1st Latin-American Workshop on Structural Health Monitoring
LATAM-SHM 2023
December, 5-8 2023

TEMPERATURE FORECASTING AGAINST SENSORS FAILURES IN AN ELECTRIC ARC FURNACE.

Diego Godoy[1], Bernardo Rueda[2], Juan Romero[2], Jersson X. Leon-Medina [3] and Diego Tibaduiza[1]*

[1] Departamento de Ingeniería Eléctrica y Electrónica, Universidad Nacional de Colombia, Cra 45 No. 26-85, Bogotá 111321, Colombia; Diego (D.F.G-R.); dtibaduizab@unal.edu.co (D.T.)

[2] South32-Cerro Matoso S.A., Km 22 Highway SO Montelibano 234001, Córdoba, Colombia; Bernardo.S.Rueda@south32.net (B.R.); JuanAlonso.Romero@south32.net (J.R.)

[3] Control, Data and Artificial Intelligence (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya (UPC), 08019 Barcelona, Spain

Key words: Electric Arc Furnace, Recurrent Neural Networks, Sensor Failures, Structural Health Monitoring, Time Series Forecasting.

Abstract. *Structural Health Monitoring in the industry is necessary to ensure worker safety, preservation, and operational efficiency of the structures used in the different processes. A robust and well-maintained structure minimizes risks, extends equipment lifespan, and prevents costly production disruptions. Additionally, proper structural health contributes to process stability, energy efficiency, and compliance with industrial regulations, ensuring a safe and efficient work environment. In the mining industry, highly complex equipment such as Electric Arc Furnaces (EAFs) are used for melting metals and other materials. These furnaces consist of casings, refractory walls, electrodes, and other key components that endure high temperatures and mechanical loads during the melting process. These furnaces have a robust sensor network that monitors and controls important variables in the melting process. Analytical techniques and artificial intelligence models have been developed to predict key variables, such as internal furnace temperature, based on other variables. However, issues can arise in information availability and prediction model performance due to possible sensor network failures. To contribute in the solutions to this problem, this work analyzes the performance of a recurrent neural network model (GRU) used in predicting the temperature of an Electric Arc Furnace in the Colombian mining industry, and two cases of sensor network failures were studied in furnace monitoring: abrupt failure (sensor calibration issue) and noise-induced failure. A comparison was made between the model's performance and the number of sensors with failures present, establishing the behavior and influence of these failures on temperature prediction. This study is important in the application field, as it showcases the outstanding performance of these new technologies in industrial use. It highlights the potential benefits and impact that these advancements could have on various industries.*

1 INTRODUCTION

Complex industrial processes such as ferronickel smelting require permanent monitoring of the temperature of the refractory lining to determine hot zones in the furnace and to be able to execute control actions to reduce the temperatures in these hot zones [1] and prevent problems that can affect the regular operation and to the workers associated to this operation. The knowledge of changes in these temperatures because of different changes in the operation conditions with some hours or days before damages can appear is a need. From this point of view, strategies based on historical data for predicting these variables are necessary.

Monitoring and predicting the temperature of the refractory lining in furnaces has been carried out in recent years using machine learning algorithms. In 2021, Leon-Medina et al., [2] developed a multivariate temperature prediction model based on a Deep recurrent neural network that combined a GRU layer with a dense one. In a time-series approach, 49 input variables were used to predict 2 hours in the future the behavior of 16 temperature outputs at different locations of an electric arc furnace. The average root mean square error (RMSE) value was 1.19 °C in the test set. Later, in the work developed by Godoy et al. [3] an attention mechanism in a Deep recurrent neural network was used to predict the behavior of 76 thermocouples distributed radially in an electric arc furnace. It was observed that as the number of thermocouples to predict increases, the average RMSE also increases. In this case, a value of 3.89 °C was found in the test set when 76 thermocouples were predicted. A study of the deterioration of the temperature prediction model over time was carried out, determining that it is advisable to retrain the Deep learning model every year to maintain the error value at the aforementioned magnitude. On the other hand, due to the continuous monitoring that must be carried out, in 2022, Leon-Medina et al. [4] developed a temperature prediction model that works with stream data. Unlike traditional batch learning, the stream learning modality takes advantage of the continuous arrival of new instances to train a multi-target regressor online. This study compared two types of multi-target regression trees based on Hoeffding trees. These two models were the Stacked Single-Target Hoeffding tree regressor (SST-HT) and the Incremental Structured Output Prediction tree regressor (iSOUP-Tree). The models were used to predict the behavior of 12 thermocouples distributed radially in different sectors of an electric arc furnace. The best model (SST-HT) found an average mean absolute error value of 3.719 °C.
As it is mentioned some works have developed to monitor and forecast some variables in the ferronickel production process; however, one of the challenges in the use of these data is the ability of models to cope with unforeseen situations and make accurate decisions under adverse conditions which is very common in industrial equipment in operation. As a contribution to this problem, this paper considers two common failures in the data captured by sensors to determine how the model is affected in predicting the variables.
This paper is organized as follows: the second section includes the theoretical background and methods used in the approach, the third section includes the results, and finally, the conclusions are summarized in the fourth section.

2 METHODS AND THEORETICAL BACKGROUND

2.1 Electric Arc Furnace Dataset

The dataset for training and validating the model comprises information collected over five years, encompassing 177,312 instances for 49 distinct attributes. These data were gathered from one of the Electric Arc Furnaces (EAF) used by Cerro Matoso S.A. (CMSA) for producing ferro-nickel. CMSA is a subsidiary of South 32 located in Montelibano (Cordoba) in Colombia. The data collection frequency was set at 15-minute intervals, spanning 1,847 days. The model uses various input variables such as electrode current, voltage, arc characteristics, power, calcine feed, chemical composition of the calcine, relative electrode positioning, and an array of 16 thermocouples which were also used as output variables for predictive purposes.

Diverse data preprocessing steps were undertaken to identify anomalous patterns within the employed variables. These data preprocessing steps start with removing duplicate data or variables, treating empty and null values, treating unique values, encoding strings, and removing negative temperatures. In addition, variables with high variance and zero variance were also removed from the dataset. From this preprocessing, 49 variables were used for model training and testing because they do not exhibit any abnormal behaviors [5], the proportion used for training and test sets are 80% - 20% respectively. Figure 1 shows the steps in the methodology, as it can be observed RMSE metrics and GRU models are used. The following subsections will introduce these concepts.

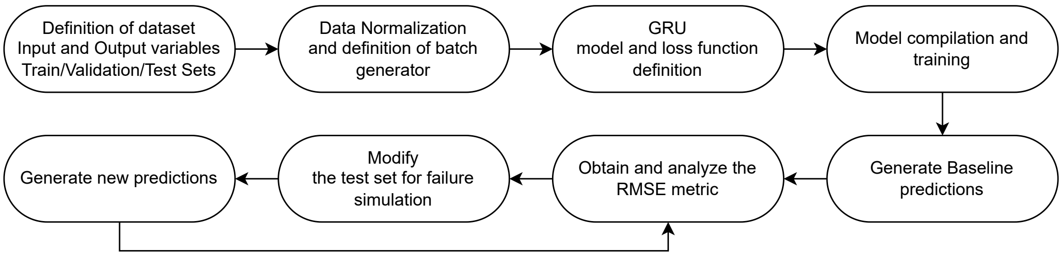

Figure 1: Step by step for the GRU model design and prediction for the failures simulation environment.

2.2 Multivariate Time Series GRU Forecasting Model

Gated Recurrent Unit (GRU) Neural Network models have proven highly effective in predicting multivariate time series due to their inherent characteristics of sequential learning and long-term memory, designed to address long sequence problems while maintaining superior computational efficiency. The key advantage of GRUs lies in their ability to capture complex temporal dependencies in sequential data, making them ideal for modeling dynamic relationships between multiple variables in time series. In particular, GRUs are well equipped to handle the high dimensionality of multivariate time series and can learn meaningful representations of the interactions between variables over time. Their internal gates allow them to decide when to refresh memory and when to forget old information, making it easier to capture short- and long-term patterns in the data. This results in highly adaptive models that can fit a wide range of time structures and make accurate forecasts in situations where the relationships between variables can be non-linear and highly dynamic [6].

The GRU model architecture used in this document consists of a layer of 300 GRU cells and a dense layer of 16 neurons that correspond to the 16 thermocouples that need to be predicted. The Adam optimization method was used to optimize the model, which is a stochastic gradient descent (SGD) method based on adaptive estimation. The selection of this model results from the findings and work done in paper [3] where they compare the performance of various RNN models for forecasting the temperature in the same context. Other training details correspond to the batch size of 250 with a sequence length of 1152 and the adaptive learning rate from 0.001 to 0.0001.

2.3 RMSE as an evaluation metric for prediction comparison

The Root Mean Square Error (RMSE) is a fundamental metric in the evaluation of Deep Learning models; this metric measures the average difference between the predicted values and the real values; also has the advantage that it penalizes wrong predictions more significantly since larger differences have a greater impact on the final result, when comparing the results between different models, a lower value indicates that the model is capable of making more accurate predictions and is therefore preferable. This metric is calculated using the following function:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{y_i} - y_i)^2} \tag{1}$$

Where $\mathbf{n}$ are the number of data points in the data set, $\mathbf{y_i}$ represents the actual or observed value at data point $\mathbf{i}$, $\hat{\mathbf{y}}_\mathbf{i}$ represents the value predicted by the model at data point $\mathbf{i}$ and $\sum$ denotes the sum over all data points, that is, the squared errors for each data point are added.

2.4 Sensor Failures

Sensor failures represent significant challenges in numerous fields that depend on accurate and reliable measurements; these can manifest themselves in different ways, from constant readings at a fixed value due to a frozen sensor to erratic or even completely out-of-range measurements due to faulty internal components. These flaws can be challenging to detect and can substantially impact the quality of the data and the decisions based on it [7]. Thermocouples, which are typical temperature-measuring devices, can also experience failures. This includes issues such as degradation of the thermoelectric wire material, loose or corroded connections, and deviations in expected thermoelectric characteristics, among others. Two types of failures identified during the initial study of the data set will be further detailed below.

2.4.1 Abrupt Failure

Abrupt failure occurs in sensors when the readings remain constant at a fixed value; it represents a critical situation that can significantly impact the measurements' precision and reliability. This failure is typically the result of a sudden interruption in the sensor's ability to capture and transmit accurate data. When the readings remain constant at a fixed value typically the last measured value as seen in the figure 2, this may be due to a physical blockage of the sensor or a malfunction in its internal electronics. Identifying and addressing these abrupt failures is essential to ensure reliable measurements and avoid potential negative consequences in specific applications where data accuracy is critical [8].

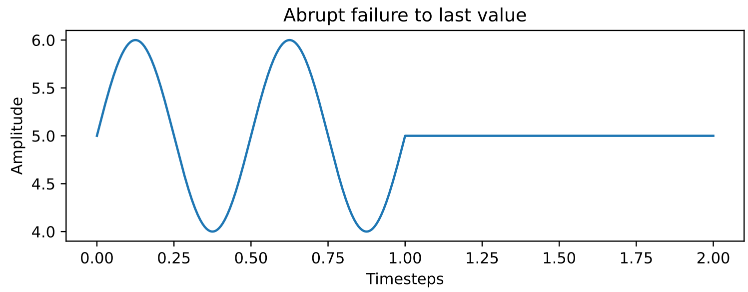

Figure 2: Representation of a sinusoidal signal that has an abrupt failure at the last recorded value.

2.4.2 Noise-induced Failure

Induced noise in the context of signals refers to any unwanted interference or disturbance mixed with the signal of interest during data acquisition or transmission steps. This phenomenon can manifest itself as random fluctuations or unsystematic variations in the signal that can distort or make it difficult to interpret the information that the signal carries accurately. In essence, noise adds uncertainty and error to measurements or data, as seen in Figure 3, making observations less reliable.

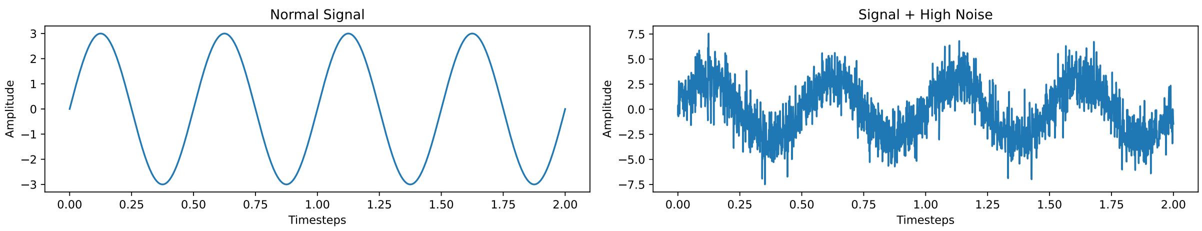

Figure 3: Representation of two sinusoidal signals, without noise and with high noise.

Noise-induced failures in sensors add an additional layer of complexity to detecting and correcting problems in measurement instrumentation. Noise, which can arise from multiple sources such as electromagnetic interference, electrical fluctuations, or even environmental vibrations, can corrupt sensor signals and lead to inaccurate or erratic measurements [9].

3 RESULTS

This section presents the results obtained through the implementation and evaluation of the Gated Recurrent Unit (GRU) neural network model in detail in the context of a wall temperature monitoring and diagnosis study of an electric arc furnace. The results presented here are based on the application of this model in three different scenarios, addressing an initial case without failures and two cases of sensor failures. The first case corresponds to the abrupt failure of the sensors and the second failure case presents a noisy environment. Throughout this section, quantitative analyses will provide information on the detection and diagnosis capacity of the GRU model, allowing a deep understanding of its performance in varied and challenging situations.

3.1 Baseline Prediction (without Sensor Failures)

Firstly, the temperature prediction was carried out in the 16 thermocouples without any type of failure in the test set as seen in figure 4, with the objective of having initial RMSE evaluation metrics that allow comparison to be carried out when failures begin to be included in the data. Figure 4 shows the good performance of the model when making predictions since in its entirety for the 16 thermocouples there is no great difference with the real behavior data, which can be verified from the RMSE obtained for each thermocouple, as evidenced in table 1, where on average a very acceptable RMSE value was also obtained.

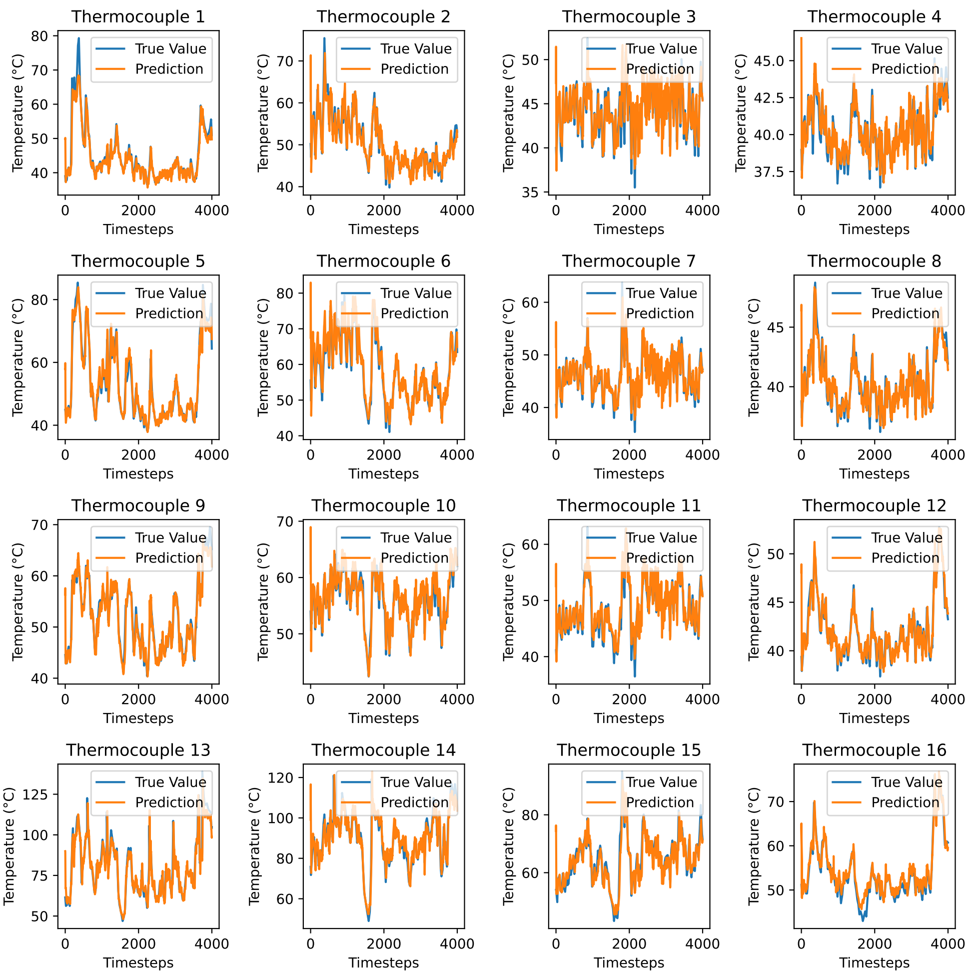

Figure 4: Prediction versus true behavior for the test set of the 16 thermocouples without failures.

Table 1: RMSE results in °C and the average for the 16 thermocouples in the train and test sets without failures.

T	Train	Test	T	Train	Test	T	Train	Test	T	Train	Test
T1	0.97	1.89	T5	1.63	2.27	T9	1.15	1.31	T13	3.58	4.91
T2	1.26	1.63	T6	1.70	2.39	T10	1.23	1.38	T14	3.38	3.99
T3	0.96	1.27	T7	1.08	1.51	T11	1.10	1.59	T15	2.34	2.75
T4	0.69	0.70	T8	0.77	0.73	T12	0.78	0.72	T16	1.38	1.57
				Average:		Train:	1.50		Test:	1.91	

3.2 Prediction case #1 (Abrupt Failure to last value)

The test set of the thermocouple #9 was modified to take the last data from the training set and remain constant over time at that value, as can be seen in figure 5.

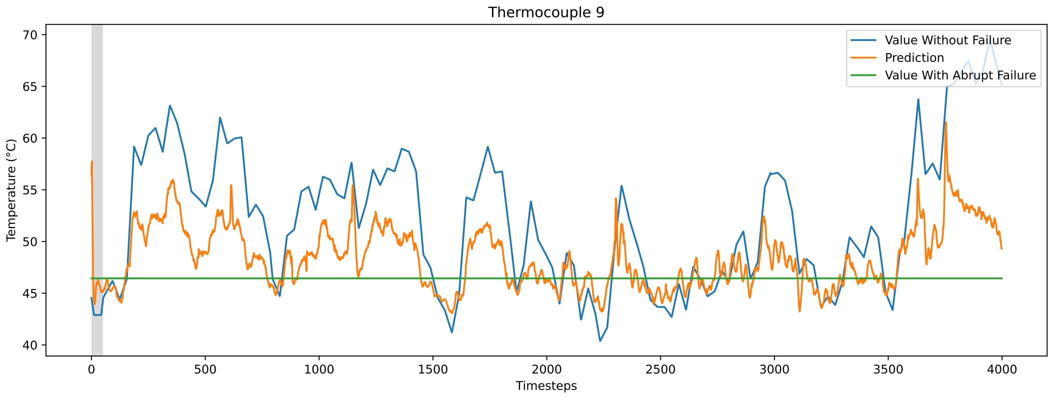

Figure 5: Prediction at thermocouple #9 when an abrupt failure sets its value to the last recorded temperature.

In figure 5 it is possible to observe that the prediction made by the model for this thermocouple in the case of failure, the prediction got worse compared to the base case; however, despite not having real information in that moment from the thermocouple, the model is capable of making a prediction very close to the real behavior, identifying its variations and trends. Now, figure 6 shows the behavior of the remaining 15 thermocouples when there is a failure in one of them and as can be seen the predictions made for these are almost not affected, which shows the good performance that the model has against this type of abrupt failures.

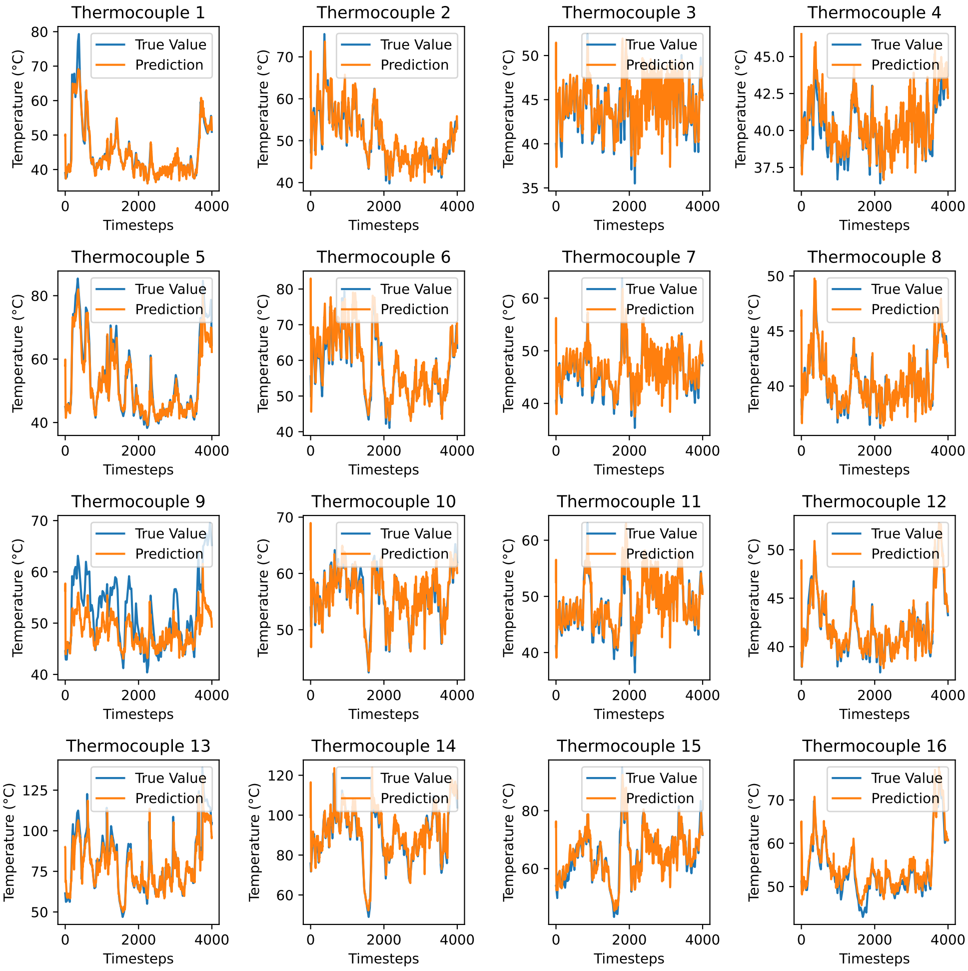

Figure 6: Prediction versus true behavior for the test set of the 16 thermocouples when an abrupt failure set the value of the thermocouple #9 to the last recorded temperature.

Table 2 shows the variation in the results of the evaluation metric for all the thermocouples. It can be concluded that errors are similar to the thermocouples that were not affected due to the failure. However, the most significant variation was that of thermocouple #9, which went from an RMSE of 1.31 to 5.61.

Table 2: RMSE comparison in °C for the 16 thermocouples with one abrupt failure in the thermocouple #9 versus the baseline results.

T	New Test	Original Test	T	New Test	Original Test
T1	1.79	1.89	T9	5.61	1.31
T2	1.65	1.63	T10	1.49	1.38
T3	1.39	1.27	T11	1.68	1.59
T4	0.84	0.70	T12	0.74	0.72
T5	3.18	2.27	T13	6.02	4.91
T6	2.52	2.39	T14	4.02	3.99
T7	1.76	1.51	T15	2.72	2.75
T8	0.75	0.73	T16	1.72	1.57
Average	New Test:	2.37	Original Test:	1.91	

Finally, the behavior of the RMSE was analyzed as more thermocouples presented the abrupt failure and as can be seen in the table 3, as the number of thermocouples with constant values increases, the RMSE value tends to increase.

Table 3: Behavior of the RMSE value in °C related to the number of thermocouples that fail abruptly to the last value.

# T. with abrupt failures	Average RMSE	# T. with abrupt failures	Average RMSE
1	2.37	5	4.12
2	2.51	6	4.78
3	2.79	7	5.50
4	2.90	8	5.91

3.3 Prediction case #2 (Failure by Noise)

As for the second type of failure, the reference test set for the thermocouple #9 was selected again, and a simulation and prediction scenario was carried out adding Gaussian noise to the original signal as shown in the figure 7

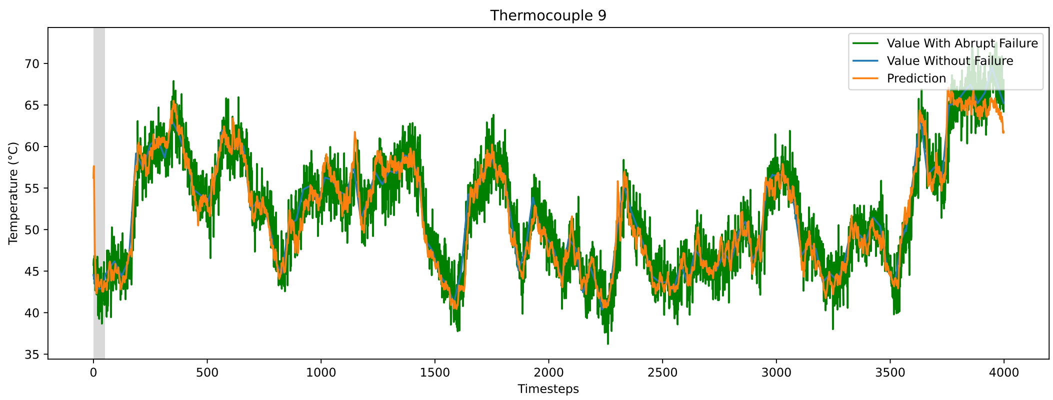

Figure 7: prediction when the signal measured by the thermocouple #9 is affected by high noise.

Table 4: RMSE comparison in °C for the 16 thermocouples with high noise present in the thermocouple #9 versus the baseline results.

T	New Test	Original Test	T	New Test	Original Test
T1	1.91	1.89	T9	1.46	1.31
T2	1.66	1.63	T10	1.44	1.38
T3	1.41	1.27	T11	1.72	1.59
T4	0.75	0.70	T12	0.74	0.72
T5	2.34	2.27	T13	5.13	4.91
T6	2.42	2.39	T14	4.02	3.99
T7	1.68	1.51	T15	2.83	2.75
T8	0.77	0.73	T16	1.58	1.57
Average	New Test:	1.99	Original Test:	1.91	

In the tables 4 and 5 it can be seen the excellent performance of the model concerning failures due to noise present in the input signal. This suggests that even when a single thermocouple presents an error of this type, the RMSE value at an individual level increases only minimally. Moreover, after each increase in the number of thermocouples with the failure, the RMSE value also increases very slowly. Even with eight thermocouples presenting the error, the predictions still have a very acceptable value.

Table 5: RMSE comparison in °C between the number of thermocouples that fail abruptly.

# T. with abrupt failures	Average RMSE	# T. with abrupt failures	Average RMSE
1	1.99	5	2.17
2	2.03	6	2.18
3	2.10	7	2.20
4	2.16	8	2.22

4 CONCLUSIONS

- The evaluation of predictive models in the event of input failures is crucial to ensure their reliability and adaptability in real-world applications. The ability of models to handle unforeseen situations and make accurate decisions under adverse conditions is paramount. By exposing models to simulated failure scenarios, weaknesses can be identified, and mitigation strategies can be developed, thus contributing to the robustness and safety of prediction-based systems. These tests are ultimately essential to enhance preparedness and confidence in the effectiveness of predictive models in various critical contexts.

- The GRU model seems to have performed well in the two simulated failure cases. The results were promising, particularly concerning abrupt failures that maintain the measurement at a constant value close to or within the normal measurement range of the sensor, as well as noise-induced failures in the measurement. The model could predict the output correctly without being affected by the noise in the input data.

- In this study, the impact of sensor failures on the predictions made by a neural network model was observed. The study analyzed both the individual level of the sensor and the collective level of the other sensors that comprise the system.

REFERENCES

[1] D. Zhou, K. Xu, J. Bai, and D. He, "On-line detecting the tuyere coke size and temperature distribution of raceway zone in a working blast furnace," *Fuel*, vol. 316, p. 123349, 2022.

[2] J. X. Leon-Medina, J. Camacho, C. Gutierrez-Osorio, J. E. Salomón, B. Rueda, W. Vargas, J. Sofrony, F. Restrepo-Calle, C. Pedraza, and D. Tibaduiza, "Temperature prediction using multivariate time series deep learning in the lining of an electric arc furnace for ferronickel production," *Sensors*, vol. 21, no. 20, p. 6894, 2021.

[3] D. F. Godoy-Rojas, J. X. Leon-Medina, B. Rueda, W. Vargas, J. Romero, C. Pedraza, F. Pozo, and D. A. Tibaduiza, "Attention-based deep recurrent neural network to forecast the temperature behavior of an electric arc furnace side-wall," *Sensors*, vol. 22, no. 4, p. 1418, 2022.

[4] J. X. Leon-Medina, J. Camacho-Olarte, B. Rueda, W. Vargas, L. Bonilla, J. Ruiz, J. Sofrony, J. A. Guerra-Gomez, F. Restrepo-Calle, and D. A. Tibaduiza, "Monitoring of the refractory lining in a shielded electric arc furnace: An online multitarget regression trees approach," *Structural Control and Health Monitoring*, vol. 29, no. 3, p. e2885, 2022.

[5] J. Camacho-Olarte, J. E. S. Torres, D. A. G. Jimenez, J. X. L. Medina, R. C. G. Vargas, D. A. V. Cardenas, C. Gutierrez-Osorio, B. Rueda, W. Vargas, D. A. T. Burgos *et al.*, "A data cleaning approach for a structural health monitoring system in a 75 mw electric arc ferronickel furnace," vol. 2, no. 1, p. 21, 2020.

[6] Y. Wang, M. Liu, Z. Bao, and S. Zhang, "Short-term load forecasting with multi-source data using gated recurrent unit neural networks," *Energies*, vol. 11, no. 5, 2018. [Online]. Available: https://www.mdpi.com/1996-1073/11/5/1138

[7] S. Hussain, M. Mokhtar, and J. M. Howe, "Sensor failure detection, identification, and accommodation using fully connected cascade neural network," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1683–1692, 2015.

[8] P. A. Samara, G. N. Fouskitakis, J. S. Sakellariou, and S. D. Fassois, "A statistical method for the detection of sensor abrupt faults in aircraft control systems," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 4, pp. 789–798, 2008.

[9] I. A. Khovanov, N. A. Khovanova, and P. V. E. McClintock, "Noise-induced failures of chaos stabilization: Large fluctuations and their control," *Phys. Rev. E*, vol. 67, p. 051102, May 2003. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.67.051102