in the lower left panel is the representation of the square root of the absolute value of the deviance residuals versus predicted values. Curvatures will indicate the absence of a quadratic term or a bad choice of link function. In this case, it seems that the data does not reasonably fit a straight line because there are some points that are more spread **for that reason we need to**

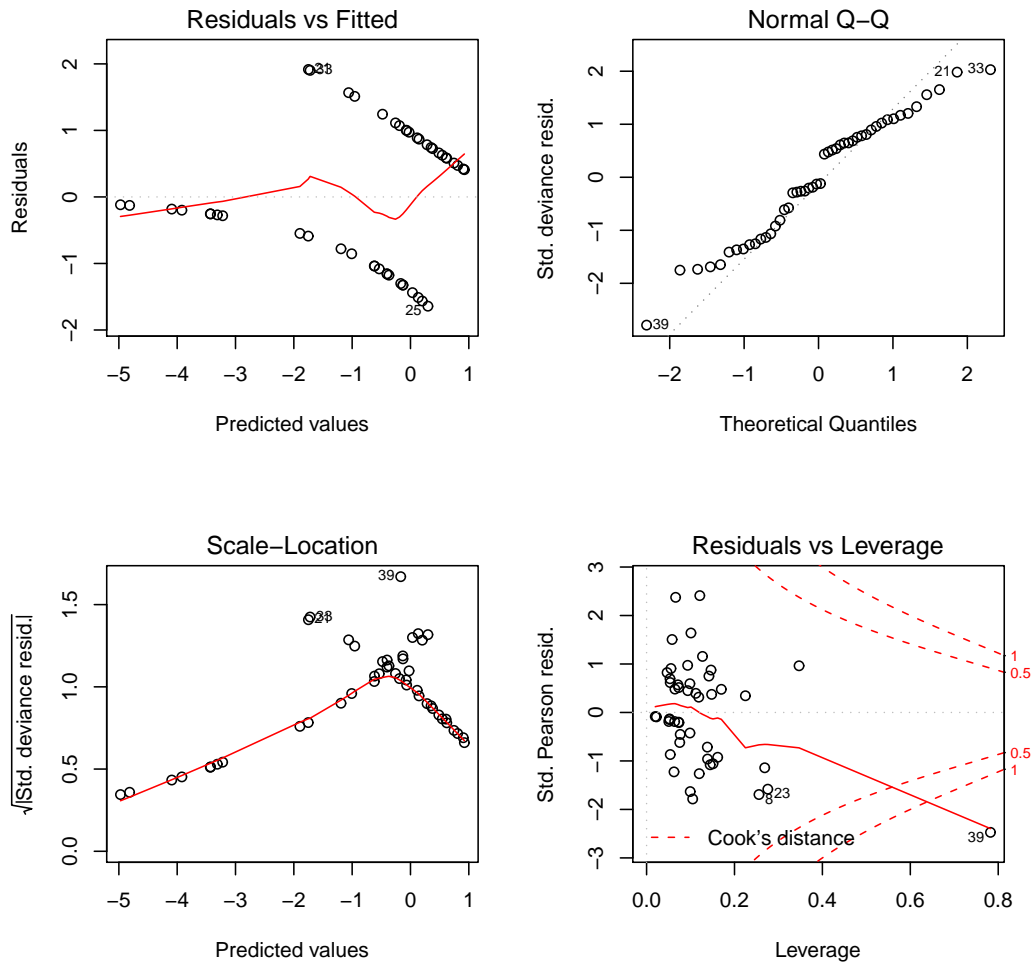


Fig. 3. Checking the assumptions for a proper GLM model.

**check the linear relationship using the GAM model in next section.**

The last graph, located in the lower right panel, determines the influential points using Cook's distances, which is a measure of the difference between the fitted values of the model and the model without each of the observations. If this statistic is very large, we face an outlier, this observation can be seen in 39, so we repeated the analysis without this observation. This is carried out to improve the model so that it considers the behavior of the majority of the data better, and is not too influenced by any one data point.

Although the selection of the link remains the same, the significance of the variables have changed, the results appear in Table 4 and their validation is shown in Figure 4 which shows that the residuals have improved their behavior significantly. In Table 4 the values listed above the double line are model coefficients including all variables and below the double line are when the non-significant variables are omitted ( $p\text{-value} \geq 0.1$ ).



Table 4  
GLM Regression without influential points

Covariables	$\beta$	$se(\beta)$	t-value	p-value
Intercept	64.47002	1052.35184	0.061	0.95115
Vmedia	0.10160	0.03296	3.082	0.00205
TP12	-1.50223	0.50878	-2.953	0.00315
T1	1.72442	0.93005	1.854	0.06372
T2	-2.42899	1.11397	-2.180	0.02922
T3	2.01210	2.63032	0.765	0.44429
Intercept	853.6715	433.5346	1.969	0.04894
Vmedia	0.1003	0.0331	3.028	0.00246
TP12	-1.3539	0.4334	-3.124	0.00179
T1	1.6650	0.9169	1.816	0.06939
T2	-2.2241	1.0352	-2.149	0.03167

Examination of the partial residual plots is also interesting. These graphs help us to understand the relationship between the response and each of the explanatory significant variables of the model. They also help identify if there are problems such as lack of linearity or constant variance and which variables cause them. The partial residual graphs were obtained in R using the faraway library (Faraway, 2009).

Figure 5 shows all independent variables versus their respective partial residuals in the GLM model. The effects of the two variables in the model are important because there is enough variation in their values. The variables in the model appear to form two distinct groups, defective coils (filled square) and non-defective (circle), with a greater linearity in non-defective. **These graphs (Figure 5) show that the model fits the data relatively well.**

**In summary, the result (Table 4 above the double line) provides us with GLM is a logistic regression model where the T3 temperature do not influence the production of defective coils (p-value  $\geq 0.1$ ). The meaningfulness of the positive coefficient  $\beta$  which corresponds to the Vmedia and T1 can be interpreted thus: the failure probability increases with the speed. On the contrary, the meaningfulness of the negative coefficient which corresponds to TP12 and T2 can be interpreted thus: the failure probability decreases when the temperature increases.**

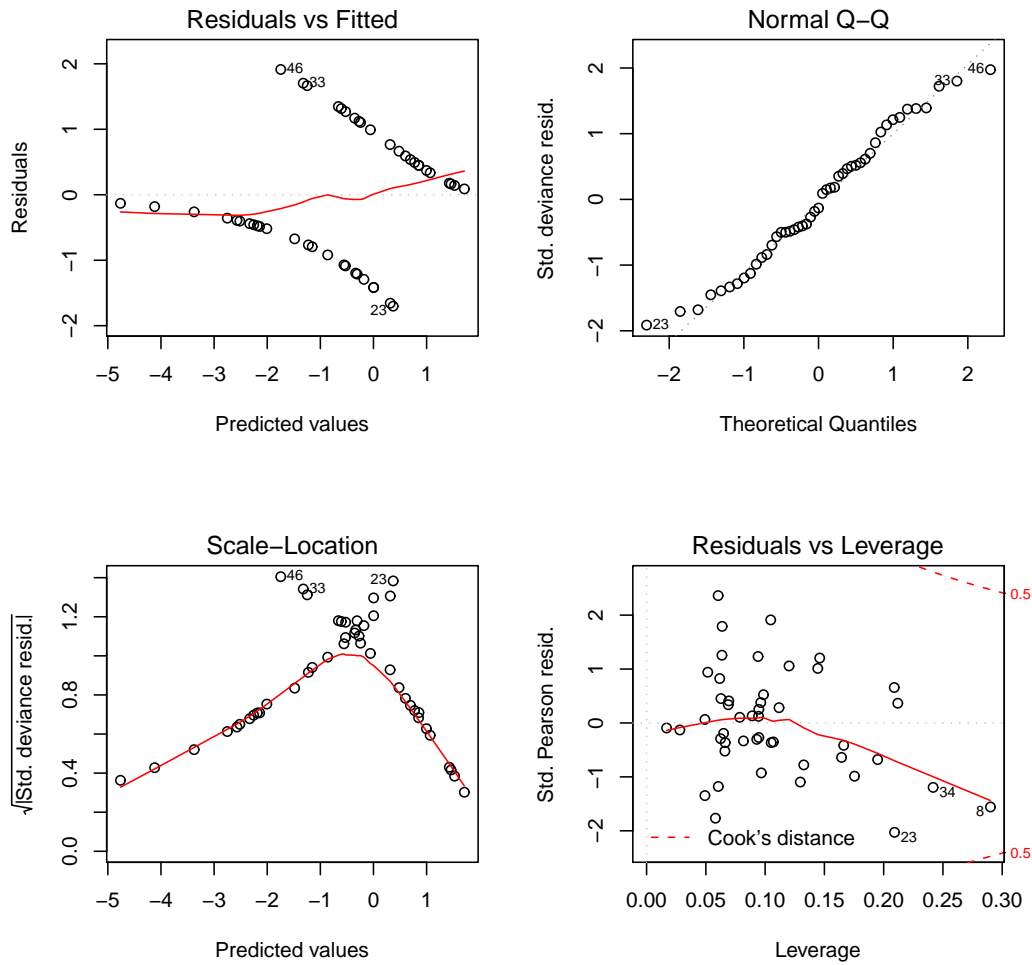


Fig. 4. Checking the assumptions for a GLM model without influential points.

#### 4.4 Generalized additive models

In order to check the consistency of these previous results, we use a different technique: Generalised Additive Model (GAM) analysis which allows us to analyze the data on the assumption that the number of defective steel coils is Binomial. They allow to us to incorporate explanatory variables in a non-linear way in the model. Our model is formally a GAM with Binomial family and cloglog link, it was fitted in R. The detailed arguments and estimations are available in Faraway (2006).

Results of the Generalised Additive Model are shown in Table 5 **with T3 non-significant variable (p-value  $\geq 0.1$ )**. The first graphics standard, which are shown in Figure 6 the partial predictions for each predictor, were performed with command `plot`. **Notice that the 95% confidence limits**

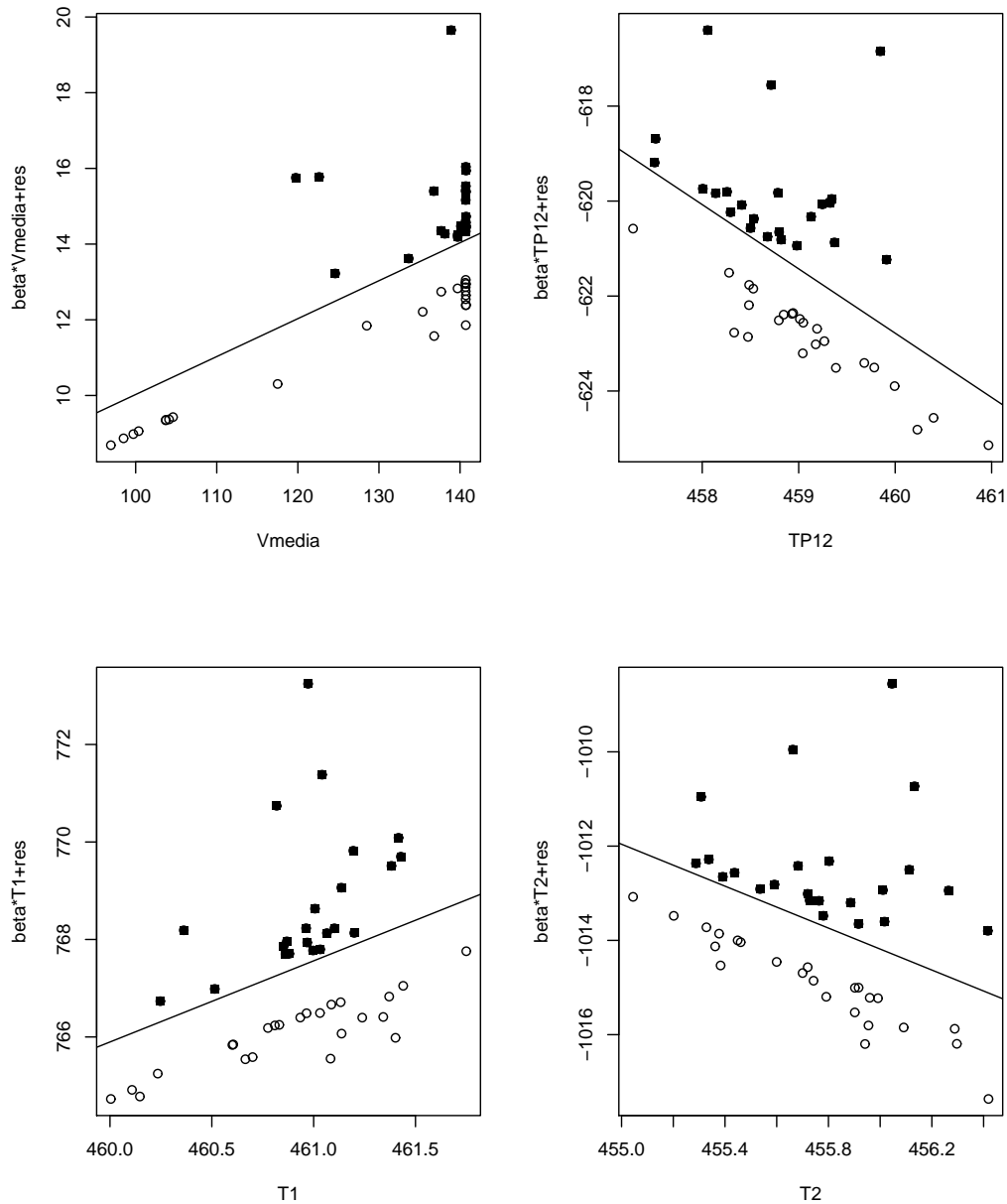


Fig. 5. Partial residuals of variables in the GLM model.

for  $T3$  cover the zero axis, confirming the insignificant of this term. The plot also shows that the partial predictions corresponding to rest of variables that have a linear pattern.

Influential points are determined with Cook's distances through the `cooks.distance` command that we used in the previous model, these distances are shown in Figure 7. It can be seen that observation 39 corresponds to a high value, so as before, we repeated the analysis without this observation.

Table 5  
GAM Regression

Covariables	edf	Ref.df	Chi.sq	p-value
s(Vmedia)	2.270	2.270	7.874	0.0257
s(TP12)	1.000	1.000	5.290	0.0214
s(T1)	1.592	1.592	5.611	0.0397
s(T2)	1.000	1.000	4.456	0.0348
s(T3)	1.000	1.000	0.516	0.4724

The significant variables have not changed, the results appear in Table 6 and their validation is shown in Figure 8 where **smooth functions is written as  $s(x, df)$ ,  $x$  is the univariate predictor, and  $df$  is the target equivalent degrees of freedom, used as a smoothing parameter (values for  $df$  should be greater than 1, with  $df=1$  implying a linear fit). This Figure 8 suggest a linear relation for all variables excepto para  $Vmedia$  para la que sugiere una curva muy suave por lo que pensamos que se confirma la linealidad que supone the GLM model.**

Table 6  
GAM Regression without influential points

Covariables	edf	Ref.df	Chi.sq	p-value
s(Vmedia)	4.142	4.142	1.005	0.91924
s(TP12)	1.000	1.000	8.698	0.00319
s(T1)	1.151	1.151	4.443	0.04360
s(T2)	1.000	1.000	6.066	0.01378
s(T3)	1.000	1.000	0.750	0.38637

#### 4.5 Classification and regression trees

Tree models are quite popular because the structure is easier for non technical people to understand. The following packages are available from the CRAN (R Development Core Team, 2009), **rpart** Recursive PARTitioning and regression trees, and **tree** Classification and regression trees. Trees can be used for several different types of response data. Our model is formally a classification tree, it was fitted in R with **rpart** function of **rpart**. The Detailed arguments and estimations are available in Faraway (2006). **Figure 9 shows the resulting tree, the interpretation of this tree is straightforward. The top circle (root node) in Figure 9 contains all steel coils (25 conforming / 22 nonconforming). This node is split based on the value of**

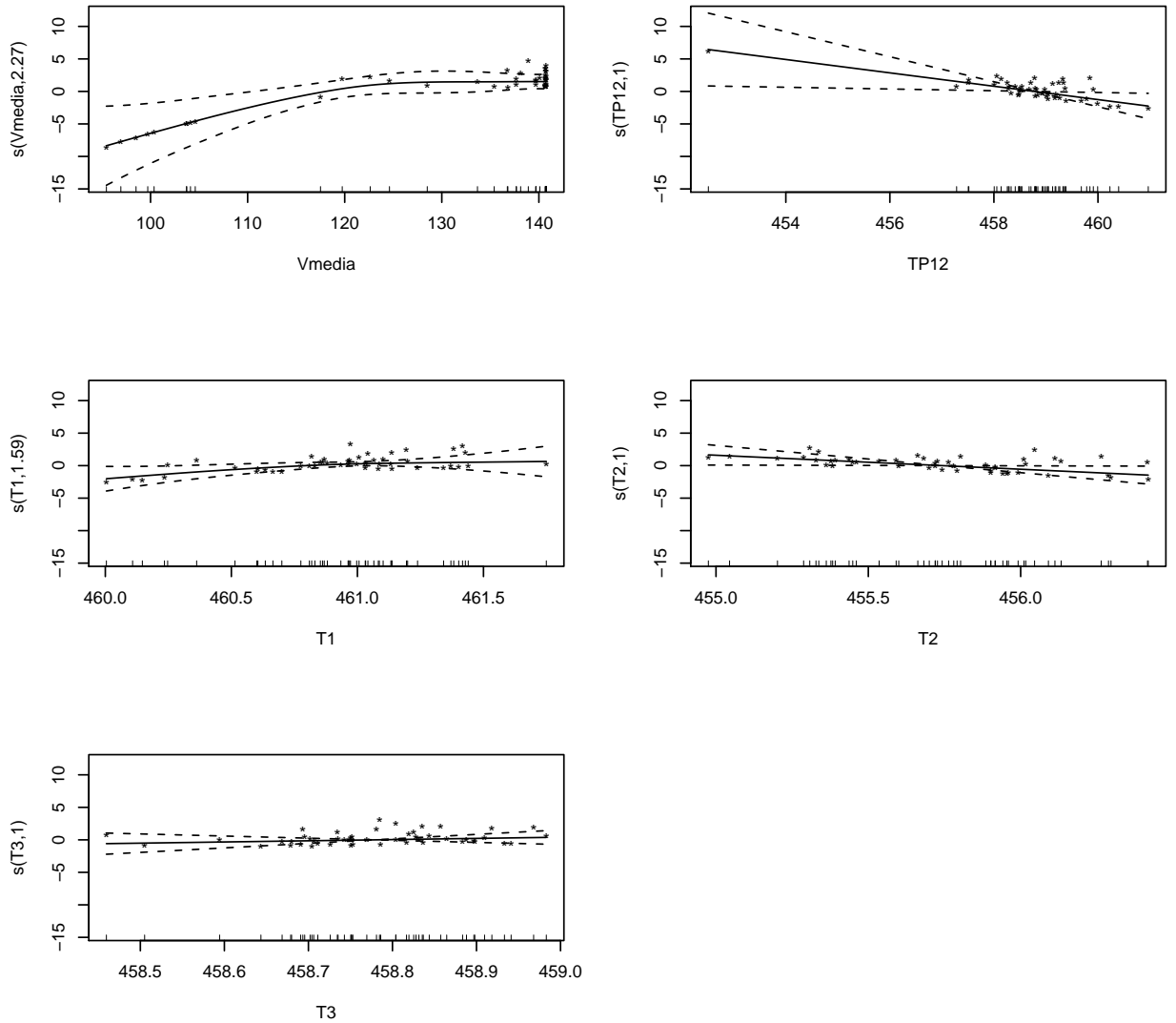


Fig. 6. Plots of the component smooth functions and partial residuals on the scale of the linear predictor

principal separation variable. The first ramification point represents the first risk factor selected by the procedure. When a split occurs, the subsamples, also called nodes, end up either in a circle or in a rectangular box. The rectangular boxes are referred to as terminal nodes and the circles are non-terminal nodes. Terminal nodes do not split further, while non-terminal nodes do. From here on, node will be used instead of subsample. The information available at each node is explained below.

- “0, 1”: category of the dependent variable (0 for conforming steel coil and 1 for nonconforming steel coil).

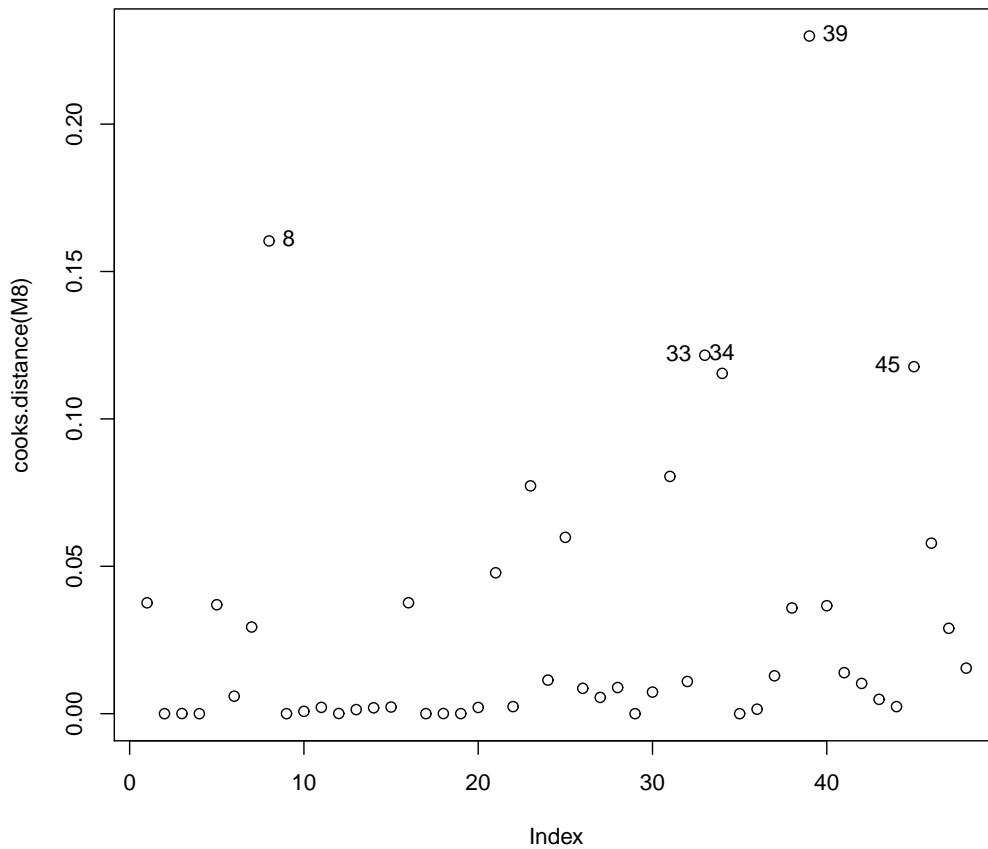


Fig. 7. Cook distances for the GAM model

- “25/22”: number of objects for each category or class of the dependent variable (0/1).
- “ $V_{media} \geq 118.6m/min$ ”: split or separation variable with the cut-off point.
- and so on for all the ramifications.

We force it to consider a larger tree and then examine the cross-validation criterion for all the trees using the command `printcp` which only displays the results in the console and `plotcp` which visualize cross-validation results as a graph in Figure 10. **Figure 10 gives a visual representation of cross-validation results, representing  $C_p$  statistics versus  $xerror$ , a good choice of  $C_p$  for pruning is often the leftmost value for which the mean lies below the horizontal line.**

We can select the size of the tree by minimizing the  $xerror$  value in the output of the command `printcp` and selecting the corresponding value of  $C_p$  with the command `prune.rpart`, **Figure 11 shows the selected tree to avoid over-**

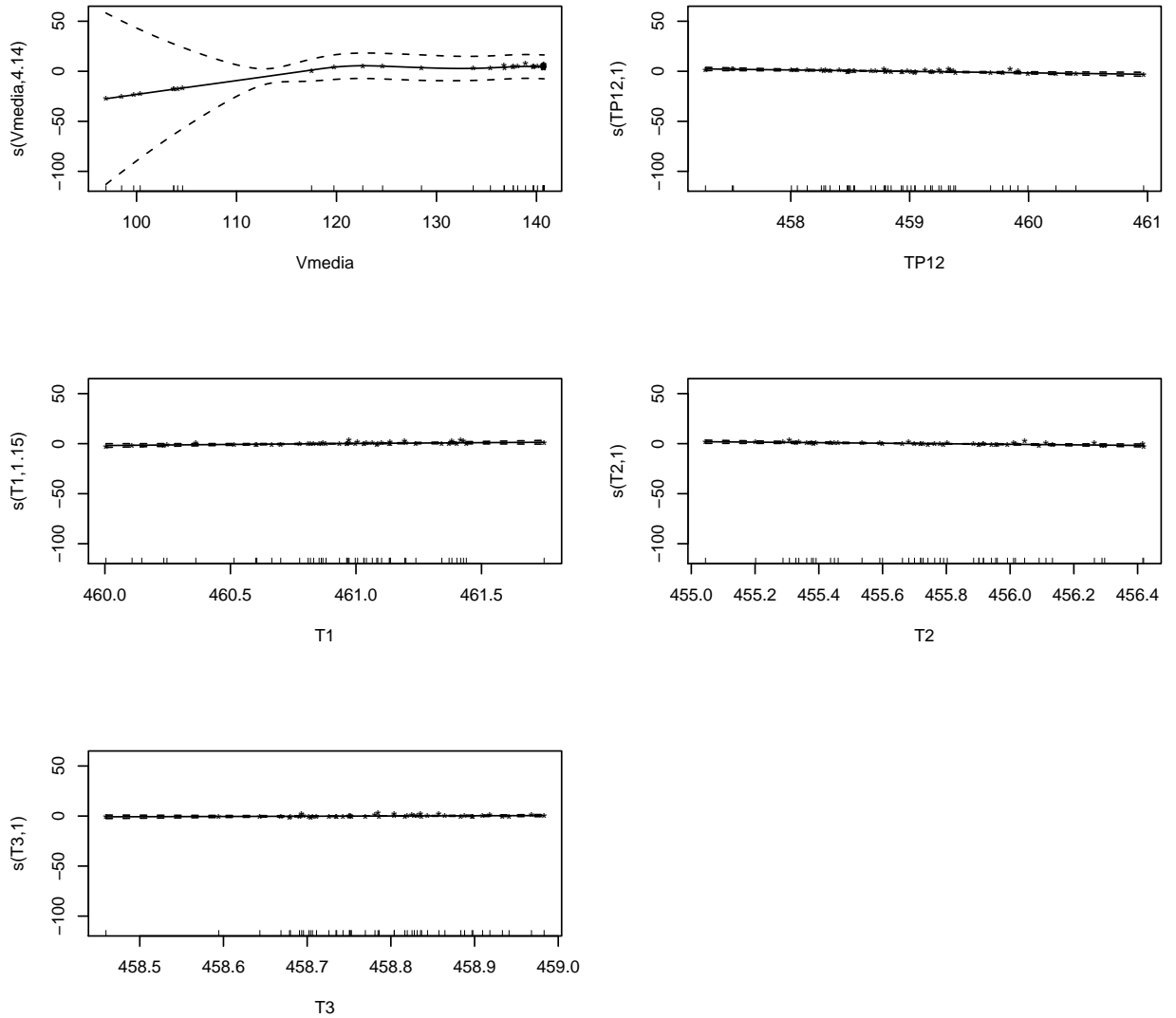


Fig. 8. Plots of the component smooth functions and partial residuals on the scale of the linear predictor

**dispersion based on the complexity parameter  $C_p$ .**

This tree is not particularly successful because the maximum tree has large “deviance residuals” (**these residuals are examined in a plot not shown here because is similar to other figures yet explained**), especially for observation 5 and the resulting pruned tree only uses one independent variable. We must point out that when fitting a tree, it is very important not to over-fit or under-fit it, because if the tree is under-fitted, then it will not be flexible enough and may overlook the important structure in the data. We propose repeating the process obtaining the clasification tree adjusted to the data without observation 5. The final tree obtained is shown in Figure 12.

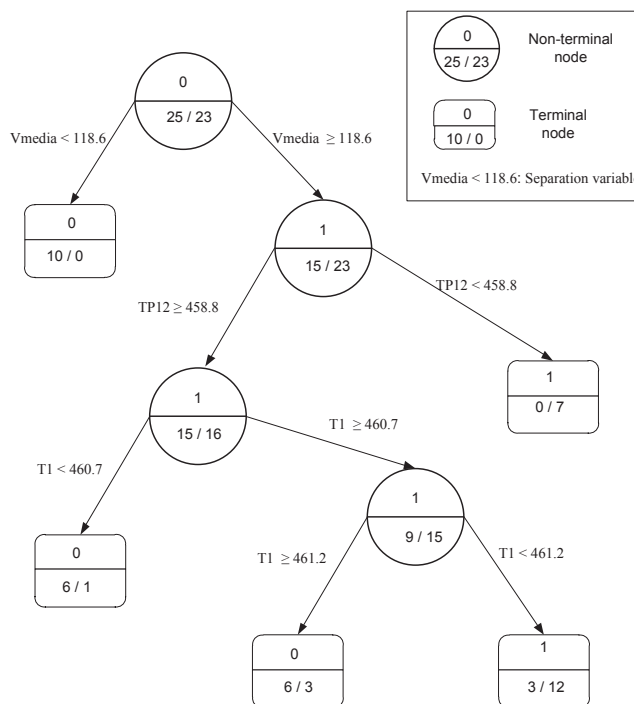


Fig. 9. Maximum tree model for the quality performance steel coils. The cases fall into two classes: "0" for conforming and "1" for nonconforming steel coil. The first ramification point (split) represents the first risk factor selected by the procedure. The values in each node are the prediction for the node (0 or 1) and the numbers  $m/n$  denotes the proportion of training cases reaching that node classified into each class (0/1).

Figure 12 shows that the conditions to obtain non defective coils are maintaining the velocity below 118.6, but if that speed is exceeded, maintaining the TP12 above 458.8, once built you have a graphical view of decision-making. **The results of final CART analysis using binary conforming quality as the outcome variable are presented in Figure 12. For Vmedia the classification tree analysis yielded a split point (threshold) of 118.6 m/min. This produced two subgroups with respective quality incidences of 10 conforming coils and zero nonconforming coils if  $V_{media} < 118.6$  m/min, and 15/22 conforming/nonconforming coils if  $V_{media} \geq 118.6$  m/min. This latter subgroup was further partitioned on the basis of the data for quality coil decline in TP12,**



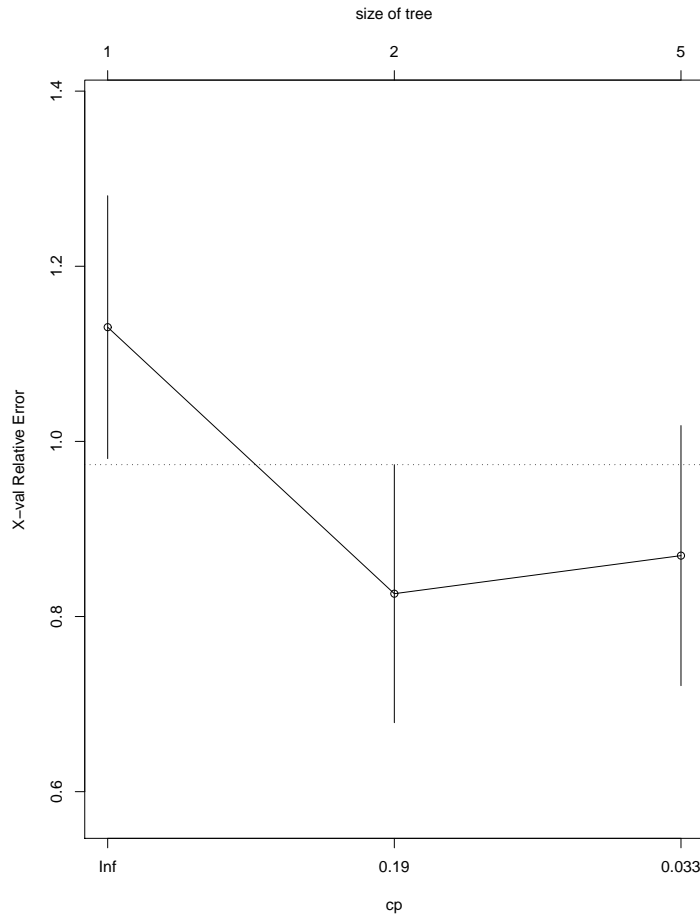


Fig. 10. Cross-validation plot for assessing predictive accuracy and for choosing the optimal size of tree. Change in cross-validation error (xerror) as the size of tree change. The tree size that minimizes the cross-validation error is selected. The tree should be pruned to include only 1 or 2 splits.

and the classification tree analysis yielded a split point (threshold) of  $458.8\text{ }^{\circ}\text{C}$ . The resultant groups had quality incidences of 11 conforming coils and seven nonconforming coils if  $\text{TP12} \geq 458.8\text{ }^{\circ}\text{C}$ , and 15/22 conforming/nonconforming coils if  $\text{TP12} < 458.8\text{ }^{\circ}\text{C}$ . The result provides us with CART is that the T1, T2 and T3 temperatures do not influence the production of defective coils. This is due to variations in zinc bath temperature measured by T1, T2 and T3 were very small, fluctuating around the target temperature of  $460\text{ }^{\circ}\text{C}$ . In this case, although the three variables are not influence in the nonconforming coils production, it is not unreasonable to conclude that the bath temperature is not important. In a modern galvanizing plant the automatic temperature control of the zinc bath is efficient enough to maintain the bath in its set point  $\approx 460\text{ }^{\circ}\text{C}$ . The analysis described above showed that steel coil which was less prone to be

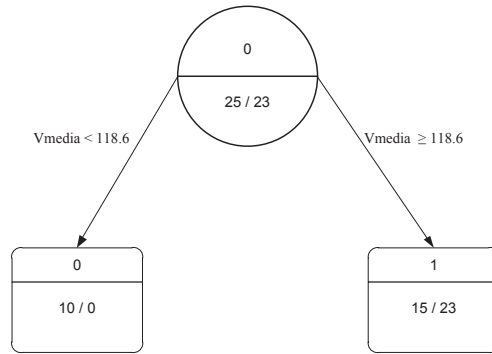


Fig. 11. Pruned Classification Tree of size 2 terminal nodes (1 split: Vmedia) for the quality performance steel coils based on minimal cross-validation.

defective is produced under the following conditions: low Vmedia, high TP12, and maintaining the zinc bath temperature close to the  $460^{\circ}C$  by monitoring of T1, T2 and T3 temperatures. The study has sought to provide insight into the impact of different variables on the risk of defect in steel coil. Our results confirm and extend previous studies (García-Díaz, 2009b) in which the influence of process variables on nonconforming coils production has been found to be associated with increased strip velocity (Vmedia) and with the input strip temperature of the zinc pot zone (TP12).

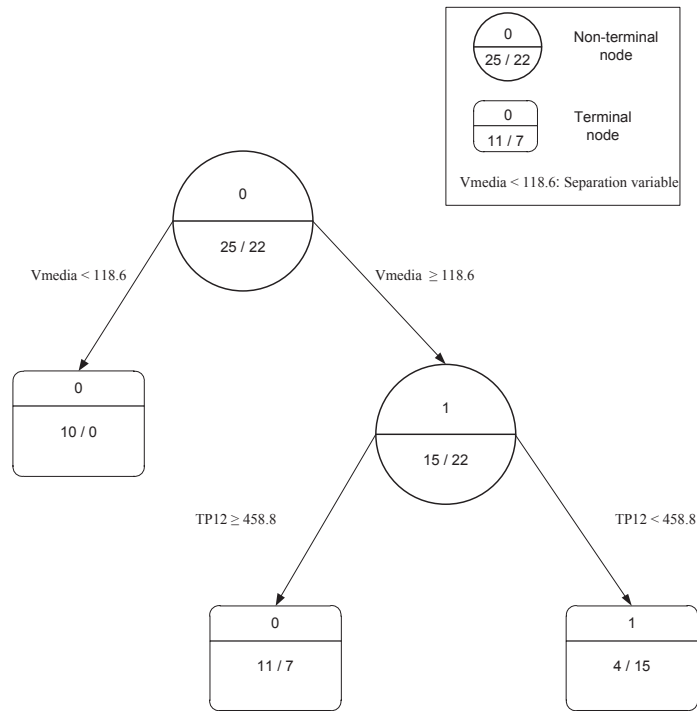


Fig. 12. Final decision tree model of size 3 terminal nodes (2 split) for the quality performance steel coils. The first ramification point (split) represents the first risk factor ( $V_{media}$ ) selected by the procedure, and the second split represents the second risk factor ( $TP12$ ). The values in each node are the prediction for the node (0 or 1).

#### 4.6 ROC curve

The above models are designed to show the impact of each variable in the defective steel coil probabilities. To evaluate the discriminate performance of the models and to compare them we can draw the ROC curve and calculate the AUC. They are typically used for any test in which there are two possible outcomes that are of interest to the researcher. Does the test give a correct positive result (A) or a correct negative result (B). In this case, it corresponds to the fact that either the steel coil is actually defective (A) or not (B). A ROC curve is considered to be a summary measurement in the sense that it uses all possible decision thresholds for the test to create the curve.

The ROC curves were obtained in R using the ROCR library (Sing et al., 2007). When comparing the ROC curves for the two models, tests with larger areas are typically considered to be better tests, it is possible for tests to have the same area and yet not be considered to be equivalent. Thus, it is important to compare ROC curves and their areas.

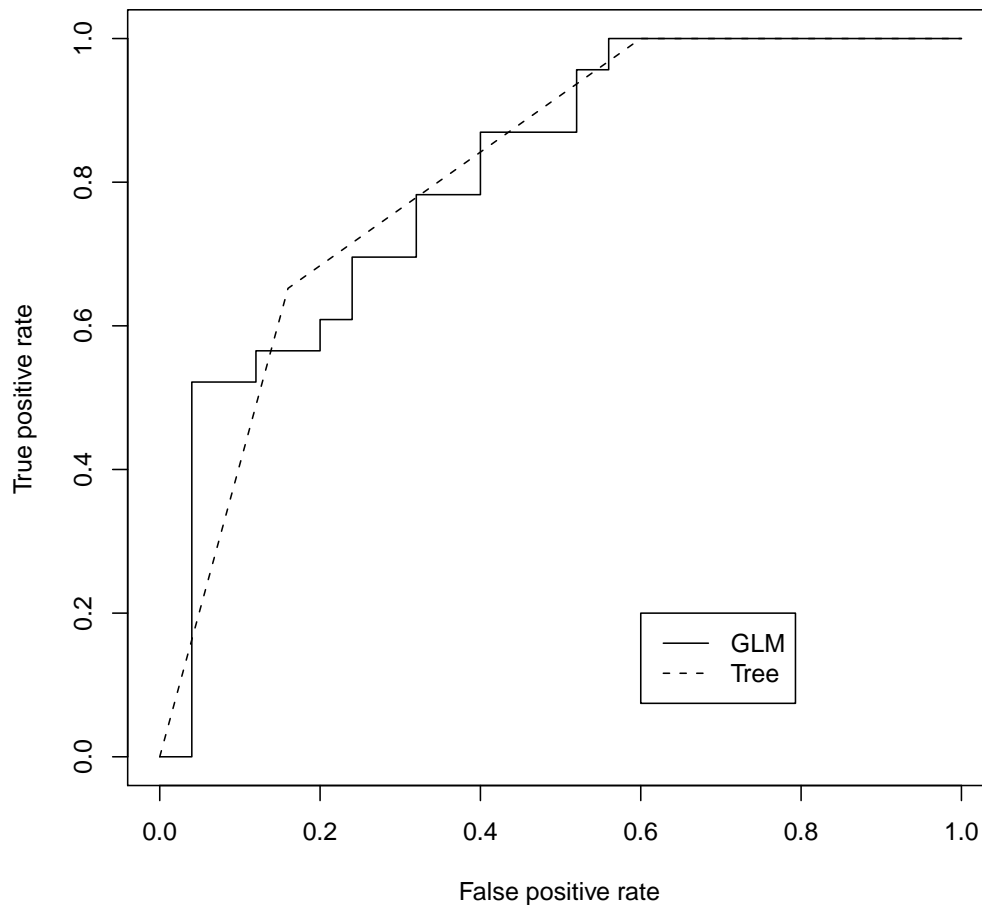


Fig. 13. ROC curves.

Figure 13 shows ROC curves based on the corresponding models whose comparison allows us to say that the defective probabilities obtained by CART discriminates better between the steel coils that are defective than those that are not, because its curve is **almost** always above the corresponding GLM model curve.

In addition, the AUC corresponding to CART is 0.8156 greater than the AUC corresponding to GLM of 0.8122. This confirms the superiority of the CART in our data. It should be taken into account that although the models were adjusted without the outliers, they were validated for all the data. Likewise,

Figure 13 shows the ROC curve of each model. To verify that the CART model provides ROC curves with greater AUC and can be considered better than the GLM for different data sets we used the bootstrap techniques as described earlier. The result of the paired t-test  $H_0 : AUC_{GLM} = AUC_{CART}$  versus  $H_0 : AUC_{GLM} < AUC_{CART}$  have t-value=-3.606 with p-value = 0.0003263(two-tailed), 0.0001632 (one-tailed), rejecting  $H_0$  showing that there is enough evidence to admit that the CART model is better than GLM.

From an economic viewpoint, quality costs have emerged as a financial control tool for management and as an aid in identifying opportunities for reducing quality cost. The idea is to find the classifier that is capable of predicting the quality measure of product based on its manufacturing parameters. CART replaces equation of usual regression type model with set of rules. This future is an important aid for the process engineer as it allows him to change the operating parameters as the line speed in order to produce high-quality galvanized steel economically. Useful rules for a process engineer are explained below. The CART analysis showed that steel coil which was less prone to be defective is produced under the following conditions: low Vmedia (Vmedia < 118.6 m/min), high TP12 (TP12  $\geq$  458.8 °C), and maintaining the zinc bath temperature close to the 460 °C by monitoring of T1, T2 and T3 temperatures.

## 5 Conclusions

Statistical analysis for estimating the quality of galvanized steel coils on the basis of short time histories of the velocity and temperature distribution of a typical galvanizing bath have been carried out using statistical regression models: generalized linear models, generalized additive models and classification trees. We have compared two different models by choosing the best fits for each one of them. **Specifically, we have compared the GLM and CART in Figure 13, not GAM which is only used to confirm the linear relationship.** The comparison is carried out by applying the ROC curve, from which we can conclude that the GLM model produces a worse **classification of galvanized steel in defective or not than CART.** Then, the methods are tested in terms of discriminative ability. According with the results of the paired test applied to the corresponding AUC to the GLM and CART model the difference between them is statistically significant. In addition, the CART model was much easier to use and interpret, because the decision rules generated could be applied without the need for mathematical calculations (Figure 12).

The study sought to provide insight into the impact of operating variables such as line speed and temperature distribution in the bath on the risk of defect in steel coils. The results show that it is essential to include the steel strip velocity and temperature near the inlet of the bath to forecast the risk of nonconforming steel coil. The temperature of the molten zinc coating bath must also be carefully controlled to avoid excessive temperature variations in the steel strips at the inlet of the bath which could cause excess alloy layer formation in the bath on the heavier zinc coated side. The presence of an ingot significantly changes the temperature distribution and also results in important variations in the local aluminum concentration, since the makeup ingot has a higher aluminum concentration. It was shown that precipitates form near the ingot surface because this region is surrounded by a solution with a temperature which is lower than the average bath temperature of 460 °C.

The results show that the CART do provide good estimates of quality coils and can be useful for quality control in manufacturing process. From this model (Fig. 12) we can conclude that steel coil which were less prone to be defective are produced under the following conditions:  $V_{media}$  less than 118.5 and TP12 greater than 458.8. **Our results confirm and extend previous studies in which the influence of process variables on nonconforming coils production has been found to be associated with increased strip velocity and with the input strip temperature of the zinc pot zone. The CART model developed in the present paper shows that the nonconforming steel coil incidence was best described by an interaction between strip velocity and input strip temperature of the zinc pot zone. The model indicates that nonconforming steel coil incidence is likely to increase in conditions above 118.6 m/min strip velocity and below 458.8 °C input strip temperature of the zinc pot zone. However, when input strip temperature is above 458.8 °C, this interaction disappears and there appears to be little effect on nonconforming steel coil incidence. The result provides us with CART is that the bath temperatures (T1, T2 and T3 variables) do not influence the production of defective coils. This is due to variations in zinc bath temperature were very small, fluctuating around the target temperature of 460 °C. The analysis described above showed that steel coil which was less prone to be defective is produced under the following conditions: low  $V_{media}$ , high TP12, and maintaining the zinc bath temperature close to the 460 °C by monitoring of T1, T2 and T3 temperatures.**

In relation to the work of other authors, we should highlight one distinctive feature of the methodology presented here, which is the possibility of comparing the different models with a simple and objective criterion. We are able to provide measures of goodness of fit statistics and a detailed study of residuals

with different graphs.

In short, we propose statistical tools which provide a clear framework for decision support in quality control in manufacturing processes.

## Acknowledgments

The authors are indebted to the anonymous referees whose suggestions improved the original manuscript.

## References

- Agresti, A., 2002. *Categorical Data Analysis*. Wiley, 2<sup>a</sup> Edición.
- Aparisi, F., García-Díaz, J. C., 2007. Design and optimization of ewma control charts for in-control, indifference, and out-of-control regions. *Computers and Operations Research* 34 (7), 2096 – 2108.
- Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. *Classification and Regression Trees*. Wadsworth International Group, Belmont, California, USA.
- Debón, A., Carrión, A., Cabrera, E., Solano, H., 2010. Comparing risk of failure models in water supply networks using roc curves. *Reliability Engineering & System Safety* 95 (1), 43 – 48.
- El Aroui, M., Lavergne, C., 1996. Generalized linear models in software reliability: parametric and semi-parametric approaches. *Reliability, IEEE Transactions on* 45 (3), 463–470.
- Faraway, J., 2006. *Extending The Linear Model With R: Generalized Linear, Mixed Effects And Nonparametric Regression Models*. Chapman & Hall/crc.
- Faraway, J., 2009. *faraway: Functions and datasets for books by Julian Faraway*. R package version 1.0.4.  
URL <http://www.maths.bath.ac.uk/~jjf23/>
- Fawcett, T., 2006. An introduction to ROC analysis. *Pattern Recognition Letters* 27.
- García-Díaz, J. C., 2009a. *Safety, Reliability and Risk Analysis: Theory, Methods and Applications*. Vol. 1. Taylor & Francis Group, London, Ch. Fault detection and diagnosis in monitoring a hot dip galvanizing line using multivariate statistical process control., pp. 201–204.
- García-Díaz, J. C., 2009b. Statistical methods for quality control of steel coils manufacturing process using glm models. In: *AIP Conference Proceedings*. Vol. 1181. American Institute of Physics, New York., pp. 299–305.
- Guh, R.-S., August 2005. A hybrid learning-based model for on-line detection

- and analysis of control chart patterns. *Computers & Industrial Engineering* 49, 35–62.
- Hamada, M., Nelder, J., 1997. Generalized linear models for quality-improvement experiments. *Journal of Quality Technology* 29 (1), 292–304.
- Hastie, T., Tibshirani, R., Friedman, J., 2001. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. Springer, New York.
- Hastie, T. J., Tibshirani, R., 1990. *Generalized Additive Models*. Chapman-Hall, London.
- Hawkins, D., 1997. FIRM: formal inference-based recursive modeling. Technical Report 546, University of Minnesota, School of Statistics.
- Himmelbau, O., 1978. *Fault Detection and Diagnosis in Chemical and Petrochemical Process*. Elsevier, Amsterdam.
- Hosmer, D., Lemeshow, S., 2000. *Applied Logistic Regression*. 2nd Edition. Wiley.
- Ilinca, F., Ajersch, F., Baril, C. and Goodwin, F., 2007. Numerical simulation of the galvanizing process during ga to gi transition. *International Journal for Numerical Methods in Fluids* 53, 1929–1646.
- Kass, G., 1980. An exploratory technique for investigating large quantities of categorical data. *Journal of Applied Statistics* 29 (2), 119–127.
- Lewis, S., Montgomery, D., Myers, R., 2001. Examples of designed experiments with nonnormal responses. *Journal of Quality Technology* 33 (3), 265–278.
- Liu, H., Davidson, R., Apanasovich, T., 2008. Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms. *Reliability Engineering & System Safety* 93 (6), 897–912.
- Liu, H., Davidson, R., Rosowsky, D., Stedinger, J., et al., 2005. Negative binomial regression of electric power outages in hurricanes. *Journal of infrastructure systems* 11, 258.
- Martínez-De-Pisón, F., Pernía, A., Jiménez-Macias, E., Fernández, R., 2010. Overall model of the dynamic behaviour of the steel strip in an annealing heating furnace on a hot-dip galvanizing line. *Revista de Metalurgia* 46 (5), 405–420.
- McCullagh, P., Nelder, J., 1989. *Generalized Linear Models*. Chapman and Hall, London.
- Montgomery, D., 2009. *Introduction to Statistical Quality Control*, 6th Edition. John Wiley and Sons, New York.
- Myers, R. H., Montgomery, D., Vining, G., 2010. *Generalized Linear Models with Applications in Engineering and the Sciences*. 2nd Edition (New York: Wiley).
- Ordieres-Meré, J., Martínez-de Pisón-Ascacibar, F., González-Marcos, A., Ortiz-Marcos, I., 2010. Comparison of models created for the prediction of the mechanical properties of galvanized steel coils. *Journal of Intelligent Manufacturing* 21, 403–421.
- Pacella, M., Semeraro, Q., Anglani, A., 2004. Manufacturing quality control by means of a fuzzy art network trained on natural process data. *Engineering Applications of Artificial Intelligence* 17 (1), 83–96.



- Pernía-Espinoza, A., Castejón-Limas, M., González-Marcos, A., Lobato-Rubio, V., 2005. Steel annealing furnace robust neural network model. *Ironmaking & Steelmaking* 32 (5), 418–426.
- Qinlan, J., 1993. *Programs for Machine Learning*. Morgan Kaufmann, The Morgan Kaufmann Series in Machine Learning.
- R Development Core Team, 2009. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.  
URL <http://www.R-project.org>
- Roshan, J., Adya, H., 2001. Prediction of yield in a multi-product batch production environment. *Quality Engineering* 14 (1), 153–159.
- Sellers, K., Lotze, T., 2011. *COMPoissonReg: Conway-Maxwell Poisson (COM-Poisson) Regression*. R package version 0.3.4.  
URL <http://CRAN.R-project.org/package=COMPoissonReg>
- Sing, T., Sander, O., Beerenwinkel, N., Lengauer, T., 2007. *ROCR: Visualizing the performance of scoring classifiers*. R package version 1.0-2.  
URL <http://rocr.bioinf.mpi-sb.mpg.de/>
- Skinner, K., Montgomery, D., Runger, G., 2004. Generalized linear model-based control charts for discrete semiconductor process data. *Quality and Reliability Engineering International* 20 (8), 777–786.
- Sonquist, J., Baker, E., Morgan, J., 1971. Searching for structure (alias-AID-III); an approach to analysis of substantial bodies of micro-data and documentation for a computer program (successor to the Automatic Interaction Detector Program). Ann Arbor, Institute for Social Research, University of Michigan.
- Tang, N., 1999. Characteristics of continuous galvanizing baths. *Metallurgical and Materials Transactions B* 30.
- Timofeev, R., 2004. Classification and regression trees (cart). theory and applications. Master's thesis, CASE - Center of Applied Statistics and Economics, Humboldt University, Berlin.
- Venables, W., Ripley, B. D., 2002. *Modern Applied Statistics with S-PLUS*. Springer, Nueva York.
- Zheng, H., D.Y.Gong, Wang, G., Liu, X., P.J.Zhang, 2007. Software of predicting mechanical properties of strip steel by using bp networks. *Journal of Iron and Steel Research* 19 (7), 54–57.