- ORIGINAL ARTICLE -

Impact of Ambient Temperature on Short-term Maximum Electrical Demand through the Performance of Forecast Models generated with Prophet

Impacto de la Temperatura Ambiente en la Demanda Eléctrica Máxima de Corto Plazo a través del Desempeño de Modelos de Pronóstico Generados con Prophet

César A. Yajure-Ramírez[1]

[1] *Basic School of Engineering, Central University of Venezuela, Venezuela*
cyajure@gmail.com

Abstract

The maximum short-term electrical demand is affected by climatic factors, including ambient temperature. To incorporate it into the forecast models, it is necessary to generate an indicator that represents the ambient temperature of the area under study. The objective of this research is to determine the impact of ambient temperature on the short-term maximum electrical demand through the performance of the forecast models, integrating into a single indicator the temperature measurements from different points of the geographical area under analysis, using as weighting factors to the proportions of regional demands with respect to total demand. The Prophet forecasting technique is used, with historical data on electrical demand and daily ambient temperature from November 2022 to November 2024. To evaluate the models, the MAE, RMSE, and MAPE metrics are used, with data outside the historical period. The forecast model considering the Weighted High Temperature indicator as a regressor variable was the one that had the greatest improvements in the metrics when comparing them with those coming from the model that did not consider temperature as a regressor variable, with improvements of 25%, 21%, and 15%, in MAPE, MAE, and RMSE, respectively.

Keywords: Correlation, electrical demand, forecast, performance metrics, temperature.

Resumen

La demanda eléctrica máxima de corto plazo se ve afectada por factores climáticos, entre ellos la temperatura ambiente. Para incorporarla en los modelos de pronóstico, se hace necesario generar un indicador que represente a la temperatura ambiente del área bajo estudio. El objetivo de esta investigación es determinar el impacto de la temperatura ambiente en la demanda eléctrica máxima de corto plazo a través del desempeño de los modelos de pronóstico, integrando en un solo indicador las mediciones de temperatura de distintos puntos del área geográfica bajo análisis, utilizando como factores de ponderación a las proporciones de las demandas regionales con respecto a la demanda total. Se hace uso de la técnica de pronóstico Prophet, con datos históricos de demanda eléctrica y temperatura ambiente diaria desde noviembre del 2022 hasta noviembre del 2024. Para evaluar los modelos se utilizan las métricas MAE, RMSE, y MAPE, con datos fuera del período histórico. El modelo de pronóstico considerando el indicador Temperatura Alta Ponderada como variable regresora fue el que tuvo las mayores mejoras de las métricas al compararlas con aquellas provenientes del modelo que no consideró a la temperatura como variable regresora, con mejoras del 25%, 21%, y 15%, en MAPE, MAE, y RMSE, respectivamente.

Palabras claves: Correlación, demanda eléctrica, métricas de desempeño, pronóstico, temperatura.

1. Introduction

Quantitative forecasting techniques are useful when historical data is available, and the behavior of the variables is relatively stable. In general terms, these techniques could have different forecast horizons depending on the use that the results will have, that is, short term, medium term, and long term. With regard to the variables of the electricity sector, some authors consider four classes: very short-term forecast that would range from a few minutes to an hour and is used for real-time control, short-term forecast with ranges from one hour to one week for the schedule of electrical generation and transmission, the medium-term forecast that ranges from one week to one year that is used to schedule maintenance, the purchase of fuel for the plants, among other uses, and the long-term forecast which ranges from one year to twenty years, which is useful for strategic planning and development of the sector [1].

In electricity markets, it is important to know the estimate of hourly and daily demand as accurately as possible, since under-contracting or over-contracting energy and then selling it in the markets can cause financial problems for electrical companies. On the other hand, if you are in a regulated system, demand forecast information is useful for short-term operational planning, in order to satisfy this demand. For the development of the short-term forecast of maximum electrical demand, there is a variety of techniques, both statistical and non-statistical, with time series analysis being widely used through ARIMA models, along with artificial neural networks [2]. It is well known that the electricity demand of a country is dependent on climatic variables, economic variables, and demographic variables. For the short term and hourly and daily resolutions, climatic and demographic variables have greater influence, two of the most important are the ambient temperature and the amount of rainfall. For example, according to studies carried out in Spain, during a weekday in spring, an increase of one degree Celsius in temperature increases energy consumption by approximately 25 GWh [3].

Now, due to the effect that temperature has on the demand of tropical countries in the short term, this work focuses on determining the impact of temperature on electricity demand. Validating the impact of temperature on demand will allow us to be proactive in the months of high temperatures in the country, when there is high demand, and plan savings in the reservoirs of hydroelectric plants, or fuel in thermoelectric plants, during the months of lower temperatures. The complexity of using the temperature variable in demand forecasting models lies in the fact that its values are specific, so we do want to estimate the electrical demand of a geographic region considering the temperature environment, a mechanism must be established to generate temperature values that represent the entire region. In this sense, the objective of this research is to determine the impact of ambient temperature on the maximum short-term electrical demand through the performance of the forecast models and integrating temperature measurements at different points of the area into a single indicator. The Prophet forecasting technique developed by Facebook is used.

In previous works, the topic of studying the correlation between ambient temperature and electricity consumption has been addressed with different approaches than that presented in this research, being innovative the obtaining of integrated temperature indicators by using the proportions of demand as factors of weighing. For example, in [4] authors develop an empirical analysis based on residential electricity consumption surveys to explore the correlation between temperature increase and household electricity consumption in China. These surveys provide valuable information about residential consumption and the types of home appliances that predominate in chinese households. Among the results obtained, they indicate the existence of a significant correlation between the variables, with an increase of 8.9% in annual electrical energy consumption when the average temperature exceeded 32°C.

In addition, in [5] was developed a study to determine the effect of temperature on electricity consumption in Mexico using fixed effects panel models. The methodology used includes the control of other variables that affect electrical energy consumption, such as: income, manufacturing production per capita, electricity and gas prices, and household size. They find that there is a non-linear relationship between temperature and electrical energy consumption, and that this consumption increases with temperature, being stronger in the hottest states. By extrapolating the results to high environmental warming scenarios, they obtain that by the end of the century electrical energy consumption would increase by 12%.

Besides, in [6] was carried out a study to determine the impact of temperature on the electricity demand of the city of Dhaka in Bangladesh. They use a multiple linear regression model, considering ambient temperature, relative humidity, wind speed, wind direction, and precipitation, as explanatory variables. They make use of the Pearson correlation coefficient and covariance to establish the relationship between the variables. They tested up to four types of functions that relate demand and ambient temperature: linear, second-degree polynomial, logarithmic, and exponential. They obtain that temperature represents 75% of the total variance, and that the exponential model is the one that best represents the relationship between the two variables, in terms of the coefficient of determination.

Moreover, authors in [7] carried out a work to establish the effects of ambient temperature on residential electricity consumption in the Gauteng province of South Africa. The historical data used corresponds to the hourly electricity consumption of 5,975 homes between January 1, 2010, and March 31, 2013. These homes have high incomes with high consumption of electrical energy for heating in the cold months and by air conditioning in the months of high temperatures. They develop regression models that include the time of day and month of the year among the explanatory variables. The results suggest a strong and robust negative log-linear relationship between temperature and electricity consumption below a heating threshold and a strong positive relationship above a cooling threshold.

Additionally, in [8] was proposed a prediction model for the total monthly electricity consumption of the Fujian province in China, developing a composite temperature index of both daily resolution and monthly resolution. The historical data corresponds to the daily values of total electricity consumption from January 2017 to November 2020, in addition to the daily values of maximum temperature of nine cities that make up the province of Fujian. The average of the maximum values is taken to obtain the daily temperature of the province, and the average of the maximum monthly values is taken to obtain its monthly temperature. The proposed model is applied and the results obtained are compared with those obtained with ARIMA models, support vector machines, and random forests. They found that the models that consider the temperature index have better performance than those that do not consider it, with the proposed model together with ARIMA models having the best performance for the daily forecast with an accuracy of 98.23%.

Authors in [9] developed a study to determine the impact of temperature on the short-term electricity demand forecast through time series analysis using traditional methods and artificial neural networks. The data corresponds to the hourly values of electrical demand read at a substation located in Kathmandu, Nepal, from April 2018 to April 2019, that is, 8,760 readings. The methodology includes different models for weekdays and weekends. The results indicate that the artificial neural network models have the best performance with a MAPE of 0.34% for weekdays and 8.04% for weekends.

In [10] was carried out a correlation study between climatic variables and electricity demand in two regions of Indonesia, the province of Central Java and the island of Bali. The climatic variables considered were temperature, precipitation, net solar radiation, pressure, wind speed, and relative humidity. The historical data used are the daily demand values for the years 2018 and 2019, as well as the daily values of the climatic variables for the same period. They obtain a positive correlation coefficient of 0.55 in the province of Central Java, and 0.71 on the island of Bali.

Also, authors in [11] developed a work to determine how climatic factors affect electricity consumption in the metropolitan region of Bangkok in Thailand, considering different temporal resolutions, and through the correlation coefficient. Separate historical data are used every 30 minutes between the years 2015 and 2017 of electricity consumption and seven climatic factors: temperature, dew point, relative humidity, wind (quantity), pressure, rain and cloud cover. The temporal resolutions considered were separated by 30 minutes, daily, weekly, and monthly. They obtain that temperature has greater weight for resolutions of up to one day, with the dew point being the most important factor for weekly and monthly resolutions.

Finally, authors in [12] presented a methodology to select and process temperature variables automatically, to improve the electrical demand forecasting process, and facilitate the interpretation of the results. It applies to data from the Spanish Electrical Network. They managed to reduce the forecast error consistently for the forecast horizon, with an improvement of 0.16% in MAPE and 59.71 MWh in RMSE. The procedure improved the interpretability of the results, since they separate the effect of temperature according to location and time.

The rest of the work is distributed as follows. Section II presents the theoretical review, as well as the methodology used in the research. Section III shows and discusses the results obtained. Section IV presents the conclusions. Finally, the bibliographical references are listed.

2. Background

2.1. Model performance metrics

To evaluate forecast models and make comparisons between them, a series of metrics are used. Among the standard statistical measures to evaluate models are: the mean absolute error (MAE) whose expression is given by Eq.1, the square root of the mean square error (RMSE) represented by Eq. 2, and the mean absolute percentage error (MAPE) represented by Eq. 3. The error being the difference between the real value and the predicted value [13].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - F_i| \qquad \text{Eq. (1)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - F_i)^2} \qquad \text{Eq. (2)}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\left(\frac{Y_i - F_i}{Y_i}\right) \times 100\right| \qquad \text{Eq. (3)}$$

In the previous equations, Y_i is the ith actual value of the variable to be predicted, F_i is the ith predicted value, and n is the number of records or historical data available.

Additionally, in order to measure the quality of fit of the models, the determination coefficient R^2 is used. This indicator is a measure of the quality of fit of the model, and indicates the proportion of the variability that is explained by the model obtained. It varies between 0 and 1, a value very close to 1 reveals that the model fit is almost perfect. While a value close to 0 is an example of poor model fit [14]. In any case, it should not be used as a sole performance measure, but rather in conjunction with the other previously defined metrics. The determination coefficient is calculated according to Eq. (4).

$$R^2 = \frac{SCR}{SCR+SCE} = 1 - \frac{SCE}{SCR+SCE} \quad \text{Eq. (4)}$$

In Eq (4), SCR is the sum of the regression squares. It is the quantity that represents the variation explained by the model. Likewise, SCE is the sum of the squares of the errors. It is the variation due to error, or the variation not explained by the model.

2.2. Prophet forecasting technique

Use is made of the Prophet technique created by the Facebook data science team during 2017. It was designed specifically for the analysis of time series, based on an additive model where non-linear trends are quickly adjusted to different types of data seasonality, such as: annual, weekly, and daily, in addition to incorporating the effect of vacations and holidays. According to the team, this method works best with time series that have strong seasonal effects and multiple periods of historical data. It is robust to missing data and changes in trend and usually handles outliers well [15].

The technique is based on a time series model made up of an element that represents the trend, another that represents regular changes, that is, seasonality, another component that represents the effect of non-working days, and finally the error term [16]. Therefore, the model is given by Eq. (5):

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad \text{Eq. (5)}$$

In Eq. (5) it is true that $y(t)$ is the time series, $g(t)$ is the trend component, $s(t)$ represents the periodic changes in the series, $h(t)$ represents the effect of holidays and vacation, and ϵ_t is the error term that represents changes that do not fit within the other components.

Regarding the trend component, the technique implements two types: growth saturation model, and a piecewise linear model. Related to seasonality, it is modeled with periodic time functions, specifically with standard Fourier series, such as the one presented in Eq. (6).

$$s(t) = \sum_{n=1}^{N} \left(a_n \cdot cos\left(\frac{2\pi nt}{P}\right) + b_n \cdot sen\left(\frac{2\pi nt}{P}\right) \right) \quad \text{Eq. (6)}$$

Where P is the period of the series, which will depend on the type of seasonality, and N is the number of terms in the series.

Taylor & Letham indicate that P is the expected regular period for the time series, taking a value equal to 7 for data with weekly seasonality and 365.25 for data with annual seasonality. From the above it is concluded that the technique works by default with data with daily resolution. Likewise, they propose that higher values of N apply to seasonal patterns that change more rapidly, and in their studies, they have found that a value of 10 is useful for annual seasonality, and 3 for monthly seasonality. According to [17], the technique defines by default an additive seasonality, but there is also the option of establishing a multiplicative seasonality.

One of the characteristics that contrasts this technique with some other time series analysis techniques is the management of non-working days, allowing the inclusion in the model of a list of holidays and vacations that are within the historical data period, but also within the forecast horizon.

It has been used in various applications, for example, in [18] they use it for predicting stock values in the Indian stock market, in [19] they apply this technique for the forecast of successful sales, and in [20] use the technique to develop temperature forecasting, among other uses.

3. Methodology

The historical values of maximum electrical demand used in the models correspond to the daily values of the period from November 2022 to early November 2024 of a South American country. Regarding temperature, measurements of both average ambient temperature and high ambient temperature were taken in each of the different regions that make up the country's electrical system. These measurements are found in [21]. A review of these data was made to ensure the absence of missing, duplicate, or outlier data, in accordance with the data preprocessing techniques presented in [22]. To detect atypical data, the Tukey test was used, since as mentioned by [23], being based on the quartiles of the data, it does not require assuming a specific distribution. In any case, no missing, duplicate, or atypical data were detected. Similarly, temperature data were processed to convert them from a scale in degrees Fahrenheit to degrees Celsius.

To obtain the temperature indicator, it is considered that each of the regions that make up the complete system represents an electrical demand whose proportion with respect to the total demand is approximately constant over time. With the historical values of the demand, the expected proportion was obtained, given by Eq. (7), for each of the 365 days of the year. These proportions represent the weights or weighting factors of each of the regions, which are used to obtain an average ambient temperature weighted by the proportion of the demand of the respective region with respect to the total demand, which is calculated with Eq. (8).

With this same procedure, the weighted high temperature was obtained, using Eq. (9). Similarly, the average of the high temperatures, the average of the average temperatures, and the median of the average temperatures, were calculated.

$$FP_{ij} = \frac{Demand\ of\ the\ region\ j\ on\ the\ day\ i}{Total\ demand\ on\ the\ day\ i} \qquad \text{Eq. (7)}$$

$$TMP_i = \sum_{j=1}^{8}\left(TM_{ij} \times FP_{ij}\right) \qquad \text{Eq. (8)}$$

$$TAP_i = \sum_{j=1}^{8}\left(TA_{ij} \times FP_{ij}\right) \qquad \text{Eq. (9)}$$

In Eq. (8) it holds that TMP_i is the weighted average temperature of day i, and TM_{ij} is the average temperature of region *j* on day *i*. Similarly, in Eq. (9) it holds that TAP_i is the weighted high temperature of day *i*, and TA_{ij} is the high temperature of region *j* on day *i*.

The next step consisted of doing a correlation analysis of all the variables among themselves, that is, all the calculated temperature indicators and the maximum electrical demand. The results of this analysis were used to define which temperature indicators to use as regressors in electricity demand forecasting models.

Prior to modeling, an exploratory analysis of the data was carried out, temporal demand graphs were developed, which allowed observing the trend and seasonality that occur during the study period considered.

In the modeling stage, the Prophet forecasting technique developed by Facebook was used through the Python programming language, obtaining forecast values of the maximum electrical demand taking into account the ambient temperature, but also without taking it into account, for purposes of comparison. To evaluate the models, the MAE, RMSE, and MAPE metrics were used, both in the test data and with additional data corresponding to the months of November and December, determining the improvement produced in these indicators when the temperature is taken into account in the forecasting models.

4. Discussion of results

4.1. Correlation analysis

For the purposes of the analysis, the following indicators are taken into account: average high temperature (TA), weighted high temperature (TAP), average temperature (TM), weighted average temperature (TMP), and median temperature (TMED), and their relationship with the maximum electrical demand (DEM). It must be taken into account that the coefficient of variation range between -1 and 1. If the value is equal to or close to -1, it indicates that the corresponding pair of variables have a strong and inverse relationship. On the contrary, if the value is equal to or close to 1, it indicates that the pair of variables under analysis have a strong and direct relationship. If the value is close to 0, it indicates that the relationship between the respective variables is null. As for intermediate values, absolute values between 0 to 0.3 indicate that the relationship is “weak”, between 0.3 to 0.7 the relationship is “moderate”, and between 0.7 to 1.0 the relationship is “strong” [24].

The analysis was carried out independently using the Pearson, Spearman and Kendall methods, given that the first of them is parametric and the other two methods do not assume a specific distribution of the data. Since the results of the three methods were similar to each other, those obtained with Spearman's method are presented in Fig. 1. In [25] the Spearman method is considered for the optimal selection of explanatory variables for forecasting models.

From Fig. 1, it can be seen that demand had a “weak” relationship with high ambient temperature, and a “moderate” relationship with the rest of the temperature indicators considered, with the coefficient being greater with the weighted average temperature 0.62, followed by the median temperature with 0.60.

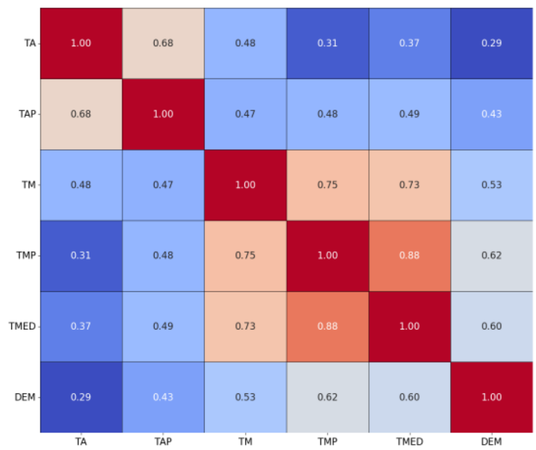

Fig. 1 Correlation Matrix.

Subsequently, Fig. 2 presents the scatter graphs of the maximum electrical demand with the temperature indicators TAP, TM, TMP, and TMED. The graphs show the type of relationship for which the coefficient of determination R^2 was higher, resulting in that for the four cases the relationship

Demand vs. Temperature is of the second-degree polynomial type. According R^2, it can be seen that the demand versus temperature relationship is stronger when considering the median temperature and the weighted average temperature.

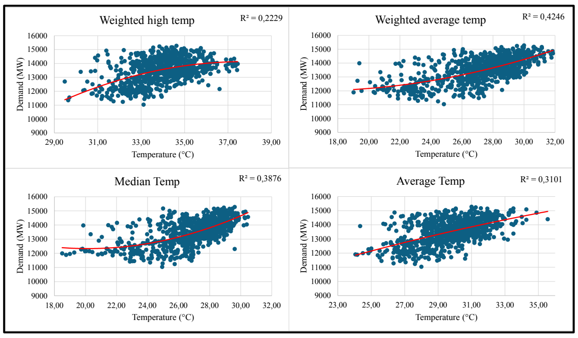

Fig. 2 Scatter plots

4.2. Exploratory data analysis

Firstly, the analysis of the demand data is carried out with hourly resolution, presenting the average curve in Fig 3. It is observed that demand presents two peaks, one around 2 pm, and another around 8 pm (daily seasonality). Likewise, it can also be seen that demand has an upward trend, mainly from 2023 onwards. The shaded part of the curves corresponds to the variability of the data, represented by a 95% confidence interval.

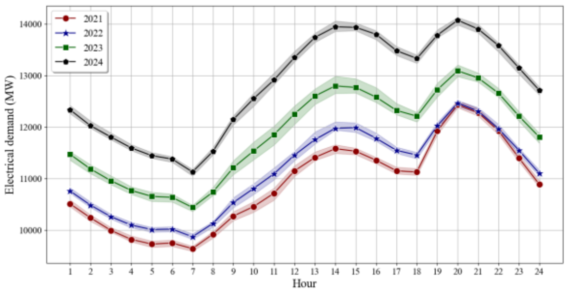

Fig. 3 Average hourly demand curve

Next, Fig 4 shows the average daily demand curve, in which “1” corresponds to Monday and “7” corresponds to Sunday. It can be seen that, for all the years considered, demand decreases during the weekend (weekly seasonality), with minimum values for Sundays. The upward trend of the data can be seen.

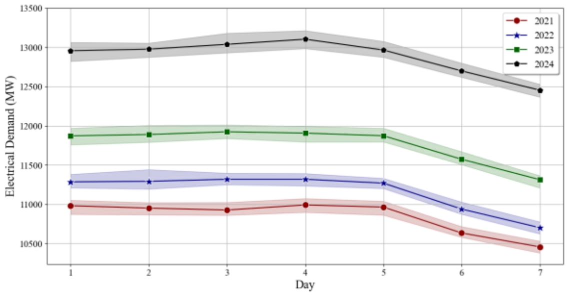

Fig. 4 Average daily demand curve

For the purposes of observing monthly seasonality, Fig 5 shows the monthly average curve, in which “1” corresponds to the month of January and “12” corresponds to the month of December. It is noted that demand has two peak values, one around the month of May and another around the month of October. Like the previous figures, an upward trend is seen, which intensifies starting in May 2023.

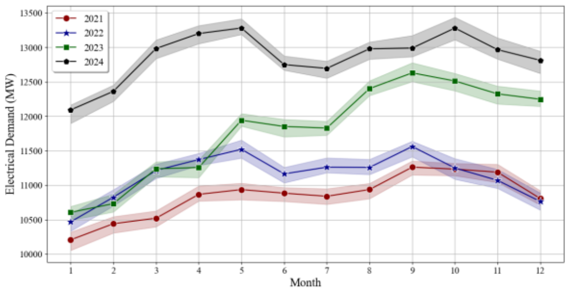

Fig. 5 Average monthly demand curve

4.3. Forecast of maximum electrical demand

The Prophet technique was used to develop the forecast of maximum electrical demand, first without considering any of the temperature indicators, and then separately considering the following temperature indicators as regressor variables: TMP, TAP, and TMED. Regarding the hyperparameters, the seasonality mode was set as multiplicative type. The robustness of the models was estimated using the cross-validation method that is incorporated in this forecasting technique. 366 days were set for the training period, 30 days for the cut-off points, and up to 7 days for the forecast horizon. For each of the forecast horizons, MAPE values of less than 3% were obtained, and coverage of 85% for the 7-day horizon. This same technique was used in [26] for the forecast of the levels of the reservoirs used for the production of hydroelectric energy.

The historical data considered corresponded to the period from 11/01/2022 to 11/03/2024, and four models were generated each week from 11/04/2024 to 12/22/2024, that is, for seven weeks (49 days). Each of the models obtained had a MAPE value of less than 5% in the testing phase. For the regressor variables, the actual values of the ambient temperature were taken. Fig. 6 shows the daily curves of the average value of the 95% prediction interval obtained for each of the temperature indicators considered, along with the real values of demand, and the daily percentage errors.

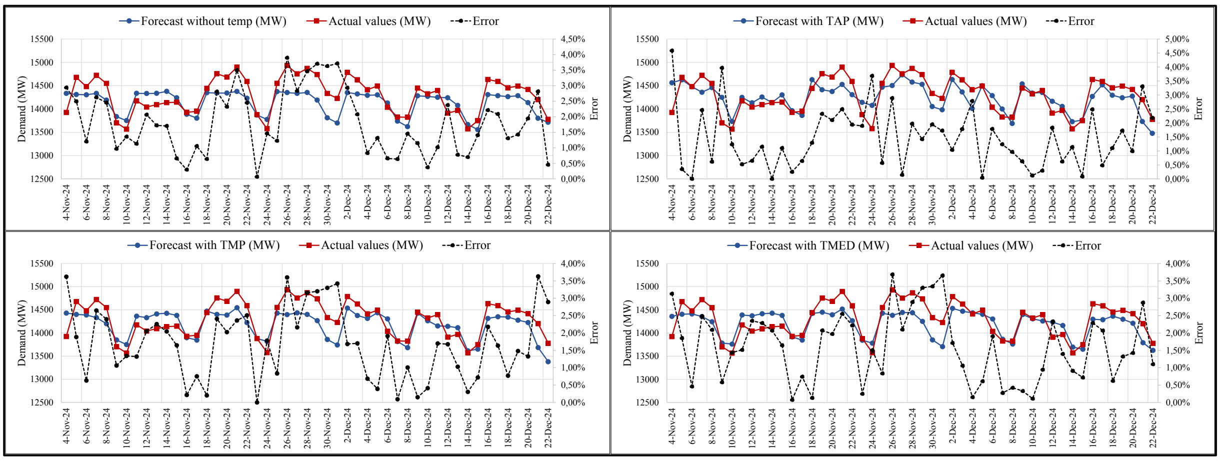

Fig. 6 Forecast vs. actual demand.

From Fig. 6, it can be seen that all cases the percentage errors are less than 5%. In the case of the forecast without considering the ambient temperature, the maximum percentage error was 3.90% corresponding to 11/26, with a minimum error of 0.07% on 11/23. Regarding the results when the TAP is considered, the maximum error was 4.59% on November 4, and the minimum value of the percentage error was 0.00% on November 6. When reviewing the results when the TMP is used, it is noted that the maximum error corresponded to the forecast of 12/21 with 3.63%, and the minimum error was that of 11/23 with 0.00%. Finally, it is observed that when the TMED was used, the maximum error was 3.69% on 11/26 and the minimum value of the error occurred on 11/16 with 0.08%.

Related to TAP, the maximum percentage error of 4.59% was obtained, but also the lowest number of days with a percentage error (e) below 3% with a total of 45 days. The results associated with this regressor had the maximum number of days with errors less than 1% with 17 days, followed by the results associated with the median temperature with 16 days. Likewise, the results associated with the TAP regressor had the highest number of days in which the daily forecast of the maximum electrical demand was above the actual value of the demand, with a total of 21 days, followed by the TMED regressor with 15 days. Complete results are presented in Table 1.

4.4. Model performance evaluation

The MAE, RMSE, and MAPE metrics are calculated by comparing the actual demand values with the corresponding forecast values, for the period from 11/04/2024 to 12/22/2024. These indicators are also used in [27] to evaluate residential electricity consumption forecasting models. The results are presented in Table 2.

Table 1. Model comparison.

Models	e < 3%	e < 1%	Pron > Real
No temperature	43	12	13
With TMP	42	14	14
With TAP	45	19	21
With TMED	44	17	15

Table 2. Model performance metrics.

Models	MAPE (%)	MAE (MW)	RMSE (MW)
No temperature	1.80	260	301
With TMP	1.68	241	283
With TAP	1.44	206	256
With TMED	1.59	229	271

From Table 2 it can be seen that the models in which the TAP regressor was considered had the lowest values of each of the three metrics considered, followed by the models that considered the TMED regressor. It can also be seen that the models in which the use of ambient temperature as a regressor variable is not considered had the highest values of each of the metrics, that is, they were the worst performing.

4.5. Determination of the impact of temperature

If the MAE metric of the models that do not use ambient temperature is compared with the models that use it, it is obtained that: it is reduced by almost 21% when using the TAP regressor, 12% when considering the TMED regressor, and 7% when taking into account the TMP regressor. Regarding the RMSE metric, the improvements are 15%, 10%, and 6%, when considering the TAP, TMED, and TMP regressors, respectively. In relation to the MAPE metric, when using the TAP, TMED, and TMP regressors, improvements of 25%, 13%, and 7% are obtained, respectively.

For comparison purposes, simple linear regression models were run with electrical demand as the objective variable, and temperature as the regressor variable. The complete data was split 70% for training the model, and 30% for testing the model. For each of the temperature indicators TAP, TMP, and TMED, positive and statistically significant coefficients were obtained, with values of 322, 234, and 240, respectively. This confirms the direct correlation between electricity demand and this climate variable.

5. Conclusions

A methodology is presented to determine the impact of ambient temperature on the maximum electrical demand of a country through the performance of forecast models. The temperature values of the different regions that make up the complete system are integrated into global indicators considering the historical proportion of the maximum demand of the regions with respect to the total demand of the country.

From the correlation analysis it is observed that all the temperature indicators considered have a direct relationship with the maximum electrical demand, this relationship being moderate with respect to the high weighted temperature, the weighted average temperature, the average temperature, and the medium temperature. The relationship of these indicators with demand turned out to be quadratic, this relationship being stronger with the weighted average temperature.

The forecast models developed using the weighted high temperature were those that had the best values of the evaluation metrics considered, with the models in which temperature was not considered those that had the highest values of the aforementioned metrics. When the TAP regressor is used, the MAPE is reduced by 25%, the MAE by 21%, and the RMSE by 15%, when compared to the values of these metrics from the models when temperature is not used as a regressor variable.

Authors' Contribution

The author confirms contribution to the paper as follows:

CAYR: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing – original draft, Writing - Reviewing and Editing.

Competing Interests

The author has declared that no competing interests exist.

References

[1] M. A. Hammad, B. Jereb, B. Rosi y D. Dragan, «Methods and Models for Electric Load Forecasting: A Comprehensive Review,» Logistics, Supply Chain, Sustainability and Global Challenges, vol. 11, nº 1, pp. 51-76. doi: 10.2478/jlst-2020-0004., 2020.

[2] G. P. Papaioannou, C. Dikaiakos, A. Dramountanis y P. G. Papaioannou, «Analysis and Modeling for Short- to Medium-Term Load Forecasting Using a Hybrid Manifold Learning Principal Component Model and Comparison with Classical Statistical Models (SARIMAX, Exponential Smoothing) and Artificial Intelligence Models (ANN, SVM): Th,» MDPI energies, vol. 9, nº 8, 2016.

[3] Aleasoft Energy Forecasting, «Las altas temperaturas y la demanda eléctrica irán a más en los veranos futuros,» 27 Julio 2023. [En línea]. Available: https://elperiodicodelaenergia.com/las-altas-temperaturas-y-la-demanda-electrica-iran-a-mas-en-los-veranos-futuros/.

[4] Y. Su y K. Ullah, «Exploring the correlation between rising temperature and household electricity consumption: An empirical analysis in China,» ScienceDirect - Heliyon, pp. 1-14. https://doi.org/10.1016/j.heliyon.2024.e30130, 2024.

[5] W. Wouter Botzen, T. Nees y F. Estrada, «Temperature Effects on Electricity and Gas Consumption: Empirical Evidence from Mexico and Projections under Future Climate Conditions,» MDPI - Sustainability, pp. 1-28. https://doi.org/10.3390/su13010305, 2021.

[6] A. Istiaque y S. Islam Khan, «Impact of Ambient Temperature on Electricity Demand of Dhaka City of Bangladesh,» Energy and Power Engineering, pp. 319-331. DOI: 10.4236/epe.2018.107020, 2018.

[7] S. B. Berkouwer, «Electric Heating and the Effects of Temperature on Household Electricity Consumption in South Africa,» The Energy Journal, pp. 209-230, 2020. https://doi.org/10.5547/01956574.41.4.sber.

[8] X. Li, W. Shang, X. Zhang, B. Shan y X. Wang, «Total Electricity Consumption Forecasting Based on Temperature Composite Index and Mixed Frequency Models,» Data Intelligence, pp. 750-766, 2023. https://doi.org/10.1162/dint_a_00215.

[9] Y. Rajbhandari, A. Marahatta, B. Ghimire, A. Shrestha, A. Gachhadar, A. Thapa, K. Chapagain y P. Korba, «Impact Study of Temperature on the Time Series Electricity Demand of Urban Nepal for Short-Term Load Forecasting,» MDPI - Applied System Innovation, pp. 1-20, 2021.

[10] S. Aisyah y A. A. Simaremare, «Correlation between Weather Variables and Electricity Demand,» de 6th International Energy Conference, 2021.

[11] Y. Ratanawaraha, T. Kaewkitipong, S. Nardviriyakul y J. Tang, «How Weather Factors Affect Electricity Consumption in the Different Times Ranges: A Study in the Bangkok Metropolitan Region of Thailand,» International Journal of Environmental Science and Development, pp. 252-256, 2019. doi: 10.18178/ijesd.2019.10.8.1182.

[12] A. Candela Esclapez, M. López García, S. Valero Verdú y C. Senabre Blanes, «Automatic Selection of Temperature Variables for Short-Term Load Forecasting,» MDPI Sustainability, pp. 1-22, 2022. https://doi.org/10.3390/su142013339.

[13] S. Makridakis, S. Wheelwright y R. Hyndman, Manual of Forecasting: Methods and Applications, 1997.

[14] R. E. Walpole, R. H. Myers, S. L. Myers y K. Ye, Probabilidad y estadística para ingeniería y ciencias, Ciudad de México: Pearson Educación de México, S.A. de C.V., 2012.

[15] Facebook, «https://facebook.github.io/prophet/docs/quick_start.html#python-api,» 4 Mayo 2024. [En línea]. Available: https://facebook.github.io/prophet/.

[16] S. J. Taylor y B. Letham, «Forecasting at Scale,» Peer J, p. https://doi.org/10.7287/peerj.preprints.3190v2, 2017.

[17] M. Peixeiro, Time Series Forecasting in Python, Shelter Island, NY: Manning Publications Co., 2022.

[18] K. Sharma, R. Bhalla y G. Ganesan, «Time Series Forecasting Using FB-Prophet,» de ACM-2022: Algorithms Computing and Mathematics Conference, Chennai, India, 2022.

[19] E. Žunić, K. Korjenić, K. Hodžić y D. Đonko, «Application of Facebook's Prophet Algorithm for Successful Sales Forecasting Based on Real-World Data,» International Journal of Computer Science & Information Technology, pp. 23-36, 2020. DOI: 10.5121/ijcsit.2020.12203.

[20] Z. Z. Oo y S. Phyu, «Time Series Prediction Based on Facebook Prophet: A Case Study, Temperature Forecasting in Myintkyina,» International Journal of Applied Mathematics, Electronics and Computers, pp. 263-267, 2020. https://doi.org/10.18100/ijamec.816894.

[21] Meteovzla, «Enlaces meteorológicos de Venezuela,» 11 12 2024. [En línea]. Available: http://www.meteovzla.rf.gd/.

[22] A. Navlani, A. Fandango y I. Idris, Python Data Analysis, Birmingham, UK: Packt Publishing Ltd., 2021.

[23] A. KOLBAŞI y A. ÜNSAL, «A Comparison of the Outlier Detecting Methods: An Application on Turkish Foreign Trade Data,» Journal of Mathematics and Statistical Science, vol. 5, nº 9, pp. 213-234, 2019. http://www.ss-pub.org/journals/jmss/vol-5/vol-5-issue-9-september-2019/.

[24] B. Ratner, Statistical and Machine-Learning Data Mining - Techniques for Better Predictive Modeling and Analysis of Big Data, Boca Raton, FL: CRC Press Taylor & Francis Group, 2017.

[25] C. A. Yajure-Ramírez, «Enfoque multicriterio para la selección óptima de variables explicativas para modelos de pronóstico de la energía eléctrica de plantas solares fotovoltaicas,» Revista Tecnológica - Espol, pp. 83-98, 2023. https://doi.org/10.37815/rte.v35n3.1045.

[26] C. A. Yajure Ramírez, «METHODOLOGY FOR TARGET FORECASTING OF WATER LEVEL IN HYDROELECTRIC PLANT RESERVOIRS UNDER CONDITIONS OF LOW INFLOW,» Journal of Engineering and Technology for Industrial Applications, pp. 43-48, 2024. DOI: https://doi.org/10.5935/jetia.v10i48.1137.

[27] C. A. Yajure-Ramírez, «Multi-criteria methodology based on data science for the selection of the optimal forecast model for residential electricity consumption,» Scientia et Technica, pp. 108-116, 2023. DOI: https://doi.org/10.22517/23447214.25335.

Citation: C. A. Yajure-Ramírez. *Impact of ambient temperature on short-term maximum electrical demand through the performance of forecast models generated with Prophet.* Journal of Computer Science & Technology, vol. 25, no. 1, pp. 16-24, 2025.
DOI: 10.24215/16666038.25.e02
Received: January 14, 2025 **Accepted:** April 21, 2025.