Original Article

Development and implementation of a mathematical model for predicting zinc coating mass in continuous hot-dip galvanizing plants

Eliomar Junior Teixeira Machado [1]
Erik de Souza Couto [1]
Luís Eduardo Smecelato Maldonado [1]
Márcio Pedroso Bastos [1]
Rodrigo Carvalho de Paula [1]
Wesley Rossi Pimenta [1*]

Abstract

Continuous galvanizing lines require precise control of zinc coating mass (g/m^2) to ensure product quality, avoid overcoating, and reduce costs. However, the physical distance between the air knives and the on-line coating gauge introduces a measurement delay that hinders real-time adjustments. To address this limitation, a multiple linear regression model was developed for Continuous Galvanizing Line 1 (LZC1) at CSN. A total of 343,573 operational data records, sampled at 10-second intervals, were collected. After data cleaning and preparation, 156,201 observations were retained. Exploratory data analysis and a stepwise variable selection procedure identified the most significant predictors, including line speed, air knife pressure and knife-to-strip distance. The ordinary least squares (OLS) model, validated through 10-fold cross-validation, achieved a coefficient of determination (R^2) of approximately 0.78. The fitted model was implemented in a dedicated function block within the line's programmable logic controller (PLC), where it continuously estimates coating mass in real time. A dynamic offset mechanism, frozen at the onset of each coating transition, compensates for steady-state error and accelerates convergence to the target value. Historical records confirm reduced coating variability, decreased transient overcoating, and lower zinc consumption, yielding both economic and environmental benefits.

Keywords: Real-time predictive control; Multiple regression; Zinc coating mass; Continuous galvanizing; PLC implementation.

1 Introduction

A continuous hot-dip galvanizing line is an industrial system designed to apply metallic coatings, usually zinc-based, to the surface of steel strip in a continuous manner, with the objective of increasing its corrosion resistance and extending its service life. The process occurs sequentially and automatically, comprising stages such as uncoiling, thermal annealing, immersion of the strip in a molten zinc bath, and subsequent removal of excess coating by air knives. Figure 1 illustrates a typical layout of these lines, which include devices such as accumulation towers, annealing furnaces, zinc pots, air knives, and coating gauges. The high operating speed of these lines (which can exceed 150 m/min) requires precise control of variables to ensure coating uniformity on both faces of the strip, meeting normative specifications and application requirements in sectors such as construction, automotive, and white goods [1-3].

One of the major problems faced in continuous galvanizing lines is related to the lag between the coating application point and the zinc layer measurement point on the steel strip. Due to the high temperatures involved in the process, coating gauges are positioned at a considerable distance from the zinc pot and air knives, being installed only after the cooling zone. This spacing causes a significant delay in obtaining coating data, making it difficult to perform real-time adjustments to operational parameters. As a consequence, the applied coating mass may not meet specified requirements, impacting product quality and causing material waste, especially through the occurrence of overcoating, which represents the application of a zinc mass above what is necessary [4].

The rationale for developing this study is based on the economic and operational relevance of coating control in the plant. Zinc is a high-value input, and its excessive consumption, frequently resulting from overcoating, represents one of the main sources of waste in these lines [1,2]. Beyond the direct impact on production costs, excess coating does not add value to the final product, being technically unnecessary and economically undesirable. In this context, building a predictive statistical model becomes a promising tool to anticipate coating mass based on operational variables

---

[1] Companhia Siderúrgica Nacional – CSN, Volta Redonda, Rio de Janeiro, RJ, Brasil.

*Corresponding author: engwesleyrossi@gmail.com
E-mails: eliomarjunior@csn.com.br; erikinf@gmail.com; luis.maldonado@csn.com.br; bastos.galvasud@csn.com.br; rodrigo.carvalho@csn.com.br;

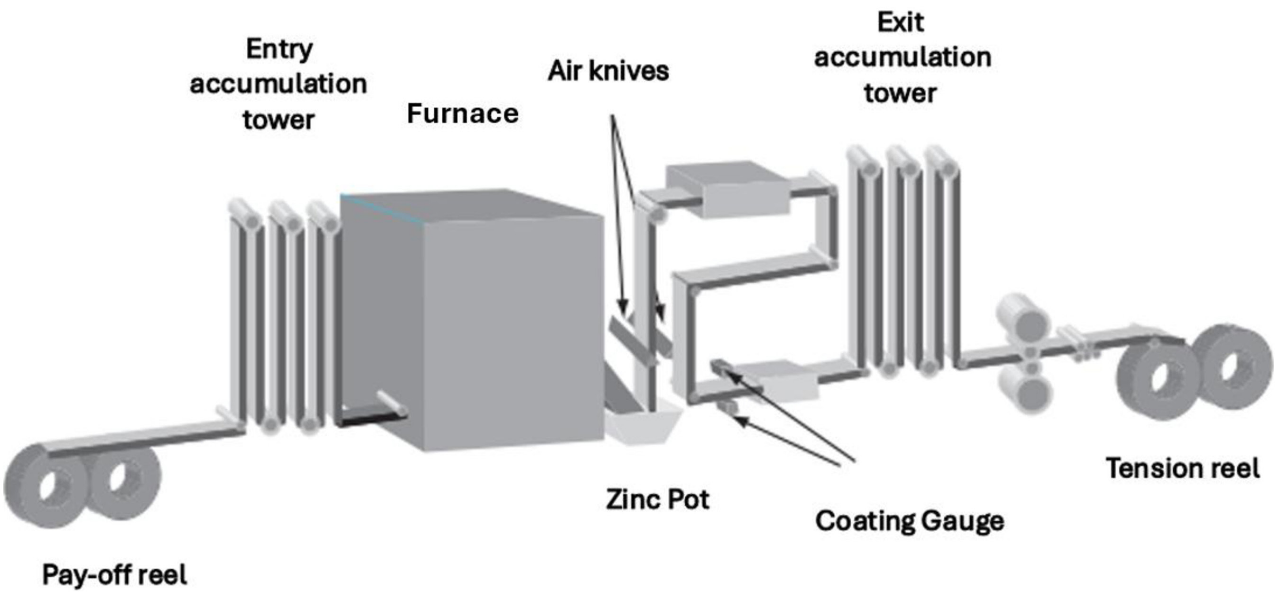


Figure 1. Typical layout of a continuous hot-dip galvanizing line [1].

available in real time. Such a solution enables faster and more accurate decision-making, promoting waste reduction, improved product quality, and greater reliability in process control, as advocated by studies focused on the application of regression models in the steel industry [5]. In recent years, data-driven modeling approaches have been increasingly applied to steelmaking and coating processes.

Pan et al. [6] developed and deployed a neural network-based coating weight control system for a hot-dip galvanizing line, achieving significant reductions in coating weight variance and transition time.

Elsaadawy et al. [7] proposed an analytical coating weight model based on pressure and shear stress correlations, validated with industrial coil average data.

These studies demonstrate the potential of data-driven models to enhance coating control. However, few studies focus on the practical implementation of such models to enable real-time control actions within industrial automation environments, which constitutes the main contribution of this work.

Therefore, the objective of this work is to develop a predictive mathematical model at Galvanizing Line 1 (LZC1) of Companhia Siderúrgica Nacional - CSN, based on multiple regression techniques, capable of estimating in real time the coating mass applied to the steel strip during the continuous galvanizing process. The proposal aims to use operational variables available at the moment of adjustment, such as line speed, air knife pressure and distance from the strip to the knives, enabling faster and more accurate decision-making. Additionally, the work includes implementing this model in the line's existing programmable logic controller (PLC) and applying a correction mechanism for steady-state error, ensuring high accuracy even during coating transitions.

2 Experimental procedure

The development of this work was based on the collection and analysis of operational data from LZC1, a plant belonging to the CSN group. The monitored variables included line speed, air knife pressure, distance between the knives, and knife height relative to the zinc pot, coating on both faces, among other parameters relevant to coating control. Data were recorded over several months of operation, ensuring coverage of different scenarios and typical conditions of the industrial process.

Data acquisition and storage were performed through computer stations equipped with IBA software, which enabled the recording and subsequent analysis of values associated with coating mass (g/m^2).

For the development of this work, different stages were adopted aimed at building a predictive model capable of predicting zinc coating mass (g/m^2). These stages include data collection, preparation, and analysis, as well as the application of statistical techniques for model construction.

Data were collected at 10-second intervals, totaling an initial volume of 343,573 data points, using specific sensors integrated into the IBA system. Each variable was continuously monitored to ensure a consistent collection frequency and to guarantee data representativeness. These data were obtained under various operating conditions to ensure comprehensive representativeness, ensuring that the model could handle the natural variability of the process.

Data preparation was essential to ensure the quality of information used in the model. All data preparation, exploratory analysis, stepwise variable selection, and regression fitting were performed using Microsoft Excel. Initially, missing values were removed, and set point transition periods were excluded, as they could introduce variability not representative of the process. Rounding of position data was also performed to reduce non-relevant fluctuations. Subsequently, a predictor analysis was conducted to verify equivalence among variables. To this end, all equivalent predictors were grouped with the objective of avoiding repeated values. Thus, distinct coating values were observed, and the average of these was calculated for each grouping, along with standard deviation analysis. After all treatments, the total number of observations was reduced to 156,201, indicating that 187,372 values were removed.

Exploratory data analysis (EDA) was conducted with the objective of investigating the relationships between monitored variables and the zinc coating mass. To quantify the intensity of these relationships, Pearson correlation was applied.

The correlation matrix, presented in Figure 2, showed that variables such as Hor_Inf_LO (knife-to-strip distance, operation side; r = −0.66), air knife pressure (r = −0.58), Hor_Sup_LA (knife-to-strip distance, action side; r = −0.58), and line speed (r = −0.42) have significant correlation with the coating mass, reinforcing their relevance as predictors in the model. The negative correlation with line speed suggests that higher operating speeds tend to occur together with other operational adjustments that influence coating mass.

Additionally, the matrix allowed the identification of multicollinearity indications among some parameters, especially between line speed and air pressure, which required attention during the variable selection stage for multiple regression.

For the construction of the predictive model, the stepwise automatic variable selection method was adopted, with the objective of identifying the most significant predictors for estimating the zinc coating mass. The criteria for variable entry and removal were based on statistical significance values (p-value), respecting predefined alpha limits to ensure model robustness. This procedure allowed the progressive inclusion and/or exclusion of variables based on their statistical contribution, optimizing the fit of the multiple

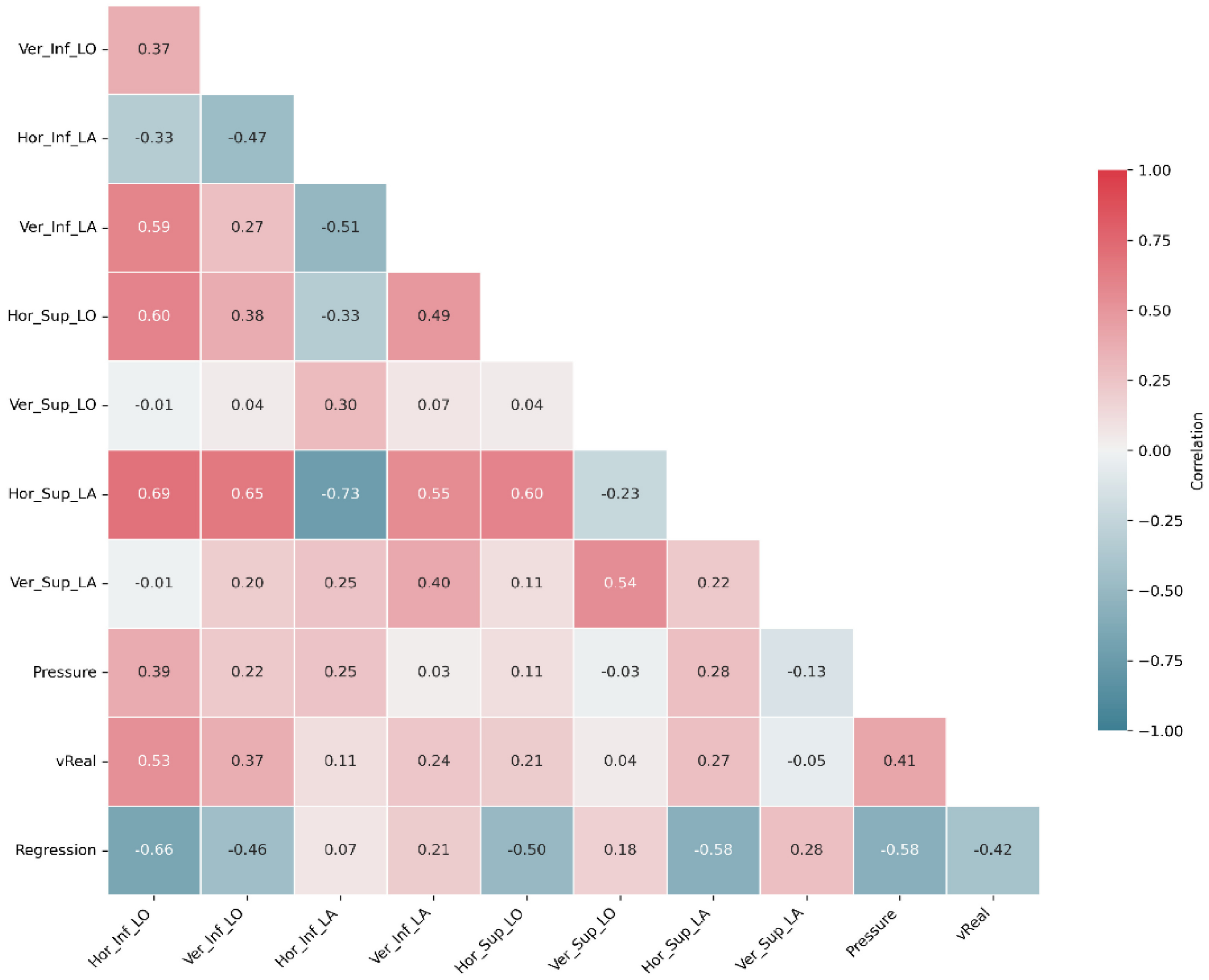


Figure 2. Correlation heatmap of process variables and zinc coating mass (g/m^2), used to identify correlations and guide stepwise variable selection for the multiple regression model.

regression model. After the iterative process, variables that demonstrated consistent statistical relevance were selected. The presence of multicollinearity among predictors was assessed using the Variance Inflation Factor (VIF), adopting an acceptability threshold to ensure there was no significant overlap of information among the selected variables. The stepwise procedure employed an entry significance level of $\alpha\_in = 0.05$ and a removal significance level of $\alpha\_out = 0.10$.

The VIF was computed for all retained predictors using a maximum acceptable threshold of 10; Although some variables presented VIF values between 5 and 10, indicating moderate multicollinearity, none exceeded the critical threshold of 10. In addition, all regression coefficients were statistically significant ($p < 0.001$), reinforcing the robustness of the fitted model.

The resulting multiple linear regression model follows the general form: $\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n$, where $\hat{y}$ is the predicted zinc coating mass (g/m^2), β_0 is the intercept, β_i are the regression coefficients, and x_i are the predictor variables. The correspondence between the process tags and the variables used in the model is presented in Table 1.

The multiple regression model was fitted using the ordinary least squares (OLS) method, which minimizes the sum of squared residuals between observed and predicted coating mass values. Model performance was evaluated through three complementary metrics using 10-fold cross-validation: the dataset was randomly partitioned into 10 equally sized subsets, with nine used for training and the remaining one for testing in each iteration. Across all 10 test folds, the model achieved a mean coefficient of determination (R^2) of approximately 0.78, indicating that the selected predictors explain roughly 78% of the coating mass variability. The mean absolute error (MAE) corresponded to approximately 9% of the mean observed coating mass, and the normalized mean squared error (MSE) corresponded to approximately 2%. The consistency of these metrics across all folds confirms that the model generalizes well to unseen data, with no evidence of overfitting.

Pareto charts of standardized effects were used to visualize the relative importance of variables in the model. The chart represented by Figure 3 indicated that Hor_Inf_LA (knife-to-strip distance, action side, lower surface) is the predictor with the highest standardized effect, followed by line speed (vReal) and air knife pressure. These variables should be rigorously monitored to maintain coating mass quality.

It is noteworthy that the ranking of predictor importance in the multiple regression model Figure 3 differs from the ranking of bivariate Pearson correlations Figure 2. For instance, Hor_Inf_LA exhibits a low bivariate correlation with coating mass ($r = 0.07$), yet it presents the highest standardized effect in the Pareto chart. Conversely, Hor_Inf_LO shows the strongest bivariate correlation ($r = -0.66$) but ranks sixth in the Pareto chart. This occurs because the standardized regression coefficients capture the partial contribution of each variable after accounting for the effects of all other predictors simultaneously.

In the presence of moderate multicollinearity among the geometric variables (VIF values between 5 and 10), suppression effects can amplify the partial contribution of variables that share variance with correlated predictors.

This behavior is well documented in multiple regression analysis and does not compromise model validity, as confirmed by the statistical significance of all coefficients ($p < 0.001$) and the stable cross-validation performance.

The regression model was deployed as a dedicated Function Block (FB) within the programmable logic controller (PLC) of LZC1. The FB incorporates the pre-trained regression coefficients, obtained from the off-line ordinary least squares fitting procedure described above, and receives as real-time inputs the ten process variables listed in Table 1: knife-to-strip heights and distances for both faces and sides (operation and action), air knife pressure, and line speed. The FB executes within the CPU's uninterrupted processing cycle, completing each evaluation in a scan time below 100 ms, which ensures that the predicted coating mass ($\hat{y}$) is updated at a rate fully compatible with the process dynamics.

In parallel, the FB computes a dynamic offset, defined as the difference between the most recent measured coating mass, obtained from the on-line coating mass gauge located downstream, subject to a transport delay inherent to the physical distance between the air knives and the measurement point, and the corresponding model prediction at the same time stamp. During steady-state operation, this offset is

Table 1. Predictor variables selected by the stepwise procedure and their corresponding process descriptions

Data	Coefficient	Description
Ver_Inf_LO	x1	Knife-to-strip height (mm) - Operation Side - Lower Surface
Hor_Inf_LA	x2	Knife-to-strip distance (mm) - Action Side - Lower Surface
Ver_Inf_LA	x3	Knife-to-strip height (mm) - Action Side - Lower Surface
Hor_Inf_LO	x4	Knife-to-strip distance (mm) - Operation Side - Lower Surface
Hor_Sup_LO	x5	Knife-to-strip distance (mm) - Operation Side - Upper Surface
Ver_Sup_LO	x6	Knife-to-strip height (mm) - Operation Side - Upper Surface
Hor_Sup_LA	x7	Knife-to-strip distance (mm) - Action Side - Upper Surface
Ver_Sup_LA	x8	Knife-to-strip height (mm) - Action Side - Upper Surface
Pressure	x9	Air knife pressure (mmH2O)
vReal	x10	Line speed (m/min)

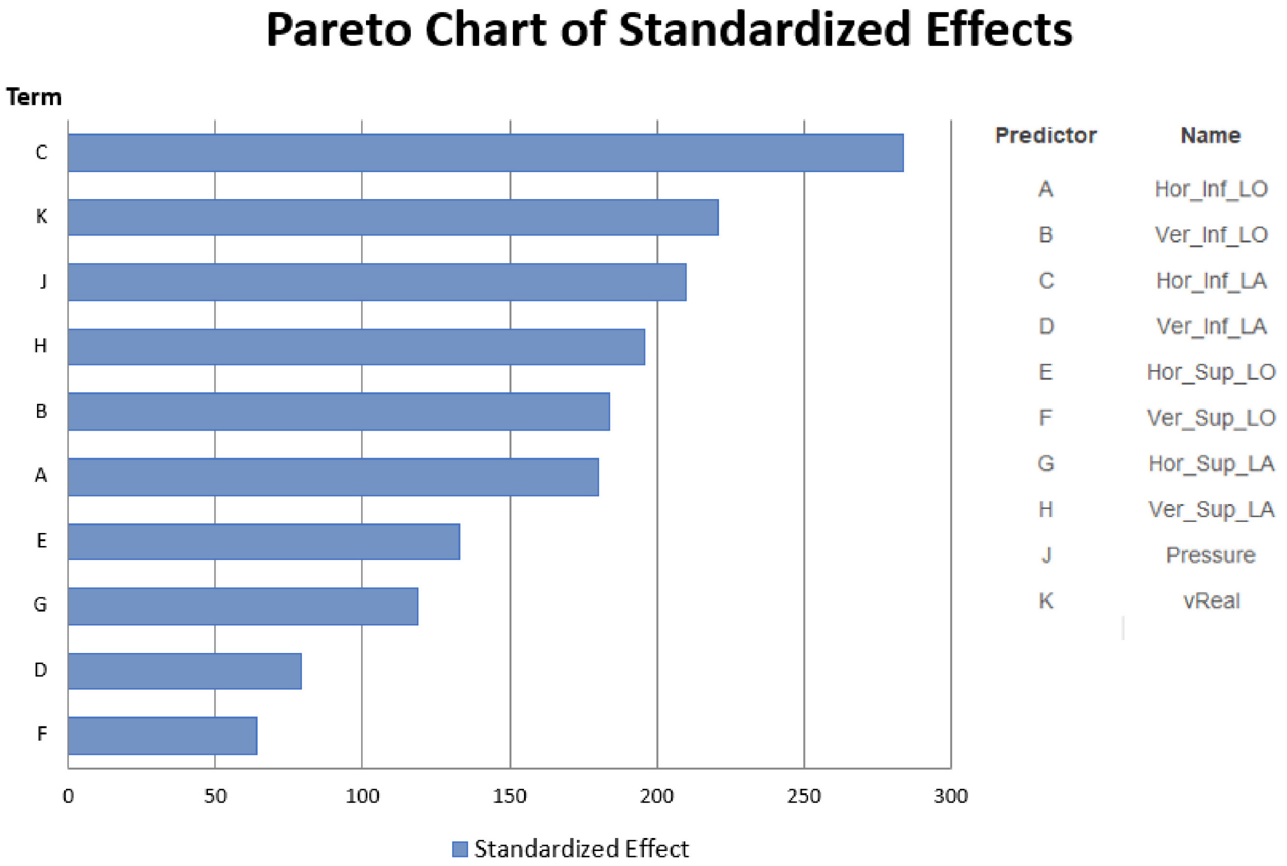


Figure 3. Pareto chart of standardized effects.

continuously updated, thereby compensating for systematic biases not captured by the regression model.

When the system detects the onset of a coating transition identified by a step change in the coating mass set point the current offset value is frozen and added to all subsequent predictions throughout the transient period. The offset remains fixed until the new steady state is confirmed by the gauge measurement, at which point continuous updating resumes. This freeze-and-hold strategy eliminates steady-state error during the critical transition window and accelerates convergence to the target coating mass.

The corrected predicted value is displayed in real time on the operator HMI (Human-Machine Interface), providing immediate visual feedback during transitions. The entire logic executes within the existing PLC hardware with no additional computational infrastructure required.

3 Results and discussion

The implementation of the predictive control system resulted in tangible improvements in coating control performance. The integration of the regression model with the PLC provided accurate real-time estimates of coating mass, enabling optimized corrections. As a consequence, a reduction in the variability of the zinc coating mass in the process was observed, with measured values converging more closely to the target zinc coating.

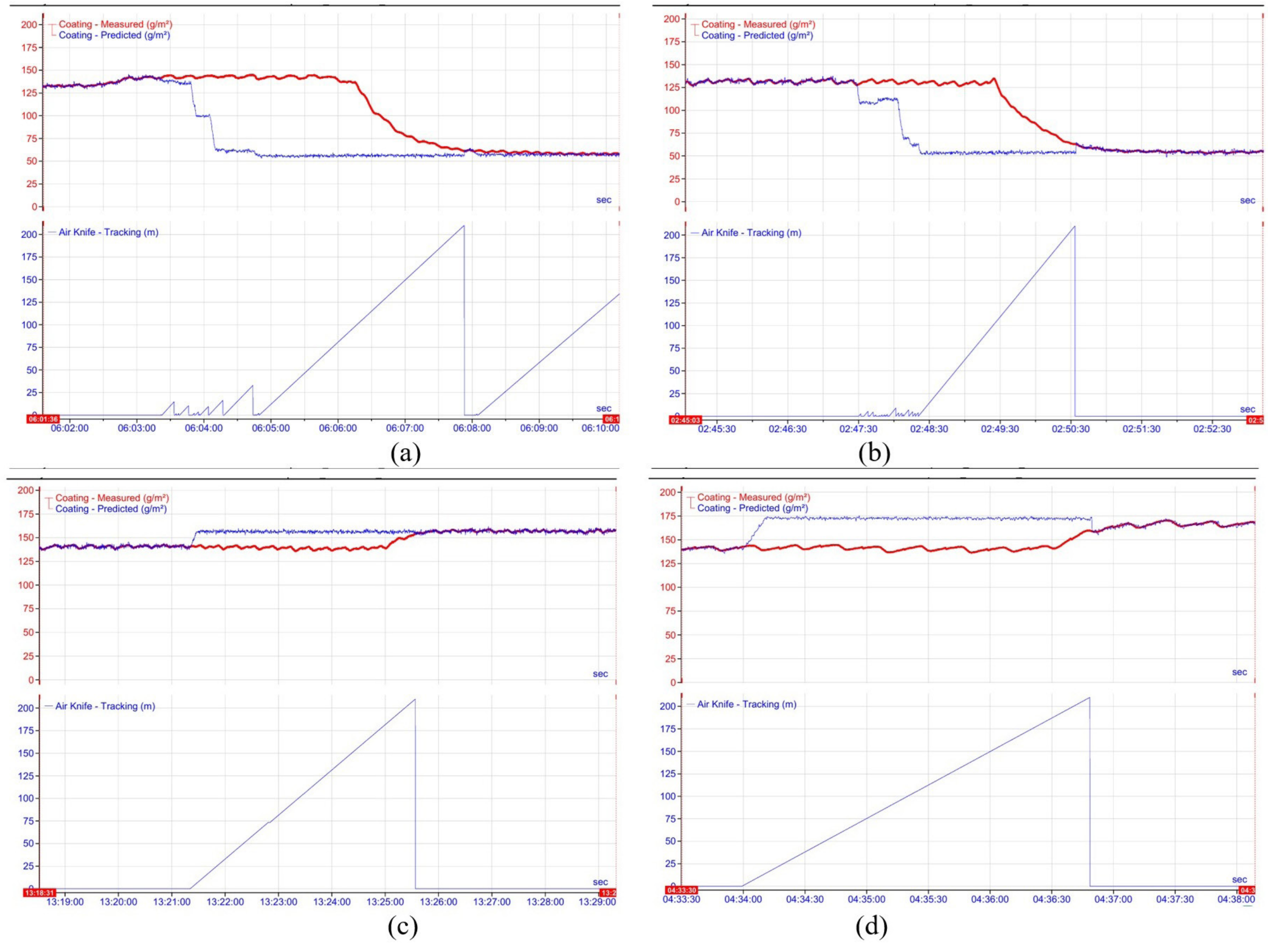
Figure 4 consolidates four historical records that demonstrate the performance of the predictive indication during distinct zinc coating mass transitions.

In Figures 4a and b, coating reduction is observed: the predicted curve (blue) and the actual measurement (red) converge rapidly to the actual coating value. In Figures 4c and d, the inverse condition is illustrated, of zinc coating mass increase, showing convergent response.

The dynamic adjustment of steady-state error proved decisive for system robustness. By eliminating this error, the mechanism indirectly compensates for process variables that could not be explicitly included in the model as predictors. In this way, control maintains high accuracy even in the presence of non-measurable factors in real time, ensuring process stability and uniformity in the applied zinc coating mass.

The model's benefits are reflected primarily in reduced zinc consumption and the consequent improvement in final product quality.

Lower over-application or overcoating of the metal triggers gains throughout the chain: it reduces the need for zinc extraction and refining, decreases energy consumption associated with these stages, and, by extension, mitigates relevant environmental impacts.



(a) (b)

(c) (d)

Figure 4. Model adjustment during coating mass transitions: (a) coating mass reduction period 1; (b) coating mass reduction period 2; (c) coating mass increase period 3; (d) coating mass increase period 4. Y-axis: zinc coating mass (g/m^2).

4 Conclusion

This study demonstrated that the application of a multiple regression model, integrated with the PLC of CSN's Galvanizing Line 1, is capable of estimating the zinc coating mass during transition periods.

The model, built from a robust set of operational data and validated by k-fold cross-validation (mean $R^2 \approx 0.78$) significantly reduced coating variability and mitigated overcoating during coating mass transitions, historically the most critical point of the line. The steady-state error compensation mechanism, continuously updated and frozen at the start of transitions, proved essential for maintaining process control within quality tolerances, while simplifying operation for the production team.

Beyond stability and quality gains, the reduction in zinc over-application brings direct and indirect benefits, such as reduced environmental impacts associated with the extraction, refining, and transport of this input. The approach, therefore, reconciles operational efficiency with sustainability.

As future work, it is recommended to: (i) incorporate additional variables, for example, bath temperature and strip thermal profile, to capture dynamic effects not yet modeled; (ii) explore nonlinear machine learning techniques, such as random forests or neural networks, for performance comparison; and (iii) establish automatic model retraining routines using continuous production data, ensuring adaptability to changes in raw materials or process conditions. These advances can further enhance control robustness and expand the benefits achieved by the present predictive solution.

Acknowledgements

The authors thank the CSN for providing access to the continuous galvanizing line facilities and the operational data that made this study possible. We extend our recognition to the Operations, Automation, and Process Engineering teams and, especially, to the dedicated Micrometer team, whose practical knowledge, technical support, and readiness were fundamental to the development and implementation of the predictive model presented here.

References

1 Linhares NJ, Campbell CHG, Mello AS, Bohrer RB, Silva JC, Santos PL. Reduction of overcoating in galvanizing lines. Tecnologia em Metalurgia e Materiais. 2008;5(2):72-76. https://doi.org/10.4322/tmm.00502002.

2 Linhares NJ, Campbell CHG, Mello AS, Bohrer RB. Reduction of zinc consumption in galvanizing lines. In: Associação Brasileira de Metalurgia, Materiais e Mineração. Proceedings of the 62nd ABM Annual Congress – International; 2007; Vitória, Brazil. São Paulo: ABM; 2007.

3 International Zinc Association. Control of coating weight [mass] for continuous hot-dip galvanized sheet products. GalvInfoNote 2.5, rev. 1.2. 2017 [cited 2025 Jun 9]. Available at: https://www.galvinfo.com

4 McKerrow PJ. Control of coating mass on a continuous hot dip galvanizing line [thesis]. Wollongong: University of Wollongong; 1978.

5 Kang H, Zhao H. Description and application research of multiple regression model optimization algorithm based on data set denoising. Journal of Physics: Conference Series. 2020;1631(1):012063. https://doi.org/10.1088/1742-6596/1631/1/012063.

6 Pan ZS, Zhou XH, Chen P. Development and application of a neural network based coating weight control system for a hot-dip galvanizing line. Frontiers of Information Technology & Electronic Engineering. 2018;19(7):834-846. https://doi.org/10.1631/FITEE.1601397.

7 Elsaadawy EA, Hanumanth GS, Balthazaar AKS, McDermid JR, Hrymak AN, Forbes JF. Coating weight model for the continuous hot-dip galvanizing process. Metallurgical and Materials Transactions. B, Process Metallurgy and Materials Processing Science. 2007;38(3):413-424. https://doi.org/10.1007/s11663-007-9037-2.

Received: 03 Feb. 2026
Accepted: 23 Apr. 2026

Editor-in-charge:
André Luiz Vasconcellos da Costa e Silva