UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS

Visión Electrónica

https://revistas.udistrital.edu.co/index.php/visele

VISIÓN ELECTRÓNICA

A RESEARCH VISION

Reliability assessment in complex systems using updated fault rate estimation

Evaluación de confiabilidad en sistemas complejos mediante la estimación actualizada de la tasa de fallas

Sebastián Aguirre-Vargas [1], *Mauricio Holguín-Londoño* [2], *Germán A. Holguín-Londoño* [3]

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:
Enviado: 12/12/2023
Recibido: 27/12/2023
Aceptado: 18/03/2024

Keywords:
Electrical systems
Estimation
Failure rate
Reliability
Time series

Palabras clave:
Confiabilidad
Estimación
Series de tiempo
Sistemas eléctricos
Tasa de fallas

ABSTRACT

The expansion and operation of the electrical network plays a significant role in ensuring service quality. Keeping the components operational requires an integrated, optimal and sustainable system focused on reliability analysis to ensure the efficiency and availability in complex systems. This paper explores a methodology to improve reliability analyses on power system components, using univariate data analysis and time series regression to develop an updated fault rate estimation. An exploratory data analysis is developed to understand the behavior and uncertain nature of the variables under study. The Simple Exponential Smoothing (SES) model and the Autoregressive Integrated Moving Average (ARIMA) are implemented to analyze and estimate the behavior of the failure rate of some electrical distribution transformers over time, where the variability of reliability is observed as the failure rate varies over time, due to external factors.

RESUMEN

La expansión y operación de la red eléctrica trae consigo grandes retos para garantizar la calidad del servicio. Mantener operativos los activos requiere de un sistema integrado, óptimo y sostenible enfocado en los análisis de confiabilidad para asegurar la eficiencia y disponibilidad de los sistemas complejos. Este documento, basado en un enfoque analítico, explora cómo mejorar los análisis de confiabilidad en los componentes de los sistemas eléctricos de potencia, mediante el análisis de datos univariante, así como metodologías de regresión de series temporales para desarrollar una estimación actualizada de la tasa de fallas. Inicialmente, se desarrolla un análisis exploratorio de datos para comprender el Autoregressive Integrated Moving Average (ARIMA), y el Suavizado Exponencial Simple (SES), además,

1 Electrical engineer, Universidad Tecnológica de Pereira (UTP), Colombia. Master's student in electrical engineering, UTP, Colombia. Current position: Engineering and budget analyst in Ergo Energía, Colombia. E-mail: sebastian.aguirre1@utp.edu.co.

2 Electrical engineer, Universidad Tecnológica de Pereira (UTP), Colombia. Ph.D. in enginnering and Master in electrical engineering, UTP, Colombia. Current position: Full Professor, UTP, Colombia.

3 Electrical engineer, Universidad Tecnológica de Pereira (UTP), Colombia. Ph.D.'s student in enginnering and Master in electrical engineering, UTP, Colombia. Current position: Full Professor, UTP, Colombia. E-mail: gahol@utp.edu.co

Citar este artículo como: S. Aguirre-Vargas, M. Holguín-Londoño, G. A. Holguín-Londoño, "Reliability assessment in complex systems using updated fault rate estimation", Visión electrónica, vol. 18, no. 1, pp. 23-31, january-june 2024, https://doi.org/10.14483/22484728.22415

analizar y estimar el comportamiento de la tasa de fallas de algunos transformadores de distribución a lo largo del tiempo, donde se observa la variabilidad de la confiabilidad a medida que la tasa de fallas presenta variaciones en el tiempo debido a factores externos.

1. Introduction

Electricity plays a very important role in the development of a country's economy due to the impact of electricity in most modern activities. It is of great interest to implement reliability analysis strategies in the planning, design, operation and maintenance of the electricity network infrastructure, for an adequate electricity supply [1]. Understanding the state of complex systems through reliability indicators can be decisive in the decision making for proactive actions that mitigate failures or interruptions, decrease the high electrical load levels and improve network voltage levels, through high-impact projects and preventive maintenance scheduling [2].

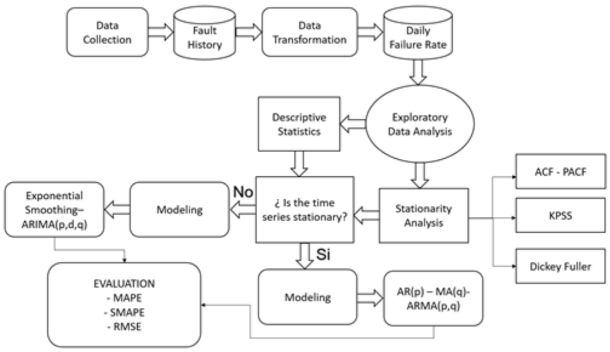
The implementation of this type of management is challenging because it requires the study of complex systems, which consist of many elements that interact with each other. The implementation of efficient asset management in power electrical systems is important due to the dynamic network configuration, quality of service, operating environment, and operating costs, among others [3] [4]. Factors such as projected demand, due to the expansion of the industrial and residential sectors, and the progress in the energy transition through Non-conventional Renewable Energy Sources (NCRES) makes it essential to plan investment projects for replenishment and expansion of assets, acquisition of new technologies and continuous improvement of electricity networks [5]. However, budgetary limitations in investment, operating and maintenance costs of the countries have led to an increasing rate of failures in power electrical systems, significantly affecting the productivity of assets and the quality of energy service [6].

The failure rate is one of the most important parameters in the statistical reliability models. In consequence, it is necessary a correct estimation. The use of prediction in time series as in the failure rate, has had great impact in the financial sector. Thus, in [7] it is presented a methodology to predict short-term stock prices through historical data with Recurrent Neural Networks (RNNs). In the home services sector, the Long Short-Term Memory (LSTM) model is implemented with the Monte Carlo method to predict the failure rate of water distribution networks [8]. In the electrical sector, the failure rate of electric meters is calculated using a Bayesian hierarchical approach to establish certain reliability requirements [9]. This paper proposes a methodology for the up-to-date estimation of the distribution transformer failure rate, using time series modelling methods [10] [11] [12] [13] [14] [15] [16].

2. Methodology

In this paper, a failure rate prediction model is presented to evaluate the variation of reliability of complex equipment or systems over time. An exploratory analysis of the database is proposed based on descriptive statistics and methods of verifying stationarity, such as the Augmented Dickey-Fuller (ADF) test or the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. In the study case, due to the nature of the data, the failure rate prediction model is implemented using an Autoregressive Integrated Moving Average (ARIMA) model and exponential smoothing methods, to finish with their respective evaluation using MAPE, SMAPE and RMSE metrics. Figure 1 shows he general framework of the proposed analysis, which can be divided in several steps.

Figure 1. General framework of analysis.



Source: own.

- Descriptive statistics. It is essential in data analysis, as they facilitate the analysis and description of the behavior of a random variable. Among the most important are:

o **Mean**: Geometric center of the data.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Where n represents the number of samples.

o **Standard deviation**: Measure the variability of the data with respect to the mean.

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

o **Asymmetry coefficient**: Measure of symmetry of the data with respect to its center. It is also generally called the coefficient of skewness.

$$AC = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{n * \sigma^3}$$

- The Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are used to define the stationarity of a time series. Augmented Dickey Fuller's test model is:

$$\Delta y_t = \alpha + p * y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_p \Delta y_{t-p} + \epsilon_t$$

- Exponential smoothing methods are used to model non-stationary time series. The simple exponential smoothing method is defined by:

$$\widehat{y_{T+1|T}} = \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots$$

Where α represents the smoothing parameter.

The evaluation of the prediction of the model is presented by metrics of the symmetric mean absolute error (SMAPE), the mean absolute error (MAPE) and the mean squared error (MSE).

3. Experimental Verification and Results

3.1. Description of data and descriptive statistics

By preprocessing and transforming the historical fault data of an electrical distribution transformer, of

Figure 2. Daily Failure Rate.

Source: own.

a network operator, connected to the Local Distribution System, the database implemented in the analysis and modeling of this study is obtained. It consists of the mobile failure rate, that is, the value of the failure rate updated periodically. Each time period in this database is a mobile quarter, where its value is adjusted according to changes over time. By performing the daily calculation, removing the oldest value from the dataset and selecting the most recent date, we obtain 1593 estimates from April 1, 2019 to August 21, 2023, as shown in Figure 2.

The failure rates are obtained considering the different incidences that are observed in the transformer, and that are considered in circular CREG 063 of 2019 as causes of planned and unplanned events, that is, the causes of unavailability that have their origin by atmospheric conditions, preventive maintenance, affected by failures, animals or branches on networks, network extensions, vandalism, among others. The above, to perform a reliability analysis of the equipment, taking into account the requirements and guidelines that network operators in Colombia must meet

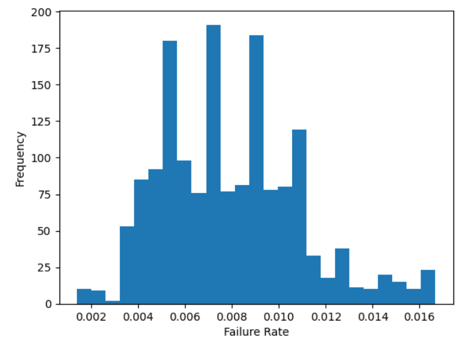
Table 1. Descriptive statistics.

Mean	Max	Min	Standard deviation	Kurtosis	Asymmetry
0.007949	0.016667	0.001389	0.002956	0.1168	0.585

Source: own.

The results of the descriptive statistics are shown in Table 1, where the failure rates range from 0.001389 to 0.0166, the overall mean of the data is 0.007949 and the standard deviation is 0.2956%. The asymmetry coefficient indicates a slight positive asymmetry, which is reflected in the histogram in Figure 3, while the kurtosis indicates, due to its near-zero value, a mesocurtic distribution. In addition, the mean is close to the median, suggesting that the mean is a good indicator of the data center. These statistics reflect the homogeneity of the data.

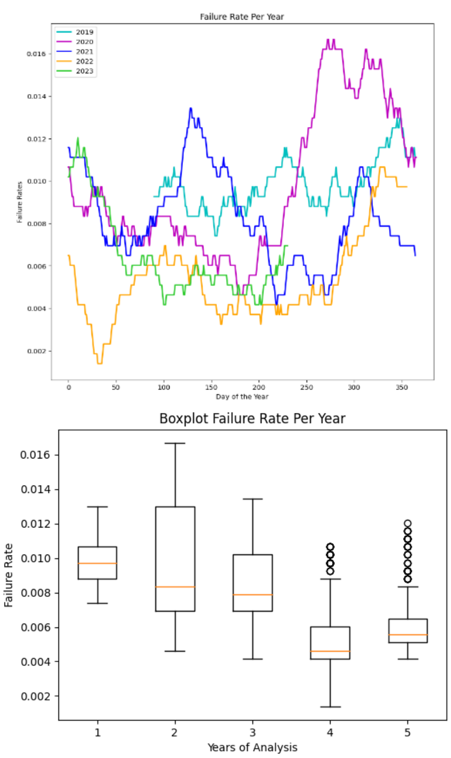
Figure 3. Failure Rate Histogram



Source: own.

When analyzing the annual failure rates, their statistical properties present variations in time, whose causality is inferred due to the operating conditions of the network, climate and other external and internal factors that greatly influence the failure of an equipment or system. Figure 4 shows the failure rates by year.

Figure 4. Failure rate per year and BoxPlot.



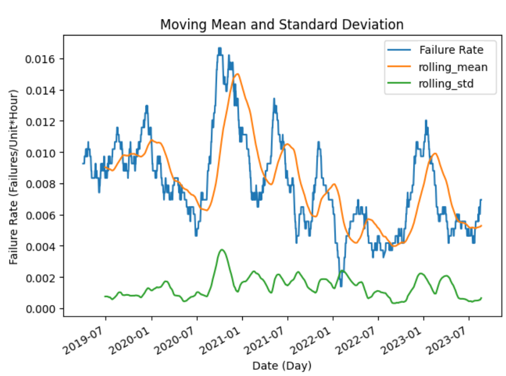
Source: own.

In the first semester of 2020, for example, the failure rate tends to decrease gradually, reaching a minimum value of 0.0045, from which an upward trend is observed, reaching a maximum value of 0.0166 in the second semester, determining the highest failure rate of the year 2020 and the analysis horizon. In 2021, a large peak of 0.0134 appears in the first half of the year, where subsequently a decreasing trend results in the minimum failure rate value in the year of 0.0041. During the years of analysis, the failure rate has fluctuated over time, where there are decreasing trends in the first quarter of the year and increasing trends at the end of the year. In the years 2019, 2020 and 2021 there is a greater instability of the quality of the energy service compared to the years 2022 and 2023, since the failure rate in the last two years, that represent the incidents in the electrical transformer, are with a lower dispersion of the data and generally present a lower average of the order of 0.0045 compared to the 0.009 of the previous years.

3.2. Stationarity

Evaluating the stationarity of a time series is a necessary analysis, because it is a desired characteristic for this type of data, since the modeling result presents in general a better adjustment and feasibility of the prognosis. The mean and moving standard deviation are shown in Figure 5.

Figure 5. Mean and Moving Standard Deviation.



Source: own.

The graph above shows a growing and decreasing nonlinear pattern in the moving average and a standard deviation with variabilities that increase and decrease over time. The Augmented Dickey-Fuller (ADF) test is used to check the stationarity of the time series; for which, if the p-value is less than 0.05, the series is stationary. The test results for the failure rate series are presented in Table 2.

Table 2. Descriptive statistics.

T statistic	-2.684042
p-value	0.076841
N° of Lags Used	8
N° of Observations Used	1584
Critical value (1%)	-3.434485
Critical value (5%)	-2.863366
Critical value (10%)	-2.567742

Source: own.

The null hypothesis, H0, of this test shows that the time series is not stationary. In Table 2, the T statistic is greater than 5% of the significance level and the p-value of 0.076841 is greater than 5% of the standard confidence level. There is no reason to reject the null hypothesis and it is concluded that the series is non-stationary.

Table 3. Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test.

Test statistic	1.825593
p-value	0.01
# of Lags Used	26
Critical value (1%)	0.739
Critical value (5%)	0.463
Critical value (10%)	0.347

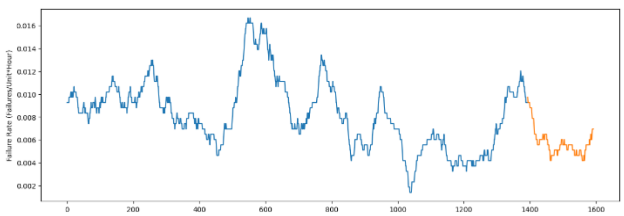
Source: own.

To confirm the non-standard nature of the time series, the Kwiatkowski-Phillips-Schmidt-Shin test is implemented. The null hypothesis, H0, of this test shows that the time series is stationary. In Table 3, the T statistic of 1.825593 is less than 5% of the standard confidence level. The hypothesis is rejected, and it is concluded that the series is non-stationary.

3.3. Modeling

To compare model predictions with actual failure rate values, the division of data into training and test sets is initially implemented, as presented in Figure 6.

Figure 6. **Training and test data.**

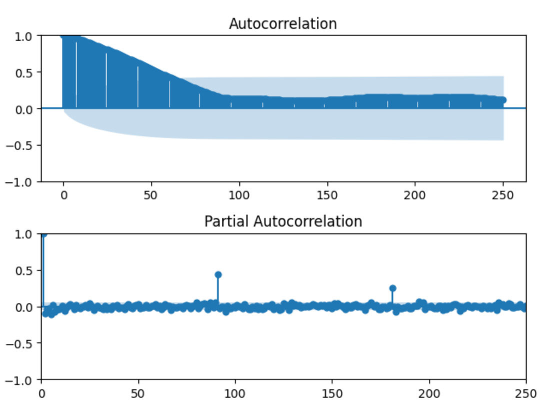


Source: own.

3.3.1. ARIMA model

To determine the order of the ARIMA model, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) are used to determine p and q. The autocorrelation and partial autocorrelation diagrams of the failure rates are presented in Figure 7. In this, 2 autocorrelations are observed with an important significance level, so it is defined initially (2,1,1) as order of the ARIMA model.

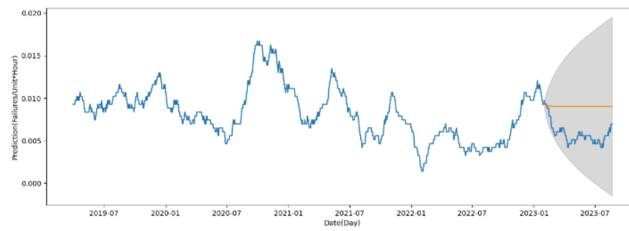
Figure 7. **Autocorrelation function and Partial autocorrelation function.**



Source: own.

When optimizing the model with respect to the Akaike Information Criterion (AIC) metric, the ARIMA model (4,1,0) is suggested with the lowest AIC. The prediction in the test data set is presented in Figure 8, with its respective confidence interval. The value of the resulting flat line is 0.006..

Figure 8. **Training and test data.**

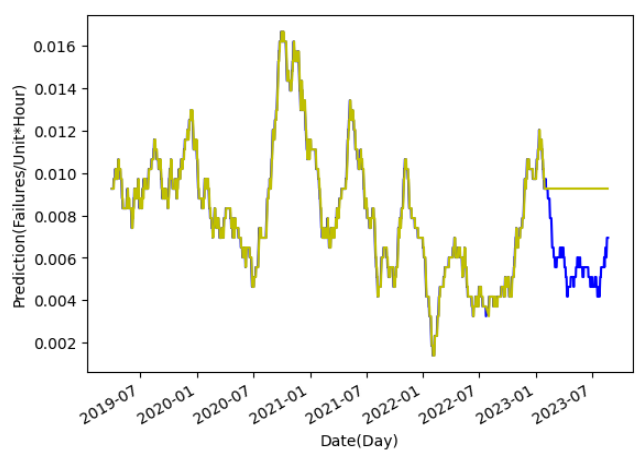


Source: own.

3.3.2. Exponential Smoothing Methods - Exponential Simple Smoothing (SES)

This model is the simplest of its kind, its forecast function is flat and has the same prediction value during its analysis horizon. When performing the simulation, the optimal resulting model is obtained with an alpha equal to 0.995. This allows the prediction corresponding to Figure 9. The forecast is a flat line with a value equal to 0.009259 which remains constant during the analysis interval of the test set.

Figure 9. **Simple Exponential Smoothing Model Prediction.**

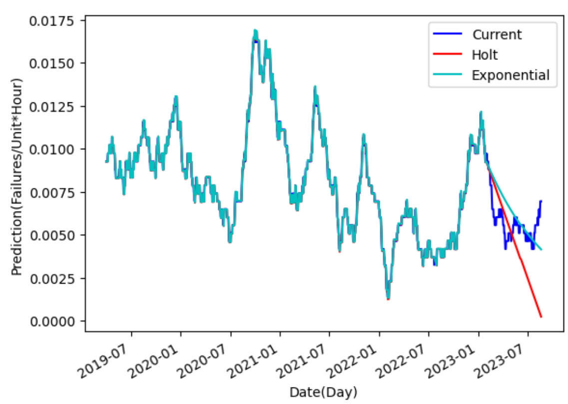


Source: own.

3.3.3. Exponential Smoothing Methods - Holt Smoothing

Other methods of exponential smoothing are the Holt method and the Exponential Holt method. For our study case, the simulation of the Holt method is initially performed and adjusted with a linear trend, resulting in the linear prediction function presented in figure 10 in red. The Exponential Holt method implemented is presented in Figure 10 in blue, which shows a decreasing behavior, that is, an improvement of these models compared to the ARIMA and SES models, because the lines are not flat and have a behavior similar to the time series.

Figure 9. Prediction models Holt.



Source: own.

3.3.4. Evaluation

To evaluate the results of the failure rate modeling, we used the symmetric mean absolute percentage error (SMAPE), the mean absolute percentage error (MAPE) and the root mean squared error (RMSE). The results are presented in Table 4.

Table 4. **Results prediction errors.**

MODEL	MAPE	SMAPE	RSME
SES	0,6673%	48,1863%	0,3701%
Holt softening	0,3443%	46,8803%	0,2407%
ARIMA	0,2680%	23,3021%	0,1561%
Smoothing of Exponential Holt	0,1879%	17,2608%	0,1323%

Source: own.

The Exponential Holt Smoothing model achieves the best prediction with lower error metrics than the other models. For example, the SES model proposes as a predictive function a constant linear relationship resulting in a SMAPE of 48%. The ARIMA model presents a flat linear trend in the prediction function (Figure 8), however, the prediction error values do not differ much from the Exponential Holt model, with 23% and 17% respectively.

3.4. Conclusions

The exploratory analysis and modeling of the failure rate of a local electrical distribution transformer shows a variability in the failure rate values over time. Although the underlying causality of the failures must be analyzed, the analyses generate indications of a growing and decreasing trend of the data during certain seasons in the analysis horizon. Therefore, reliability models will present an important variability by considering the updated failure rate, which reflects the current state of the complex system. Regarding the modeling performance, it is observed that the models Exponential Holt Smoothing and ARIMA present better performance and are viable for the prediction of short-term future values. However, the previous study opens the doors to future prediction analysis with more robust models to analyze the percentage of improvement compared to previous models.

Acknowledgments

This work was financed by the Universidad Tecnológica de Pereira, especially by the Master's program in Electrical Engineering and the research group: Gestión de Sistemas Eléctricos, Electrónicos y Automáticos (GISEEA).

References

[1] M. Mahdavi, M. S. Javadi, and J. P. S. Catalão, "Integrated generation-transmission expansion planning considering power system reliability and optimal maintenance activities," *Int. J. Electr. Power Energy Syst.*, vol. 145, p. 108688, 2023. DOI: 10.1016/j.ijepes.2022.108688. [Online]. Available: https://doi.org/10.1016/j.ijepes.2022.108688

[2] L. Ding, H. Wang, K. Kang, and K. Wang, "A novel method for SIL verification based on system degradation using reliability block diagram," *Reliab. Eng. Syst. Saf.*, vol. 132, pp. 36-45, 2014. DOI: 10.1016/j.ress.2014.07.005. [Online]. Available: https://doi.org/10.1016/j.ress.2014.07.005

[3] ISO, "ISO 55001:2014 Asset Management. Management systems – Requirements," The British Standards Institution, 2014.

[4] B. Dhilon, *Applied Reliability and Quality: Fundamentals, Methods, and Procedures*, New Jersey, USA: Springer, 2007.

[5] M. Firouzi, A. Samimi, and A. Salami, "Reliability evaluation of a composite power system in the presence of renewable generations," *Reliab. Eng. Syst. Saf.*, vol. 222, 2022. DOI: 10.1016/j.ress.2022.108396. [Online]. Available: https://doi.org/10.1016/j.ress.2022.108396.

[6] R. Yajun and M. Xiurui, "The reliability evaluation of the power system containing wind farm using the improved state space partition method," 2014 International Conference on Power System Technology, Chengdu, China, pp. 36-41, 2014. [Online]. Available: https://doi.org/10.1109/POWERCON.2014.6993498

[7] S. Anbazhagan and N. Kumarappan, "Day-ahead deregulated electricity market price forecasting using neural network input featured by DCT," *Energy Convers. Manage.*, vol. 78, pp. 711-719, 2014. [Online]. Available: https://doi.org/10.1016/j.enconman.2013.11.031.

[8] X. Fan, X. Zhang, and X. B. Yu, "Uncertainty quantification of a deep learning model for failure rate prediction of water distribution networks," *Reliab. Eng. Syst. Saf.*, vol. 236, p. 109088, 2023. DOI: 10.1016/j.ress.2023.109088. [Online]. Available: https://doi.org/10.1016/j.ress.2023.109088.

[9] W. Qiu, Q. Tang, Z. Teng, W. Yao, and J. Qiu, "Failure rate prediction of electrical meters based on weighted hierarchical Bayesian," *Measurement*, vol. 142, pp. 21-29, 2019. DOI: 10.1016/j.measurement.2019.04.062. [Online]. Available: https://doi.org/10.1016/j.measurement.2019.04.062.

[10] J. Stock, *Introduction to Econometrics, Madrid*, Spain: Pearson Educación S.A., pp. 373-411, 2012.

[11] G. *Box,Time Series Analysis Forecasting and Control*, New Jersey, USA: John Wiley & Sons Ltd, pp. 2-43,2016.

[12] S. *Raschka, Machine Learning con PyTorch y Scikit-Learn*, Madrid, España: Alphaeditorial, pp. 290-307, 2023.

[13] Y. Chen, M. Ma, and Y. Zou, "Forecasting hourly electricity demand with nonparametric functional data analysis," *Procedia Comput. Sci.*, vol. 214, pp. 428-436, 2022. DOI: 10.1016/j.procs.2022.11.195. [Online]. Available: https://doi.org/10.1016/j.procs.2022.11.195

[14] Ye Zhu, Shiwen Xie, Yongfang Xie, Xiaofang Chen, "Temperature prediction of aluminum reduction cell based on integration of dual attention LSTM for non-stationary sub-sequence and ARMA for stationary sub-sequences," *Control Eng. Pract.*, vol. 138, p. 105567, 2023. DOI: 10.1016/j.conengprac.2023.105567. [Online]. Available: https://doi.org/10.1016/j.conengprac.2023.105567.

[15] M. Rahimi y M. Rausand, "Failure rate prediction in various life cycle phases: A framework for updating," en 2012 IEEE International Conference on Industrial Engineering and Engineering Management, Hong Kong, China, pp. 762-766, 2012. [Online]. Available: https://doi.org/10.1109/IEEM.2012.6837842

[16] M. Tabesh, J. Soltani, R. Farmani, y D. Savic, "Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling," *Journal of Hydroinformatics*, vol.11, no.1, pp. 1-17, 2009. [Online]. Available: https://doi.org/10.2166/hydro.2009.008