# Machine Learning-Driven Architecture for Real-Time Optimization in Industrial Processes

Filipe A. T. Silveira
*Aeronautics Institute of Technology (ITA)*
São José dos Campos, Brazil
filipeandersonn@gmail.com

Elcio H. Shiguemori
*Institute for Advanced Studies (IEAv)*
São José dos Campos, Brazil
elciohs@ita.br

Filipe A. N. Verri
*Instituto Curvelo*
São José dos Campos, Brazil
filipe.verri@institutocurvelo.org.br

***Abstract*—This paper presents a modular framework and a predictive modeling approach to enhance Real-Time Optimization (RTO) strategies in industrial combustion processes using Machine Learning (ML). The proposed architecture defines a generalizable methodology for integrating ML models into RTO systems, supporting scalable and adaptable applications in energy-intensive operations. A case study on continuous furnaces is used to demonstrate the methodology, where predictive models are developed to estimate optimal initial setpoints for the Lower Calorific Value (LCV) of the fuel mixture. Historical process data are used to train and compare several supervised learning algorithms, aiming to reduce reliance on manual configurations and improve the efficiency of the optimization system. Among the tested models, the Random Forest regressor achieved the best performance, combining low prediction error with high explanatory power. These predictions contributed to a noticeable reduction in natural gas consumption when used to initialize the RTO. The results confirm the potential of ML to improve optimization outcomes and support sustainable, data-driven decision-making in industrial processes.**

***Index Terms*—Real-Time Optimization, Machine Learning, Predictive Modeling, Modular Architecture, Industrial Energy Efficiency.**

## I. Introduction

Real-Time Optimization (RTO) plays a critical role in modern industrial operations by dynamically adjusting setpoints to improve efficiency and reduce costs [1]. In processes with slow dynamics—such as continuous combustion furnaces—the choice of the initial setpoint is crucial, as poor initialization can delay convergence and diminish potential energy savings.

Machine Learning (ML) techniques have shown increasing potential to enhance decision-making in industrial environments through predictive modeling and soft sensing [2], [3]. Our previous work [1] demonstrated that predicting the Lower Calorific Value (LCV) of the fuel mix using ML can accelerate RTO convergence and improve fuel efficiency in a steel reheating furnace operating with Natural Gas (NG) and Blast Furnace Gas (BFG).

This paper advances that approach in two ways. First, we propose a modular architecture to generalize the integration of ML into RTO systems. Second, we expand the modeling scope by evaluating a broader set of supervised learning algorithms for LCV prediction. These contributions aim to strengthen the predictive layer of RTO and promote scalable applications in energy-intensive industrial processes.

This article expands upon a previous technical publication (DOI: 10.37885/241218374), introducing new experiments, enhanced methodology, and a generalized integration framework.

## II. Related Work

Machine Learning (ML) has been widely applied in industrial contexts, demonstrating significant improvements in predictive maintenance, energy efficiency, and process control. [2] presents a case study on real-time soft sensing and anomaly detection in chemical operations. [3] explores supervised learning models to reduce energy consumption in thermal processes. [4] proposes a modular reference architecture for integrating ML into industrial control loops, focusing on lifecycle management and scalability.

While these studies illustrate the benefits of ML in industry, they often target monitoring or high-level optimization scenarios, without directly addressing integration with Real-Time Optimization (RTO) strategies.

Recent works have begun to explore the synergy between ML and optimization in real-time control contexts. [5] proposes a hybrid architecture that combines neural networks and mechanistic models for improved setpoint tracking in nonlinear systems. [6] discusses how ML-based surrogates can accelerate RTO convergence while maintaining operational constraints. A broader review by [7] outlines how ML enhances model predictive control (MPC) and optimization pipelines, including case studies in energy systems and chemical processes.

These works reinforce the relevance of ML for industrial optimization tasks. However, they typically focus on general architectures or simulated scenarios.

Our contribution builds upon this literature by proposing a modular ML-RTO architecture tailored for combustion furnace control and by validating its predictive layer using real industrial data. While the proposed system has not yet been deployed in a production environment, the results demonstrate its potential for improving setpoint initialization and accelerating RTO convergence.

## III. Contextualization

Considering the continuous heating process and its characteristics defined in [8], the furnace has 8 zones distributed across 3 units. These furnaces are used to perform heat treatment on materials. To accomplish this, the energy required for the heat treatment must be delivered through gas burners distributed along the zones and furnaces. The gas used is a mixture of NG and BFG. A mixing station controls the ratio of these gases according to the LCV (Lower Calorific Value) specified in operational practice. The blend is adjusted by controlling the flow of NG mixed with BFG.

The furnace system includes an LCV adjustment of the mixed gas, composed of blast furnace gas enriched with natural gas, which determines the fuel gas's burning power without considering its condensation. The mixing station optimizes the use of BFG, which is produced in the blast furnace process. Therefore, the objective of this station is to optimize the LCV to consume as little NG as possible, which has a high cost, without harming the quality of the treated material. That is, NG must be used when BFG is scarce and/or as energy complementation depending on the operation rate of the process.

To increase the energy efficiency of this process and consequently produce at the lowest possible cost, [8] proposed an optimal control strategy, through a technique of Real-Time Optimization for the fuel gases, defining an optimization problem as shown in (1) and (2). This strategy aims to maximize the use of BFG, which does not have a significant cost, while minimizing NG consumption.

$$\begin{aligned} &\text{Minimize:} \\ &\qquad \delta e_{\min(q_{\mathrm{mg_i}}, lcv_{\mathrm{mg}})} = \\ &\qquad |(\gamma \cdot q_{\mathrm{max_i}} - q_{\mathrm{mg_i}}) \cdot lcv_{\mathrm{mg}}| \\ &\text{w.r.t.} \quad lcv_{\mathrm{mg}} \end{aligned} \tag{1}$$

$$\text{s.t.:} \begin{cases} \text{Furnace Pressure } \leq p_{\max} \text{ mmCA} \\ \text{BFG flow } \leq bfg_{\max} \text{ Nm}^3/\text{h} \\ \text{NG flow } \leq ng_{\max} \text{ Nm}^3/\text{h}, \\ lcv_{\min} \leq \text{LCV} \leq lcv_{\max} \text{ kcal/Nm}^3 \end{cases} \tag{2}$$

where $\boldsymbol{\delta e}_{\min}$ is the function representing the minimum increment of heat transfer rate among zones to be minimized, taking into account the established percentage of the maximum zone flow ($\gamma$), the maximum flow rate ($\boldsymbol{q_{max_i}}$) of zone $i$, the gas flow of the current mixture of zone $i$ ($\boldsymbol{q_{mg_i}}$), and the lower calorific value of the mixed gas ($\boldsymbol{lcv_{mg}}$). Additionally, $p_{\max}$ denotes the maximum furnace pressure, $bfg_{\max}$ is the maximum BFG flow, and $lcv_{\min}$ and $lcv_{\max}$ represent the lower and upper limits of the LCV, respectively.

Figure 1 graphically presents the relationship between the LCV of the fuel gas mixture and the NG and BFG flow rates. It can be seen that by minimizing the LCV of the gas mixture, the NG flow rate is reduced and, consequently, the production cost.

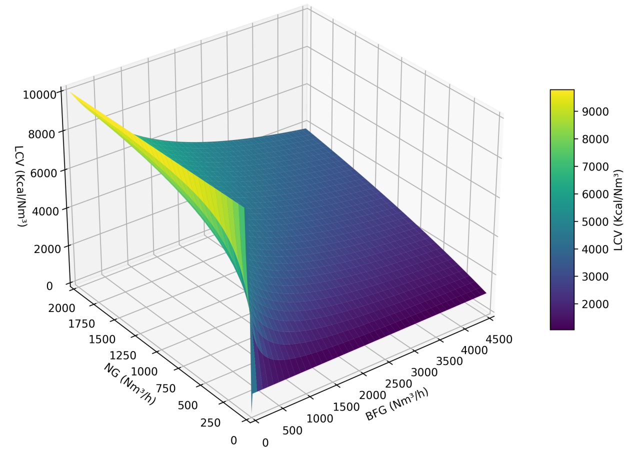


Fig. 1. LCV based on the NG and BFG gas flow [8].

## IV. Research Problem

The primary challenge addressed in this research is the optimization of the LCV in continuous combustion furnaces. These furnaces, which are critical for maintaining the desired properties of materials during industrial heating processes, pose significant control and optimization challenges due to their inherent complexity and high energy consumption.

In the current operational framework defined in [8], the LCV, which is a measure of the energy content of the fuel gas mixture, is adjusted manually based on operator experience and standard practices. This approach, while functional, is suboptimal as it does not fully account for the dynamic nature of the furnace environment, where variations in material type, width, thickness, and production speed continuously influence the required energy input.

Furthermore, while Real-Time Optimization systems have been implemented to enhance energy efficiency by adjusting the fuel mix dynamically, these systems still rely on initial setpoints that may not be optimal. The process of adjusting from these suboptimal initial conditions to an optimized state can result in delayed energy savings and inconsistent performance.

The research problem, therefore, revolves around the need to predict the optimal initial LCV setpoint using historical furnace data and machine learning techniques. By accurately predicting the LCV starting point that corresponds to optimal fuel efficiency, the transition period during which the RTO system fine-tunes the fuel mix can be minimized. This would lead to more immediate and consistent energy savings, particularly in reducing natural gas (NG) consumption, which is the more expensive component of the fuel mix.

In summary, the research problem is centered on enhancing the Real-Time Optimization of continuous combustion furnaces by developing a predictive model that accurately determines the optimal LCV setpoint. This approach is expected to improve the efficiency of the RTO system, leading to reduced energy costs and enhanced operational performance.

### A. Objective

This work aims to enhance the energy efficiency and operational performance of continuous combustion furnaces by integrating Machine Learning (ML) techniques into Real-Time Optimization (RTO) strategies. To this end, it proposes a structured methodology for ML-RTO integration, based on a modular architecture that defines distinct components for data acquisition, feature engineering, predictive modeling, and optimization. As a concrete application, the methodology is employed to predict the optimal initial setpoint of the LCV of the fuel mixture, enabling the RTO system to initiate operation closer to its optimal condition.

In industrial processes with slow thermal dynamics, such as continuous furnaces, initializing the optimization routine from a suboptimal LCV setpoint leads to delayed convergence and limits energy savings. By training supervised learning models on historical process data—including material characteristics, temperature setpoints, and operational conditions—this study aims to estimate more accurate LCV starting points compared to manual adjustments traditionally made by operators, thereby minimizing the reliance on high-cost NG while maximizing the use of BFG.

To achieve this, the following steps are undertaken in this work: data collection and preprocessing, feature engineering, model development using multiple ML algorithms, performance evaluation, and preparation for integration with the RTO system. Additionally, a modular and reusable architecture is proposed to generalize this ML-RTO integration strategy for broader industrial applications.

## V. Methodology

Following the data science lifecycle outlined by [9], this work is structured into four main stages: (i) defining the goal, (ii) collecting and preparing the data, (iii) developing predictive models, and (iv) conducting experiments and performance evaluation. To support the systematic implementation of these steps and promote reusability across industrial applications, we introduce a modular framework for integrating Machine Learning models into RTO systems in the subsequent section.

### *A. Define the Goal*

The goal of this project is to enhance the RTO of continuous combustion furnaces by predicting the LCV using machine learning models. This prediction will serve as an input for optimizing the fuel mix, thus reducing the consumption of natural gas and improving energy efficiency.

Since the RTO requires time to adjust from the initial setpoint defined by the operator to the optimal LCV, it was observed that starting with an optimal LCV could enhance the NG reduction, moving from the percentage reduction observed at "RTO Reduction" to the one at "Final Reduction" shown in Table I. These results were obtained from real process data collected in March 2024, during periods when the RTO was active. Therefore, this work aims to leverage historical data to train machine learning models capable of predicting this initial setpoint, allowing the system to begin operation closer to the optimal LCV. This approach is expected to further reduce energy consumption, aligning the results more closely with the "Final Reduction" values.

### *B. Data Collection and Management*

The process data for the furnace were collected from a time series data historian (PIMS – Plant Information Management System), covering the period from `2024-01-24 16:30:00` to `2024-04-23 11:25:00`, to build the dataset. The following variables were considered in the dataset:

- **material_type** (categorical): indicates the type of material being processed in the furnace.
- **tic_sp_z1** to **tic_sp_z8** (numeric, °C): temperature setpoint from each of the eight furnace zones.
- **width** (numeric, mm): the width of the material strip being processed.
- **thickness** (numeric, mm): the thickness of the material strip.
- **LCV_oper** (numeric, kcal/Nm$^3$): the initial LCV set by the operator.
- **speed** (numeric, m/min): the speed at which the material moves through the furnace.
- **RTO_state** (binary): indicates whether the RTO system is active (1) or inactive (0).
- **LCV** (numeric, kcal/Nm$^3$): the actual LCV of the fuel mix.

As each variable was stored with a different sampling rate, it was necessary to establish a consistent sampling period of 5 minutes to synchronize the data. Periods where the RTO was inactive or the furnace was idle or not producing valid material were excluded to ensure data relevance for model training.

### *C. Exploratory Analysis*

An initial descriptive analysis was conducted on the dataset to understand the distribution and characteristics of the collected data. This analysis included both numerical and categorical features.

Table II presents a statistical summary of the numerical features in the dataset, with each feature summarized across key statistical metrics such as count, mean, standard deviation, minimum, maximum, and percentiles. This summary provides insight into the distribution and variability of the temperature setpoints across the eight furnace zones, as well as the physical characteristics of the material being processed, and the operational parameters.

The dataset also includes categorical variables. The variable `material_type` captures the types of materials processed. For confidentiality reasons, material type names were anonymized and do not correspond to real product codes. Figure 2 shows the frequency distribution of each material type. A noticeable imbalance is observed, with some types (e.g., BMC41) appearing much more frequently, which may influence model training.

TABLE I
NATURAL GAS CONSUMPTION AND REDUCTION RESULTS

| Start Date | End Date | Material | NG consump. OFF | NG consump. ON | Final NG consumption | RTO Reduction | Final Reduction |
|---|---|---|---|---|---|---|---|
| 10/03/2024 16:00 | 10/03/2024 18:20 | Material A | 456 $Nm^3/h$ | 426 $Nm^3/h$ | 411 $Nm^3/h$ | 6.5% | 9.8% |
| 10/03/2024 18:50 | 10/03/2024 19:50 | Material B | 580 $Nm^3/h$ | 541 $Nm^3/h$ | 529 $Nm^3/h$ | 6.7% | 8.8% |
| 11/03/2024 14:00 | 11/03/2024 17:30 | Material C | 415 $Nm^3/h$ | 394 $Nm^3/h$ | 390 $Nm^3/h$ | 5.0% | 6.0% |
| 23/03/2024 12:45 | 23/03/2024 15:15 | Material D | 908 $Nm^3/h$ | 860 $Nm^3/h$ | 828 $Nm^3/h$ | 5.2% | 8.8% |

TABLE II
STATISTICAL SUMMARY OF NUMERICAL FEATURES

| Statistic | Count | Mean | Std | Min | 25% | 50% | 75% | Max |
|---|---|---|---|---|---|---|---|---|
| **tic_sp_z1** | 10323 | 1002.90 | 159.56 | 390 | 950 | 1070 | 1110 | 1205 |
| **tic_sp_z2** | 10323 | 1008.25 | 163.97 | 386 | 950 | 1080 | 1120 | 1191 |
| **tic_sp_z3** | 10323 | 1039.75 | 120.19 | 150 | 960 | 1090 | 1125 | 1180 |
| **tic_sp_z4** | 10323 | 1044.02 | 121.93 | 150 | 970 | 1100 | 1130 | 1190 |
| **tic_sp_z5** | 10323 | 1055.88 | 119.13 | 150 | 987 | 1108 | 1130 | 1200 |
| **tic_sp_z6** | 10323 | 1063.94 | 120.01 | 211 | 1015 | 1095 | 1135 | 1218 |
| **tic_sp_z7** | 10323 | 1072.36 | 121.86 | 204 | 1035 | 1090 | 1145 | 1225 |
| **tic_sp_z8** | 10323 | 1077.64 | 124.36 | 204 | 1050 | 1085 | 1155 | 1226 |
| **width** | 10323 | 1240.41 | 63.85 | 1030 | 1240 | 1240 | 1270 | 1345 |
| **thickness** | 10323 | 1.51 | 0.61 | 0.40 | 1.19 | 1.20 | 1.98 | 4 |
| **LCV_oper** | 10323 | 3180.57 | 415.07 | 2000 | 2900 | 3200 | 3600 | 4000 |
| **speed** | 10323 | 28.79 | 8.27 | 8 | 22 | 32 | 35 | 45 |
| **LCV** | 10323 | 2487.33 | 405.48 | 1744 | 2134 | 2384 | 2756 | 4161 |

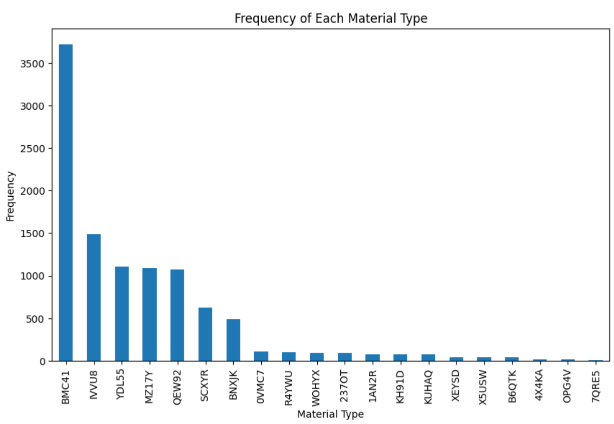


Fig. 2. Distribution of materials processed.

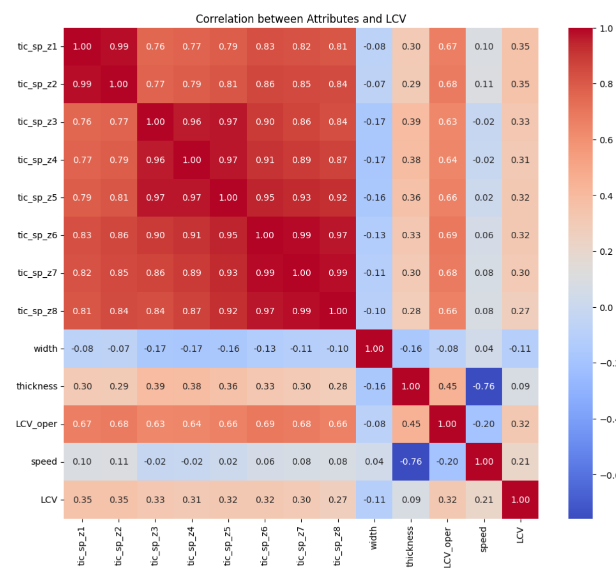


Fig. 3. Correlation between numerical attributes.

Additionally, a correlation analysis was conducted to explore relationships between variables, as shown in Figure 3. A moderate positive correlation was identified between `LCV` and `tic_sp_z1`, `tic_sp_z2`, and `LCV_oper`. High correlation among furnace zone temperatures suggests multicollinearity, which needs to be addressed during model development.

## *D. Building the Model*

The process of building the predictive model involved critical steps, including data preparation, preprocessing, model selection, and evaluation. Below, we detail each of these steps and the methodologies employed.

### *1) Data Preparation*

The dataset was divided into features (`X`) and the target variable (`y`), which is the `LCV` (Lower Calorific Value). The features included both numerical variables, such as temperature setpoints and material characteristics, and categorical variables, such as the type of material processed. The data was then split into training and testing sets using an 80-20 split to ensure that the model could be effectively trained and evaluated.

*2) Preprocessing*

Given the complexity of the data, a comprehensive preprocessing pipeline was developed. The numerical features (`tic_sp_z1` to `tic_sp_z8`, `width`, `thickness`, `speed`) underwent several transformations, including the exclusion of observations with missing values and standardization using `StandardScaler`, which scales each feature to have zero mean and unit variance. This ensures that all variables contribute equally during model training, avoiding dominance by variables with larger numeric ranges.

Additionally, since a high degree of multicollinearity was identified among the temperature setpoints across different furnace zones, dimensionality reduction via Principal Component Analysis (PCA) was employed. PCA transforms the correlated variables into a new set of uncorrelated components. In this context, PCA improves numerical stability and mitigates multicollinearity, which can otherwise lead to unstable coefficient estimates in linear models or hinder learning in more complex models. The PCA was configured to retain 95% of the variance in the data.

The categorical feature `material_type` was handled separately, with missing values imputed with a placeholder, followed by one-hot encoding. This approach ensured proper integration of categorical variables into the modeling pipeline.

Figure 4 illustrates the correlation matrix of the principal components after applying PCA. During the pre-processing phase, the PCA retained 4 principal components, which collectively explained 95% of the variance in the dataset. This dimensionality reduction was critical for improving the model's efficiency and interpretability, without sacrificing significant predictive power.


Fig. 4. Correlation Matrix of Principal Components After PCA

*3) Model Selection and Evaluation*

To identify the most effective predictive model, several supervised regression algorithms were evaluated. These models were grouped into three categories according to their learning approach:

**1) Linear and Regularized Models**:

- **Linear Regression:** Included as a baseline due to its simplicity and interpretability.
- **Ridge Regression:** Selected to address multicollinearity with L2 regularization.
- **Polynomial Regression:** Considered to model potential nonlinearities in a low-complexity format.

**2) Ensemble Learning Models**:

- **Random Forest Regressor:** Chosen for its robustness and ability to capture complex nonlinear relationships through bagging.
- **Gradient Boosting Regressor:** Included for its strong predictive performance using sequential additive modeling.
- **XGBoost Regressor:** Added to explore advanced gradient boosting techniques known for efficiency, built-in regularization, and superior generalization on tabular data.
- **LightGBM Regressor:** Evaluated for its gradient-based, leaf-wise growth approach, offering speed and accuracy for large datasets.

**3) Neural Network Architectures**:

- **Fully Connected Neural Network (DNN):** Incorporated for its high modeling flexibility and ability to capture complex feature interactions when given sufficient data and tuning.
- **Convolutional LSTM (CNN-LSTM):** Considered as an advanced hybrid architecture capable of learning spatial patterns through convolutional layers and sequential dependencies through LSTM units. Its inclusion aimed to evaluate whether temporal dynamics or local structural patterns in the data could be leveraged to improve regression performance.

Each model was embedded into a pipeline with consistent preprocessing. Hyperparameter optimization was conducted via `GridSearchCV` (a scikit-learn tool that performs an exhaustive search over specified parameter values using cross-validation) for all regression models and via `Keras Tuner` (a library for optimizing neural network hyperparameters using search algorithms) for the neural network architectures, including tuning the number of layers, units, activation functions, dropout rates, and learning rate. This step is crucial because the performance of machine learning models heavily depends on the proper selection of hyperparameters, which control the learning dynamics, model complexity, and generalization capability. A systematic search for optimal configurations ensures that each model reaches its best potential under the constraints and characteristics of the dataset, avoiding underfitting or overfitting and improving robustness in unseen scenarios.

All models were validated using 5-fold cross-validation and evaluated on the test set using Root Mean Squared Error (RMSE) and the coefficient of determination ($R^2$) as primary metrics, both of which are standard for regression tasks involving continuous variables. RMSE emphasizes large deviations and is particularly relevant in industrial settings where prediction errors can have significant operational impacts. $R^2$ indicates the proportion of variance explained by the model, offering a normalized view of predictive quality.

*E. Machine Learning-Driven Architecture for RTO*

To generalize the integration of machine learning techniques into Real-Time Optimization (RTO) systems, a modular and extensible architecture is proposed. This framework defines distinct layers that encapsulate data handling, model prediction, and control execution, enabling flexible deployment across a variety of industrial processes.

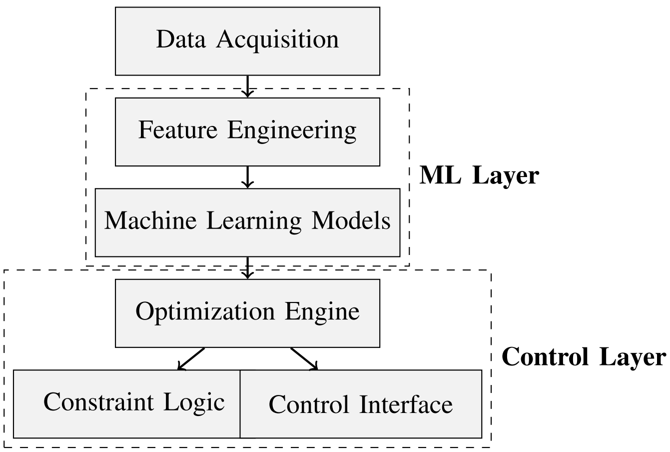


Fig. 5. Modular architecture for integrating Machine Learning models into Real-Time Optimization systems.

The proposed framework defines a conceptual architecture for augmenting Real-Time Optimization systems with machine learning capabilities. While the complete architecture illustrated in Figure 5 was not deployed in the industrial setting, its ML layer was fully developed and evaluated using real operational data in the case study. The control layer—including the optimization engine and constraint logic—had already been implemented in previous work by the authors [8], focused on optimal control of fuel gases in continuous annealing furnaces. This prior deployment indicates that a compatible control infrastructure already exists, providing a foundation for potential future integration of the ML components.

*1) Architecture Components*

The framework comprises six main functional blocks:

1) **Data Acquisition**: Interfaces with SCADA, PLCs, or data historians to collect real-time and historical process data. This block is essential for providing reliable and timely inputs that reflect the actual state of the process, serving as the foundation for all subsequent analytics and optimization tasks.
2) **Feature Engineering**: Converts raw signals into meaningful features suitable for predictive modeling, including temporal aggregates and statistical transformations. Proper feature construction enhances model performance by revealing relevant patterns and reducing noise, which is critical in industrial environments with complex dynamics.
3) **Machine Learning Models**: Use structured inputs to predict critical latent variables (e.g., LCV, heat demand), informing the optimization engine. These models extend visibility into unmeasured or uncertain aspects of the process, enabling more informed and proactive decision-making.
4) **Optimization Engine**: Solves constrained problems based on ML predictions to compute optimal control actions. This module transforms insights into actionable strategies by balancing performance objectives with system constraints, directly influencing process efficiency and cost.
5) **Constraint Logic**: Ensures process safety, regulatory compliance, and feasibility of optimization outputs. It acts as a safeguard layer that validates the applicability of optimization results, preventing unsafe or non-compliant operations.
6) **Control Interface**: Integrates with supervisory or lower-level control systems to apply setpoints or corrective actions. This final component closes the loop by translating high-level optimization outcomes into real-world actuation, ensuring that improvements are effectively realized in the physical process.

*2) Operational Enhancements*

Although the proposed architecture was not fully implemented in the industrial environment, this section presents a conceptual framework for operational enhancements that would support robust deployment in real-time scenarios. These enhancements aim to increase the reliability, transparency, and resilience of ML-RTO systems and are intended as guidelines for future implementation:

- **Model Lifecycle Management**: Supports versioning, retraining, and automated validation of predictive models. This ensures that models remain accurate and aligned with evolving process conditions, facilitating maintainability and regulatory compliance over time.
- **Monitoring**: Enables drift detection and confidence interval tracking in real-time. These mechanisms are essential for detecting model degradation or shifts in data distribution, allowing timely interventions to prevent suboptimal or unsafe decisions.
- **Fallback Mechanisms**: Allows manual override or reversion to default control behavior in case of model or sensor faults. This increases system resilience by ensuring operational continuity and safety in the event of failures or unexpected behavior.
- **Interpretability**: Implements model explainability tools (e.g., SHAP) to provide insight into ML predictions. By making the model's decisions transparent, this fosters trust among operators and engineers, supporting informed human-in-the-loop interventions.

*3) Benefits*

Compared to traditional RTO architectures, the proposed framework introduces a data-driven layer powered by Machine Learning (ML), enabling predictive capabilities and more responsive optimization. By incorporating ML models into the decision-making pipeline, the system can anticipate optimal setpoints, reduce convergence time, and improve operational consistency from the start of each production cycle.

This approach brings several specific advantages:

- **Predictive Initialization**: Uses ML forecasts to estimate optimal starting points for the optimization routine, par-

ticularly important in slow-dynamics processes such as industrial furnaces.
- **Proactive Adaptability**: Enhances the RTO's ability to adapt to changing conditions by predicting future states and disturbances before they occur.
- **Improved Optimization Convergence**: Reduces the time and energy spent in transition states by starting the optimization from a more favorable point, improving efficiency during critical process ramp-ups.
- **Resilience**: The modular design supports the isolation and retraining of ML components without disrupting the core optimization structure.
- **Scalability and Reusability**: The architecture supports growing datasets and sensor networks, while enabling easy transfer to other energy-intensive processes.
- **Technology Agnosticism**: Compatible with a wide range of ML frameworks, optimizers, and deployment environments.

The proposed framework defines a conceptual architecture for augmenting Real-Time Optimization systems with machine learning capabilities. While not implemented directly in the case study, this design supports a structured approach to integrating predictive models into industrial control pipelines, offering a scalable path for future deployments in real industrial settings.

## VI. Experiments and Results

In this section, we present the results obtained for the predictive models evaluated in this study, encompassing representatives from linear models, regularized regressors, ensemble learning techniques, and neural networks. These categories were selected to provide a balanced comparison across different levels of model complexity and interpretability, enabling the identification of approaches best suited to the characteristics of industrial process data.

Table III summarizes the performance of representative models from each methodological category, providing a comparative overview of their effectiveness in estimating the LCV based on historical process data.

TABLE III
Model Comparison

| Model | RMSE (Test) | $R^2$ (Test) |
|---|---|---|
| Random Forest Regressor | 158.45 | 0.85 |
| XGBoost Regressor | 162.78 | 0.84 |
| LightGBM Regressor | 164.12 | 0.83 |
| Gradient Boosting Regressor | 173.03 | 0.82 |
| Neural Network (DNN) | 221.65 | 0.71 |
| CNN-LSTM | 409.40 | 0.61 |
| Polynomial Regression | 320.89 | 0.38 |
| Linear Regression | 349.93 | 0.26 |
| Ridge Regression | 350.04 | 0.26 |

The results demonstrate that among all models evaluated, the **ensemble learning approaches** achieved the best overall performance. The **Random Forest Regressor** emerged as the most accurate model, followed closely by the **XGBoost Regressor** and **LightGBM Regressor**. These ensemble-based models effectively captured complex nonlinear patterns in the data and demonstrated robust generalization across test samples, making them well-suited for predictive tasks in industrial environments.

Within the **neural network category**, the **fully connected Deep Neural Network (DNN)** exhibited competitive performance, confirming its ability to model nonlinear interactions in structured industrial data. In contrast, the **CNN-LSTM hybrid architecture**, despite its higher architectural complexity, underperformed significantly. This outcome suggests that the convolutional-recurrent design did not align well with the structure of the input features, which lack the temporal or spatial dependencies typically required to leverage the strengths of sequence-based architectures.

The **linear and regularized regression models** delivered the lowest predictive performance overall. While **Polynomial Regression** was able to capture some degree of nonlinearity in the data, its performance remained inferior to that of ensemble and deep learning approaches. Both **Linear Regression** and **Ridge Regression** demonstrated limited capacity to model the underlying complexity of the process, confirming that their simplistic assumptions are insufficient for accurate prediction in this industrial context.

These results reinforce that, in the context of this case study, ensemble learning methods—particularly Random Forest and boosting variants—demonstrated superior accuracy and generalization for predictive modeling tasks involving high-dimensional industrial process data. Their ability to model complex, nonlinear relationships proved advantageous for the LCV setpoint prediction problem, where multiple process variables interact in subtle ways. In the specific scenario of continuous combustion furnaces analyzed here, these models contributed to enhanced operational decision-making by providing more accurate initial conditions for the RTO system, thereby reducing fuel consumption, improving energy efficiency, and supporting more sustainable production practices.

### A. *Optimal Hyperparameters*

During the model training process, optimal hyperparameters were identified for each model to enhance their performance. GridSearchCV was employed for classical and ensemble models, while Keras Tuner was used for deep learning architectures. The following configurations yielded the best results:

- **Ridge Regression**: The best-performing regularization strength (`alpha`) was set to 1.0.
- **Polynomial Regression**: The optimal polynomial degree was found to be 2.
- **Random Forest Regressor**: The optimal configuration used 500 estimators with `max_features` set to `sqrt`.
- **Gradient Boosting Regressor**: The best performance was obtained with 500 estimators and a learning rate of 0.2.
- **XGBoost Regressor**: The model achieved optimal results using 200 estimators, a maximum tree depth of 7, a subsample ratio of 1.0, and a learning rate of 0.2.
- **LightGBM Regressor**: The optimal configuration included 200 estimators, a maximum tree depth of 10, and a learning rate of 0.2.
- **Neural Network (DNN)**: The best model architecture featured three layers with 64 units (ReLU), 384 units (tanh), and 416 units (ReLU), and dropout rates of 0.1,

0.4, and 0.1, respectively. The optimizer used a learning rate of 0.001.

- **CNN-LSTM**: The hybrid architecture was optimized with 64 convolutional filters, an LSTM layer with 128 units, a dense layer with 128 units, and dropout rates of 0.2 and 0.1. The model used the Adam optimizer with a learning rate of 0.01.

### *B. Error Analysis and Model Interpretation*

A detailed analysis of the prediction errors for the **Random Forest Regressor**, the best-performing model, was conducted. Figure 6 presents the error distribution and a plot of prediction errors across test samples.

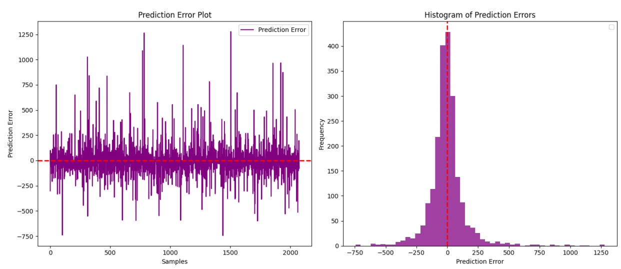


Fig. 6. Prediction Error Plot and Histogram of Prediction Errors for the Random Forest Regressor.

The prediction errors are mostly centered around zero, with few large deviations and a slight tendency to underestimate the LCV. Errors fall within acceptable operational variability, indicating no systematic failures. These results confirm the model's robustness for real-time RTO initialization.

## VII. Conclusion and Future Work

This study demonstrated the effectiveness of integrating machine learning techniques into RTO strategies for industrial applications. A modular framework was proposed to generalize the use of ML for enhancing decision-making in dynamic and data-rich environments. This architecture supports seamless interaction between data acquisition, predictive modeling, and optimization logic, promoting reusability and adaptability across different industrial domains.

A variety of supervised learning models were evaluated, including linear models, ensemble learning techniques, and neural network architectures. Among them, ensemble models consistently provided the best performance for the studied case. In particular, the **Random Forest Regressor** achieved the highest predictive accuracy, reinforcing its suitability for structured industrial process data where complex nonlinearities are present.

The case study focused on continuous combustion furnaces showed that the proposed ML-enhanced RTO strategy has potential to improve energy efficiency and operational performance. Although not yet implemented in a real-time control environment, the methodology lays a solid foundation for deployment in production systems.

Moreover, the results obtained from historical test campaigns (as detailed in Table I) show that the predictive LCV model enables the RTO routine to start from a more efficient initial setpoint. This could lead to a significant reduction in natural gas consumption—yielding a **mean additional savings of approximately 3%** beyond what is achieved by the RTO alone. This estimated saving is projected based on the data observed in the test scenarios summarized in Table I. This reinforces the economic viability of the approach, especially in energy-intensive processes, and confirms the added value of integrating ML into existing RTO strategies.

Based on the findings of this study, the following directions are proposed for future research:

- **Real-time deployment**: Implement the proposed predictive model within the live RTO control system of the continuous furnace, enabling real-time support for setpoint initialization.
- **Framework generalization**: Extend and validate the proposed ML-integrated RTO architecture in other industrial processes that require dynamic and data-driven optimization strategies.
- **Online learning**: Develop mechanisms for continuous learning to allow the predictive models to adapt to process drift and evolving operational conditions.
- **Temporal validation**: Future studies should incorporate time-aware validation methods (e.g., expanding window) and explore temporal feature engineering to better address operational variability and enable models such as LSTM.

In conclusion, this work demonstrates the feasibility of integrating machine learning into Real-Time Optimization through a modular and generalizable framework. The proposed architecture supports predictive setpoint initialization, improving adaptability, responsiveness, and energy efficiency, and enabling intelligent decision-making in energy-intensive processes.

## References

[1] F. A. Silveira and E. H. Shiguemori, "Enhancing real-time optimization with machine learning predictions of lower calorific values," in *Projetos de Pesquisa em Data Science—Estudos de Caso do Curso CEDS IBAPE-PR*. Editora Científica Digital, 2025, vol. 1, pp. 445–466.

[2] A. B. Ige, R. Singh, and A. J. Khan, "Machine learning in industrial applications: An in-depth review and future directions," *International Journal of Multidisciplinary Research and Growth Evaluation*, vol. 6, no. 1, pp. 36–44, 2025.

[3] N. P. Lawrence, A. Grushin, R. Prabhu, Q. Han, and S. J. Qin, "Machine learning for industrial sensing and control: A survey and practical perspective," *arXiv preprint arXiv:2401.13836*, 2024.

[4] M. Mowbray, M. Vallerio, C. Perez-Galvan, D. Zhang, A. Del Rio Chanona, and F. J. Navarro-Brull, "Industrial data science – a review of machine learning applications for chemical and process industries," *Reaction Chemistry & Engineering*, vol. 7, no. 7, pp. 1471–1509, 2022.

[5] X. Zhang, C. S. Edrington, M. Mohanpurkar, T. Kuruganti, G. Subbaraman, and T. Elgindy, "Real-time optimization and control of nonlinear processes using machine learning," *Mathematics*, vol. 7, no. 9, p. 890, 2019. [Online]. Available: https://www.mdpi.com/2227-7390/7/9/890

[6] I.-C. Wang and C.-C. Chien, "Machine learning for industrial optimization and predictive control: A patent-based perspective with a focus on taiwan's high-tech manufacturing," *Processes*, vol. 13, no. 11, p. 2256, 2025. [Online]. Available: https://www.mdpi.com/2227-9717/13/11/2256

[7] R. B. Gopaluni, M. Aghamohammadi, T. Chugh, and R. Yuan, "Modern machine learning tools for monitoring and control of industrial processes: A survey," *IFAC-PapersOnLine*, vol. 53, no. 2, p. 370–381, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896320303827

[8] F. A. Silveira and G. V. Raffo, "Optimal control of fuel gases on continuous annealing furnaces of steel coil strip," 2017.

[9] J. Mount and N. Zumel, *Practical data science with R*. Simon and Schuster, 2019.