Prediction of residual calcium carbonate ($CaCO_3$) in the lime kiln of a kraft pulp mill

TAÍS RESENDE COSTA, CARLA CRISTINA ARAÚJO PARREIRA, GUSTAVO MATHEUS DE ALMEIDA, MARCELO CARDOSO, FALAH ALOBAID, AND ESA VAKKILAINEN

ABSTRACT: In a kraft pulp mill, the rotary lime kiln is responsible for converting calcium carbonate ($CaCO_3$) into calcium oxide (CaO) to be reused in the causticizing plant. An important parameter of the calcination efficiency and the product quality is the amount of $CaCO_3$ not converted to CaO, commonly defined as residual carbonate. This parameter is usually determined through laboratory analysis, which introduces delays in process control and limits the ability to make timely operational adjustments.

This work presents a predictive model that functions as a soft sensor for the residual carbonate in a lime kiln of a Brazilian kraft pulp mill. This model was able to estimate residual $CaCO_3$ every 30 min using routinely measured online process variables, which is a considerable reduction from the 6-h average laboratory measurement. The correlation coefficient (*r*) between the model predictions and the mill values was 0.83, with a mean absolute error (MAE) of 0.33%. Considering the effect of the residence time throughout the lime kiln proved to be decisive in obtaining a satisfactory prediction result.

After model validation, a variable importance analysis showed that the carbon monoxide (CO) concentration in the stack flue gas (in ppm), secondary air temperature (in °C), and fan speed (in %) were the most influential variables in predicting the residual carbonate content. Combined with commonly used operation expertise, this subset of variables can serve as additional information to achieve smoother lime kiln operation.

Application: The model developed in this study demonstrates how routinely available process data can be used to estimate residual carbonate in an industrial lime kiln, providing a basis for operational assessment and further development of soft-sensing approaches in pulp mills. This can help better manage fuel efficiency, operational stability, carbon dioxide (CO_2) emissions, CaO quality, and residual carbonate content.

The rotary lime kiln is an essential part of the chemical recovery cycle in kraft pulp mills. It is responsible for the conversion of calcium carbonate ($CaCO_3$) back into calcium oxide (CaO) at high temperatures of above 850°C. The regenerated CaO is reused in the subsequent causticizing plant for the recovery of sodium hydroxide (NaOH), one of the main compounds in white liquor for wood delignification. The calcination process presents significant operational challenges, due to its multivariable nature, multiple physicochemical phenomena, and complex and variable dynamics [1,2].

In modern kraft pulp mills, a lime kiln is the only operational unit where fossil fuels are still needed to sustain the high temperature requirement [3]. Therefore, more effective monitoring and control of the lime kiln is important for improving operational efficiency, reducing atmospheric emissions, and supporting the pulp and paper industry's transition towards greater environmental sustainability.

A key parameter in lime kiln operation is the residual carbonate, defined as the amount of $CaCO_3$ not converted to calcium oxide. According to the literature, an ideal range for balancing energy efficiency and quality of the lime produced is between 1.5% and 2.5% [4]. In practice, it is usually kept below 4.5% by pulp mills in general. This value is commonly controlled by adjusting the temperature in the kiln calcination zone based on the fuel flow rate. As a reversible reaction, temperature and carbon dioxide (CO_2) concentration affect the kinetics of the calcination [5]. Relatively lower temperatures result in under calcination, with excess residual carbonate and low-reactivity lime. Conversely, relatively higher temperatures lead to over calcination, with burned lime and again low reactivity. Residual carbonate is therefore an important variable of the calcination process and an indirect measure of efficiency. It can also be used as a practical measure of the lime quality [6].

Current determination of the residual carbonate in pulp mills worldwide is generally done by laboratory analysis, typically every 4 h to 6 h or 8 h [6,7]. This scenario presents a limitation that makes it difficult to immediately control the process and make quick decisions. This challenge can be overcome by developing predictive models using online measurements of operational variables for a real-time estimation of the residual carbonate content. Such data-driven models, known as soft-sensors, differ from conventional physical sensors that directly measure parameters like flow, pressure, or temperature. In effect, a soft sensor is an alternative solution to predict a key process parameter that is difficult or even impossible to measure directly in the industrial process [8].

Data-driven predictive models have been used for various purposes in the pulp and paper industry. General applications include diagnosing fouling in kraft boilers [9], improving natural gas consumption efficiency and reducing sulfur oxide (SO_x) and nitrogen oxide (NO_x) emissions in lime kilns [10], as well as improving pulp washing [11]. More specifically, regarding soft sensor applications, Shan, Liu and Tang [12] used neural network and logistic regression for predicting alkali residue in washing pulp and the Baumé degree of the extracted black liquor. Karlström, Hill, and Johansson [13] also used an artificial neural network to predict pulp properties, such as shives (wide), fiber length, and freeness. He et al. [14] employed Random Forest, Gradient Boosting Regression, Ridge Regression, and K-Nearest Neighbors (KNN) to monitor folding endurance, bursting strength, smoothness, and transverse ring compressive strength in a papermaking process.

Ribeiro [6] reported an empirical model based on an artificial neural network to predict lime availability using simulated data. Guzzo [15] also employed the same technique to predict the calcination efficiency, but it has not been applied to a real kiln yet. Borges et al. [16] reported an advanced kiln control strategy that uses a prediction of the residual carbonate to reduce energy consumption and improve process stability, but the model used is not described.

The present work proposes a data-driven predictive model to estimate the residual carbonate content in a lime kiln of a kraft pulp mill in Brazil. Before presenting and discussing the results, the case study is described and the methodology is outlined. After model validation, a sensitivity analysis study investigated the most influential variables used as model inputs on residual carbonate. In sequence, practical examples of using this soft sensor for residual carbonate during lime kiln operation are discussed.

CASE STUDY

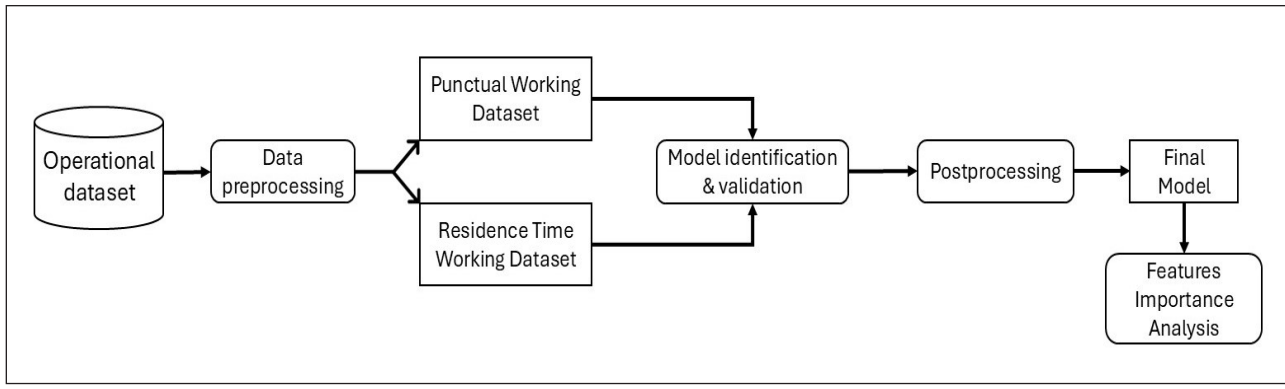
The case study of this work refers to operational data collected from a rotary lime kiln of a Brazilian pulp mill, which has a lime kiln equipped with a lime mud dryer (LMD). The dataset covers a period of approximately two years, from June 1, 2023, to June 3, 2025. The operational variables measured online were sampled every 30 min, resulting in 39648 observations for each of them, while the laboratory analyses for the residual carbonate content were performed every 6 h on average, totaling 1659 samples. From an initial number of 17 operational variables, a subset of nine variables was used as model input (**Table I**). We aimed to capture key physical, chemical, and operational aspects of the kiln operation while ensuring the present work was not overly specific to a single pulp mill in terms of predictor variables. This subset of variables encompasses kiln inputs (fuel, air, and feed), internal operating conditions (temperatures and flow), and kiln performance indicators (flue gas composition and energy efficiency), making them potentially suitable for predicting residual carbonate levels in real time.

Input Variable	Unit
Specific energy	MJ/t
Bed temperature	°C
Primary air temperature	°C
Secondary air temperature	°C
Flue gas temperature	°C
Fan speed	%
Residual oxygen	%
Lime mud feed rate	m^3/h
Carbon monoxide (CO) (in the kiln stack)	ppm
Output variable	–
Residual carbonate content	%

I. Variables considered in the data-driven predictive model for the residual carbonate content (t: metric ton).

Specifically, specific energy in megajoules per metric ton (MJ/t) was selected as an integrated indicator of kiln thermal demand, representing the total energy required per unit of lime produced. Although production rate is influenced by lime mud feed rate, specific energy reflects overall thermal efficiency and captures the combined effects of fuel input, combustion conditions, and heat transfer performance. As such, it is widely used in industrial practice to assess kiln energy efficiency and provides a fuel-independent performance metric. Heat load (MJ/t) is the metric routinely used by operators to assess lime kiln energy efficiency. This choice, therefore, improves interpretability and ensures applicability of the methodology to mills with different fuel configurations. The lime mud feed rate (m^3/h) was included to represent the material input and production load imposed on the kiln, directly influencing residence time, heat demand, and calcination dynamics.

Temperature variables characterize the thermal profile of the system. Bed temperature is directly related to calcination kinetics and residual carbonate formation. Primary and secondary air temperatures influence combustion stability and heat distribution, while flue gas temperature reflects overall heat transfer efficiency and potential energy losses. The fan speed (%) of the main lime kiln induced draft (ID) fan, located at the cold end, was used as an indicator of draft conditions and internal gas flow. These factors directly influence heat exchange inside the kiln. Residual oxygen (%) and carbon monoxide (ppm) in the kiln stack were selected as combustion quality indicators. Residual oxygen reflects excess air levels, while carbon monoxide



1. Methodology steps.

(CO) concentration indicates incomplete combustion and local reducing conditions, both of which may influence heat availability and calcination efficiency.

METHODOLOGY

Figure 1 shows the methodology steps, each of which are described below.

Data preprocessing

Successful data preprocessing relies on the integration of data analysis techniques and process knowledge. However, working with real process measurements brings an additional challenge. Initially, in this step, an exploratory data analysis (EDA) was carried out to search for patterns, relationships, and discrepancies in the data. Next, erroneous records (e.g., negative flows and zero residual carbonate content), and the kiln idle, startup, and shutdown periods were removed. The remaining anomalous data were disregarded using the Isolation Forest algorithm, a type of machine learning technique [17]. Subsequently, the resulting time series segments were smoothed with an adapted Savitzky–Golay filter [18]. One positive aspect is that this filter does not significantly distort the original signal.

We investigated two data-driven modeling approaches in this work. In the first, used as a baseline, residence time was not considered. In this case, both online variables and residual carbonate analysis were considered at the same sampling time. To account for residence-time effect, Järvensivu et al. [1] applied long moving averages (150–180 min) of selected variables to represent the kiln's slow dynamics, while filtering out short-term fluctuations. Following this idea and based on plant operating practice and consultation with kiln specialists, an average residence time of 2.5 h was assumed in this approach. In the industrial operation analyzed, the kiln is actively controlled to maintain this target value by continuously adjusting the rotational speed in response to variations in lime mud feed rate. Moreover, the residual carbonate content used for model development is obtained from laboratory analyses performed at approximately 6-h intervals, representing an integrated process response rather than instantaneous kiln conditions. To ensure consistency between model inputs and outputs, the input variables were defined as median values over the 2.5 h preceding each laboratory measurement. Given this temporal aggregation and the plant control strategy, short-term fluctuations in residence time are not expected to significantly affect the measured residual carbonate. In addition, we investigated a 6-h time window, corresponding to the sampling interval between laboratory analyses.

Model building and validation

To find a data-driven predictive model for the residual carbonate content in the lime kiln, we investigated three different machine learning techniques; namely, RandomForest (RF); Artificial Neural Network (ANN); specifically Multilayer Perceptron (MLP); and eXtreme Gradient Boosting (XGBoost) [19]. All of these techniques have been used in industrial applications, as reported in the literature. To train and test the models, the working dataset was randomly split into two subsets, a training set containing 80% of the samples and a testing set with the remaining 20%. The first subset is used to build the model, while the second is used for final validation check. Model training used the k-fold cross-validation procedure, with $k = 5$, combined with a hyperparameter grid search. Final model selection was based on the root mean square error (RMSE), correlation coefficient (r), and residual analysis combined with process knowledge. All consider the discrepancies between mill values and model predictions.

Postprocessing of the model outputs

This step assesses the impact of signal smoothing (used in the data preprocessing step) on model performance. To do so, we compared the predicted values of the final selected model with the original target values. The average and the standard deviation of the resulting residuals were calculated to estimate, respectively, the bias and uncertainty in the predictions, considering 0.5% intervals across the entire residual carbonate range. The model's predicted values were then adjusted using both statistics. In practice, these values would be those shown to the lime kiln operational team.

Features importance analysis

This step seeks to determine the effect of each model input variable on the residual carbonate content. We investigated three different methods to assess different aspects of the predictors' contributions. The *F*-score provides the relative influence of each variable on model structure. The Native Importance metric shows how much each variable contributes to reducing prediction error. The Permutation Importance evaluates how much the model depends on each variable for accurate predictions [20]. It is important to use a set of techniques, as they can present different ranks for the variables. These results were then associated with the kiln operation conditions.

RESULTS AND DISCUSSION

Data preprocessing

Manual detection of erroneous records and the Isolation Forest algorithm were initially used for anomalous data removal. The Savitzky–Golay smoothing filter was then used, where it is desirable to define a window length (equal to 5-time units in this work) that balances noise reduction without losing information about the trend and variability of the kiln operation. The resulting working dataset contains 1483 samples for the punctual approach and 1524 samples for the residence time approach. This represents a 10.6% and 8.1% reduction from the initial number of 1659 residual carbonate content samples in the raw data collected at the mill, respectively.

Model building and validation

In the punctual approach, the models were trained with 1186 samples and tested with 297 samples. In the 2.5-h residence time approach, the corresponding numbers were 1219 and 305 samples, respectively. We found out that considering a 6-h time window did not affect nor improve the model performance, so we opted to follow with the original 2.5-h approach, which is closer to the actual residence time in this kiln's operation.

Table II presents the best result of each model on the test dataset for each approach. Higher correlation coefficients and lower RMSE values were obtained considering the 2.5-h residence time, compared to the punctual approach. This highlights the importance of accounting for residence time when predicting residual carbonate content [21]. Under this consideration, XGBoost yielded the best result, with a correlation coefficient (r) considerably high, equal to 0.83 ± 0.01. The relatively low standard deviation suggests stability during model training, as desirable. The mean absolute error (MAE) was equal to 0.23 ± 0.01%, an acceptable error for residual carbonate content in practice. Other works have reported the use of this machine learning technique as a prediction tool in the pulp and paper industry [22,23].

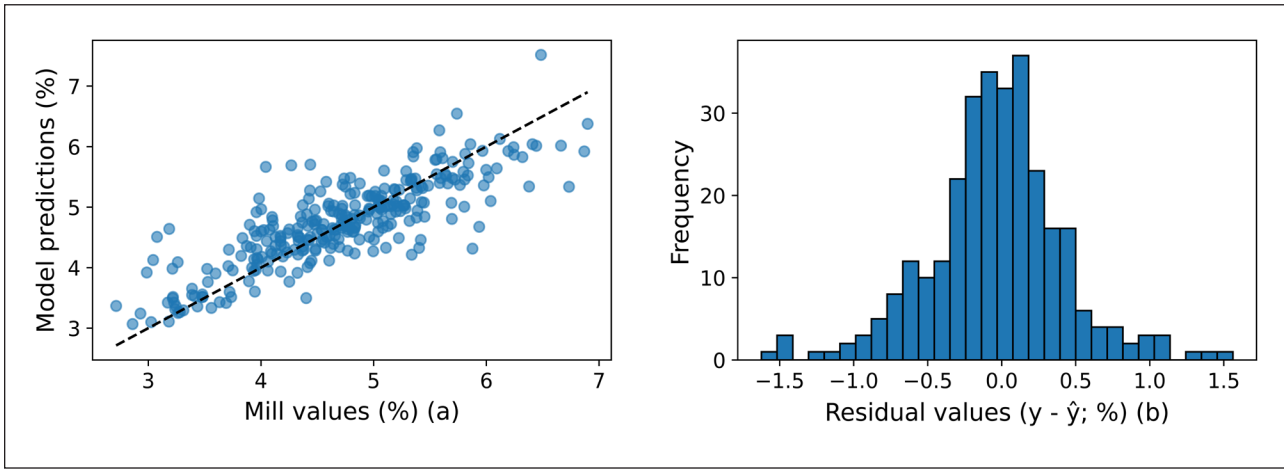
Figure 2 illustrates the results for XGBoost on the test dataset. Figure 2a depicts the high association between the model predictions (y-axis) and the mill values (x-axis) for the residual carbonate content. The largest differences appear at the upper extreme in the mill's values (above 6), which likely come from the limited number of samples. Since they fall far from the usual operating conditions of the kiln, such deviations were expected. Figure 2b shows the histogram of the residuals, with an approximate normal distribution and a mean value close to zero, as desired. Additionally, approximately 53% of the residuals fall within the range [-0.25, +0.25], and nearly 78% within [-0.5, +0.5]. These relatively small residual values suggest that the model could complement the routine laboratory analyses. In other words, it can provide information on the trend of kiln operating conditions in real time.

It is important to acknowledge that the uncertainty associated with laboratory measurements of residual carbonate may influence the interpretation of model performance. Although a quantitative assessment of measurement error was not available in this study, the analytical procedures are part of the mill's routine quality control and are conducted in accordance with standardized protocols, including regular instrument calibration and operation by trained personnel. These practices contribute to measurement consistency and reproducibility. Consequently, it is reasonable to assume that the variability captured by the model predominantly reflects actual process dynamics rather than random measurement noise, thereby representing operational trends under typical kiln conditions.

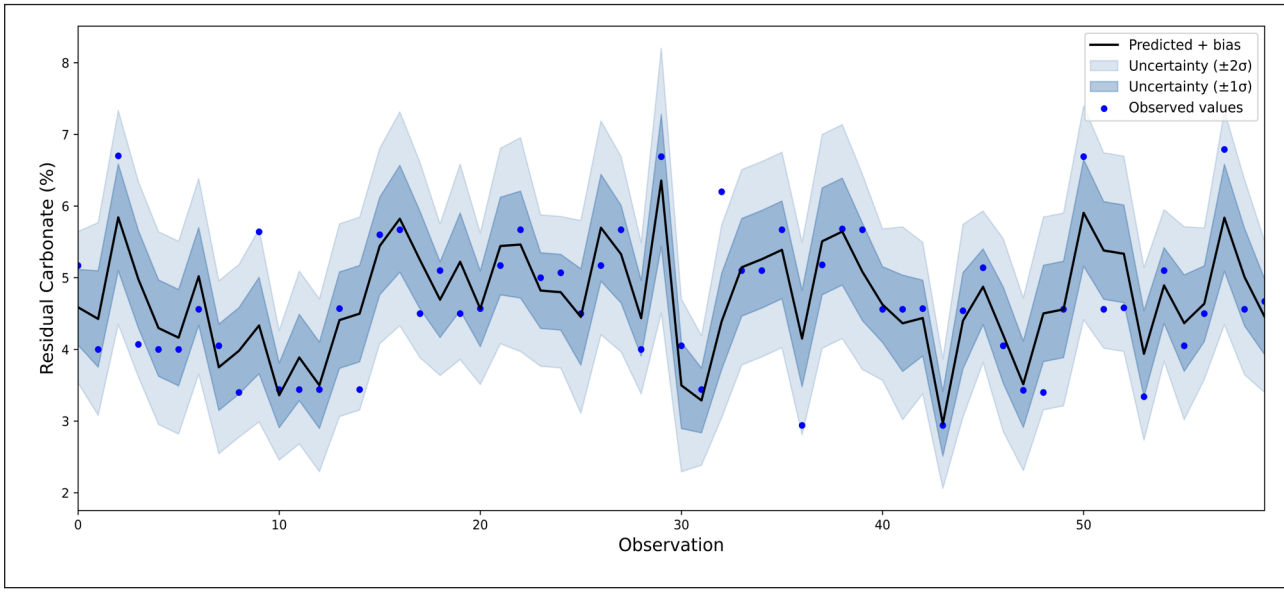
Model	Punctual Approach			Residence Time Approach (2.5 h)		
	r	MAE	RMSE	*r*	MAE	RMSE
ANN	0.29 ± 0.03	0.69 ± 0.01	0.97 ± 0.01	0.45 ± 0.06	0.56 ± 0.02	0.72 ± 0.02
Random Forest	0.63 ± 0.01	0.53 ±0.01	0.75 ± 0.01	0.80 ± 0.01	0.36 ± 0.01	0.48 ± 0.01
XGBoost	0.73 ± 0.01	0.46 ± 0.01	0.61 ± 0.01	0.83 ± 0.01	0.34 ± 0.01	0.46 ± 0.01

ANN: artificial neural network; XGBoost: eXtreme gradient boosting; *r*: correlation coefficient; MAE: mean absolute error; RMSE: root mean square error.

II. Performance results of the models in the test subset.



2. XGBoost results for the residual carbonate content in the test dataset: (a) model predictions vs. mill values, and (b) residuals distribution.



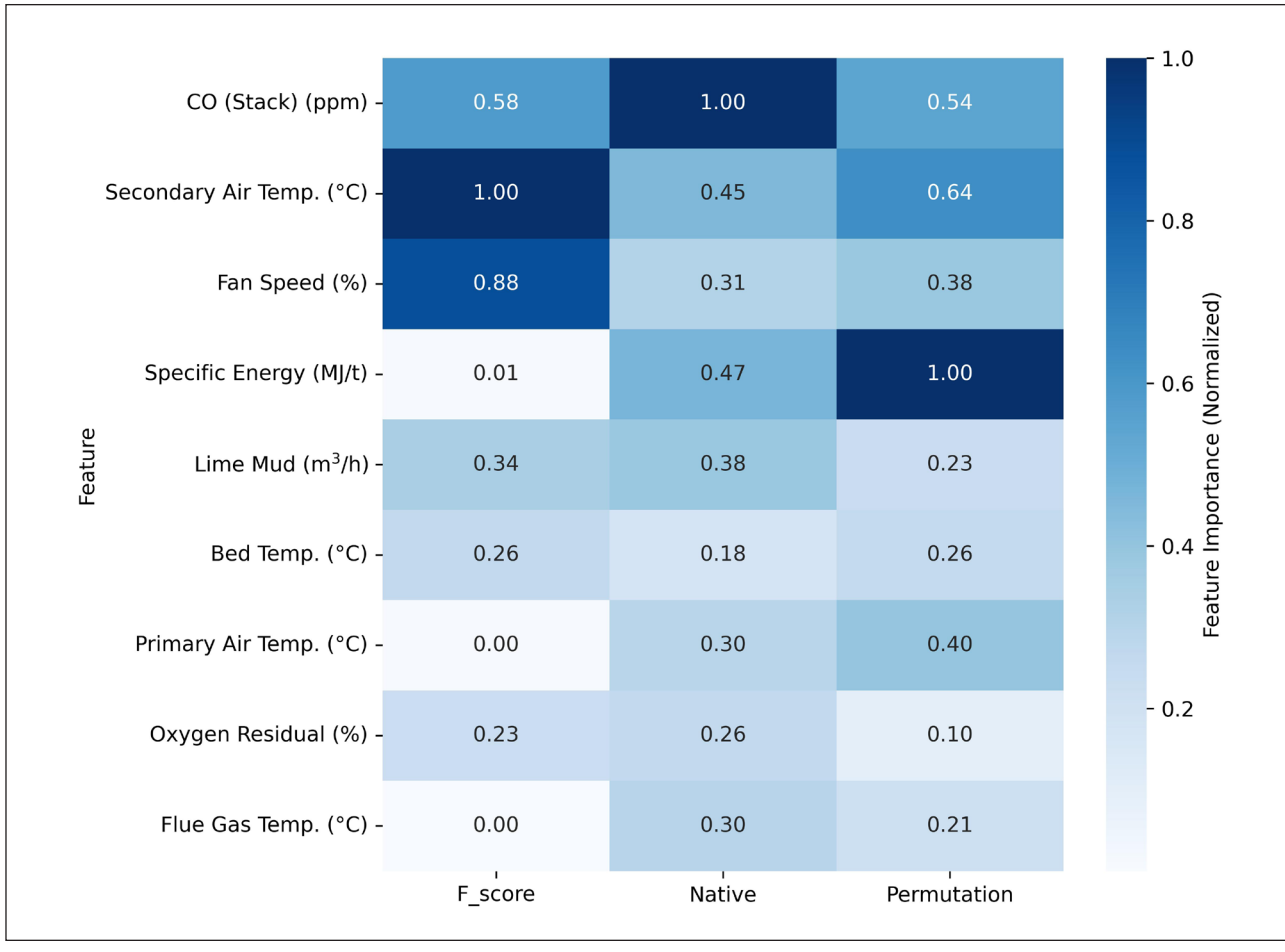
3. XGBoost results for the residual carbonate content in the test dataset: mill values (blue dots); model predictions (black line); and approximate normal uncertainty bands (±1 standard deviation, dark shadow; and ± 2 standard deviations, light shadow).

Postprocessing of the model outputs

Figure 3 presents the adjusted predicted values (black line) by the previously selected XGBoost model. The original values of residual carbonate content, collected at the mill, are also shown (blue dots). Additionally, intervals corresponding to one (dark shadow) and two (light shadow) standard deviations are also plotted. This means that the predicted value occurs in these intervals with approximately 68% and 95% chance, respectively. These bands are useful for quantifying the uncertainty in the model's prediction. Around 73% of the mill data (blue dots) fall within the first interval and 94% within the second. Mill values above and below the predicted values suggest that the model is consistent; that is, it does not suffer from the problem of overfitting. This result also suggests that the smoothing filter (data preprocessing step) did not compromise the accuracy of the model. Thus, the residual carbonate content in the lime kiln was satisfactorily described by the XGBoost model using a set of nine process variables measured online. The figure only shows a specific range of test data for visualization purposes, but this pattern was observed across the entire test set.

Feature importance analysis

The feature importance analysis revealed that, in this case, CO concentration in the stack flue gas, secondary air temperature, and fan speed are the most influential variables for predicting residual carbonate content. These variables show high relevance in all three techniques (**Fig. 4**). The closer to unity (the darker the color), the greater the importance of the variable. The significance of the three influential values is as follows:



4. Feature importance analysis using F-Score, Native Importance, and Permutation Importance techniques.

- *CO concentration (ppm)* can be a sensitive indicator of combustion efficiency. Higher CO levels signal incomplete fuel oxidation and temperature deficiencies that favor the $CaCO_3$ retention [7,24,25].
- *Secondary air temperature (°C)* is relevant due to its governing role on both the stability and completion of the combustion process and the efficiency of heat transfer to the bed [26]. Both directly affect the availability of energy for the decomposition of calcium carbonate.
- *Fan speed (%)* is a key driver of airflow patterns and gas circulation, impacting flame stability, heat distribution, and gas residence time within the kiln chamber, all of which affect the thermal environment required for efficient calcination [27, 28]. It also controls the CO_2 concentration in the gaseous phase.

These variables can also be directly related to the calcination kinetics, since they are strongly influenced by the temperatures and CO_2 partial pressure (controlled by the fan speed). The reaction rate constant is a function of temperature. Although higher temperatures are required to initiate calcination in a high CO_2 environment, the reaction occurs much faster than in a low CO_2 atmosphere [5]. Also, when the temperature drops below 800°C, the reverse reaction takes place and recarbonation occurs, so the external CaO reacts with the CO_2, regenerating $CaCO_3$ [4].

Other variables of significance were specific energy and lime mud feed rate as follows:

- *Specific energy (MJ/t)* appears to have a strong but unclear effect on carbonate predictions, as it exhibits high predictive power despite being used only occasionally.
- *Lime mud feed rate (m³/h)* has a limited impact on carbonate predictions but reflects the fundamental effect of kiln load, as higher loads increase energy and air demand and, if not properly balanced, lead to incomplete calcination. These results are consistent with the idea that both material load dynamics and combustion-related parameters must be carefully balanced to achieve optimal stable kiln operational performance and residual carbonate content.

The remaining variables presented less predictive power; however, they remain important for interpretability, as they also cover the physical, chemical, and operational aspects of the kiln.

Potential uses

This section discusses potential uses of the data-driven predictive model for residual carbonate content obtained earlier in this work.

Anomaly detection

The first concerns a soft sensor application. The model can provide information about the residual carbonate content earlier than laboratory analysis. In this work, the measurement time was reduced from 6 h to 30 min, considering the sampling interval of the model input variables in the case study. Early detection of anomalies can aid diagnostic tasks, more directly linking changes in residual carbonate to potential root causes. Consequently, early interventions during lime kiln operation, when necessary, benefit production, environmental, and safety aspects. This can also contribute to generating new insights into the dynamics of kiln operation.

Process control

Integrating the model into the kiln fuel control system can favor fuel consumption, operational stability, lime quality, and the optimal residual carbonate range. From an environmental perspective, this can reduce greenhouse gas emissions (GHG) from fuel combustion, especially CO_2, and reduce the release of incomplete combustion products. As part of the calcium cycle, this could ensure a more consistent lime quality for the downstream causticizing plant, with minimal variability, reducing excess purge and reprocessing of material.

Process monitoring

In this specific case, it refers to monitoring the data-driven predictive model itself. The model predictions and laboratory analyses every 6 h (considering the sampling interval in this work) can serve as redundant sensors. For example, increasing discrepancies between these two residual carbonate content values over time could serve as a warning sign that the model needs to be updated. This information is very useful for maintaining the reliability of the model and, consequently, kiln operation stability, as the model is used for decision-making (anomaly detection application described earlier in this section). A major challenge in machine learning projects is precisely how to maintain model reliability, as equipment and process conditions change over time. Ultimately, this contributes to monitoring the process itself.

CONCLUSIONS

This work aimed to build a predictive model for the residual carbonate content in a rotary lime kiln of a Brazilian pulp mill. Given the complexity of the calcination process, we combined process knowledge with a data-driven approach. The final XGBoost model using nine process variables was able to satisfactorily predict the residual carbonate content. The correlation coefficient (r) was equal to 0.83 and the mean absolute error (MAE) to 0.23% in the test dataset.

Lime kilns in kraft pulp mills can benefit from online residual carbonate estimation and having this information ahead of commonly used laboratory analyses. In this work, this means every 30 min compared to the 6-h average for laboratory analysis. Early interventions can benefit production, environmental, and safety aspects; for example, regarding fuel management, fossil fuel consumption, thermal energy losses, lime quality for the causticizing plant, CO_2 emissions, and waste disposal related to excessive purging. Practical uses of the predictive model, including anomaly detection, process control, and process monitoring, were also discussed.

The methodology proposed in this work can be directly applied to other lime kilns with minimal adaptation. It can also be extended to other rotary kiln applications, including the cement and lime industries, since they have similar thermal processes that involve carbonate decomposition.

ACKNOWLEDGMENTS

This work was the result of a cooperation between the Federal University of Minas Gerais (UFMG) in Brazil and the Lappeenranta-Lahti University of Technology (LUT) in Finland. The authors thank the Minas Gerais State Agency for Research and Development (FAPEMIG) and LUT University for the financial support, and the Brazilian pulp mill for providing data and all information on the operation of the lime kiln.

LITERATURE CITED

1. Järvensivu, M., Saari, K., and Jämsä-Jounela, S.-L., *Control Eng. Pract.* 9(6): 589(2001). https://doi.org/10.1016/S0967-0661(01)00017-X.
2. Charos, G.N. and Arkun, Y., *Tappi J.* 74(2): 203(1991).
3. Lipiäinen, S., Kuparinen, K., Sermyagina, E., et al., *Sustainable Prod. Consumption* 29: 421(2022). https://doi.org/10.1016/j.spc.2021.10.029.
4. Tran, H., *TAPPI Kraft Recovery Operations Course*, TAPPI Press, Atlanta, 2008.
5. Wu, R., Carrejo, E., Reza, M.S., et al., *Fuel* 373: 132372(2024). https://doi.org/10.1016/j.fuel.2024.132372.
6. Ribeiro, B., "Prediction of the lime availability on an industrial kiln by neural networks," *Int. Jt. Conf. Neural Networks, Proc., IEEE World Congr. Comput. Intell.*, IEEE, New York, 1998, p. 1987. https://doi.org/10.1109/IJCNN.1998.687164.
7. Tran, H., Vollicks, S., Gauthier, M., et al., *TAPPI J.* 6(4) : 9(2007).
8. Oikonomou, D., Leontaris, L., Dimitriou, N., et al., *IEEE Sens. J.* 25(4): 7681(2025). https://doi.org/10.1109/JSEN.2025.3526362.
9. Belisário, A.B., Edberg, A., Björk, M., et al., *TAPPI J.* 22(3): 162(2023). https://doi.org/10.32964/TJ22.3.162.

10. Suárez, V.M.G., García-Gonzalo, E., Bayón, R.M., et al., *Int. J. Adv. Manuf. Technol.* 100: 1549(2018). https://doi.org/10.1007/s00170-018-2773-4.
11. Lianhua, H., Xinping, L., Wei, T., et al., "The application of fuzzy control in pulp washing process," *ICMTMA 2011*, IEEE, New York, 2011, p. 686. https://doi.org/10.1109/ICMTMA.2011.743.
12. Shan, W.J., Liu, B., and Tang, W., *J. Korea TAPPI* 54(4): 57(2022). https://doi.org/10.7584/JKTAPPI.2022.08.54.4.57.
13. Karlström, A., Hill, J., and Johansson, L., *BioResources* 19(1): 1030(2024). https://doi.org/10.15376/biores.19.1.1030-1057.
14. He, Z., Qian, J., Li, J., et al., *J. Cleaner Prod.* 372: 133803(2022). https://doi.org/10.1016/j.jclepro.2022.133803.
15. Guzzo, G.S., "Prediction of calcination efficiency based on lime kiln input variables in a pulp mill," *Annu. Pulp Pap. Congr., 57th*, ABTCP, São Paulo, 2025.
16. Borges, A.S., Luiz, A.C., Medeiros, C.C.R.M., et al., "Lime kiln optimization using Aspen DMC3 advanced control: energy consumption reduction and process stabilization," *Annu. Pulp Pap. Congr., 57th*, ABTCP, São Paulo, 2025.
17. Liu, F.T., Ting, K.M., and Zhou, Z.H., "Isolation forest," *IEEE Int. Conf. Data Min., 8th*, IEEE, New York, 2008, p. 413. https://doi.org/10.1109/ICDM.2008.17
18. Savitzky, A. and Golay, M.J.E., *Anal. Chem.* 36(8): 1627(1964). https://doi.org/10.1021/ac60214a047.

ABOUT THIS PAPER

Cite this article as:
Costa, T.R., Parreira, C.C.A., De Almeida, G.M., et al., *TAPPI J.* 25(5): 301(2026).
https://doi.org/10.32964/TJ25.5.301

DOI: https://doi.org/10.32964/TJ25.5.301

ISSN: 0734-1415

Publisher: TAPPI Press

Copyright: ©TAPPI Press 2026

About this journal

19. Russell, S.J. and Norvig, P., *Artificial Intelligence: A Modern Approach*, 4th edn., Pearson, London, 2022.
20. Scikit-learn developer, "Inspection," in *Scikit-Learn 1.7.2 Documentation*," scikit-learn.org, 2025, Chap. 5. Available [Online] https://scikit-learn.org/stable/inspection.html <08May2025>.
21. Berg, E., Hermansson, S., Wetterling, J., et al., *TAPPI J.* 23(6): 315(2024). https://doi.org/10.32964/TJ23.6.315.
22. Liu, X., Hong, J., Zhang, M., et al., *Nord. Pulp Pap. Res. J.* 40(1): 149(2025). https://doi.org/10.1515/npprj-2024-0066.

ABOUT THE AUTHORS

We studied this topic because most kraft pulp mills these days still estimate the residual carbonate through laboratory analysis, making lime kiln control difficult. This research complements previous studies by applying predictive modeling specifically to residual $CaCO_3$ levels, offering a data-driven approach to kiln optimization.

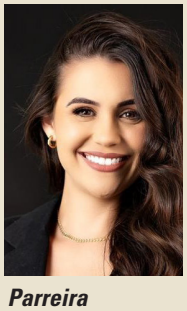
The most challenging aspect of this research was developing the model using real industrial data. This demanded thorough preprocessing of data to ensure accuracy and reliability.

Through our study, we were able to connect the most influential variables for predicting residual $CaCO_3$ with the underlying process knowledge. An interesting aspect of the research was that the model provided not only a prediction, but also an associated uncertainty range indicating the region within which the actual residual $CaCO_3$ value is likely to fall.

Mills could use this model to optimize kiln operation, improving lime quality and reducing energy consumption, process inefficiencies, and environmental impacts. Our next step is to test the model in an actual industrial environment.



Costa



Parreira



De Almeida



Cardoso



Alobaid



Vakkilainen

Costa is master's student, Parreira is a Ph.D. candidate, and Cardoso is professor at the Federal University of Minas Gerais (Universidade Federal de Minas Gerais – UFMG) in Belo Horizonte, MG, Brazil. De Almeida is a researcher, Alobaid is professor, and Vakkilainen is professor at Lappeenranta University of Technology in Lappeenranta, Finland. Email Resende Costa at tais.resende.costa@lut.fi.

23. Rodrigues, J.A., Farinha, J.T., Mendes, M., et al., *Energies* 15(17): 6308(2022). https://doi.org/10.3390/en15176308.

24. Fernández, J.R., Turrado, S., and Abanades, J.C., *React. Chem. Eng.* 4: 2129(2019). https://doi.org/10.1039/C9RE00361D.

25. Cui, Y., Yu, C., Shi, D., et al., *Appl. Therm. Eng.* 235: 121299(2023). https://doi.org/10.1016/j.applthermaleng.2023.121299.

26. Teguh, N.H., Yuliati, L., Siswanto, E., et al., *Data Brief* 52: 109931(2024). https://doi.org/10.1016/j.dib.2023.109931.

27. Lee, H.J., Park, J., and Lee, S.H., *Mater. Trans.* 58(9): 1322(2017). https://doi.org/10.2320/matertrans.M2017148.

28. Ryan, J., Bussmann, M., and DeMartini, N., *Processes* 10(8): 1516(2022). https://doi.org/10.3390/pr10081516.

