Modelado y predicción de la demanda eléctrica: comparación de enfoques estadísticos y de aprendizaje automático
Mariela N. Uhrig, Leandro D. Vignolo, and Omar V. Müller
SADIO Electronic Journal of Informatics and Operations Research (EJS) Vol. 24 No. 1 (2025) e-ISSN 1514-6774
https://doi.org/10.24215/15146774e067 | https://revistas.unlp.edu.ar/ejs
Sociedad Argentina de Informática e Investigación Operativa | Universidad Nacional de La Plata | Buenos Aires | Argentina

Modelado y predicción de la demanda eléctrica: comparación de enfoques estadísticos y de aprendizaje automático

Mariela N. Uhrig[1], Leandro D. Vignolo[2], and Omar V. Müller[3]

[1] Centro de Investigación Científica y de Transferencia Tecnológica a la Producción, CICYTTP-CONICET-PROV. E. R.-UADER
[2] Research Institute for Signals, Systems and Computational Intelligence, sınc(i), FICH-UNL/CONICET
[3] Centro de Estudios de Variabilidad y Cambio Climático, FICH-UNL/CONICET
mnuhrig@conicet.gov.ar, ldvignolo@sinc.unl.edu.ar, ovmuller@unl.edu.ar

Resumen Una predicción confiable de la demanda eléctrica es esencial para mejorar la gestión del sistema de distribución, optimizando la utilización de recursos, racionalizando la planificación operativa y reduciendo las interrupciones del servicio. La fluctuación de la demanda eléctrica está influenciada por diversos factores externos, como las condiciones climáticas, sin embargo, las asociaciones intrincadas y no lineales entre la demanda y estas influencias presentan desafíos significativos para la predicción. En este estudio, nos proponemos predecir la demanda eléctrica examinando su relación con las variables meteorológicas en la provincia de Entre Ríos, Argentina. Se emplea una red neuronal recurrente, específicamente utilizando arquitectura de memoria a corto y largo plazo (LSTM), para modelar esta relación compleja directamente a partir de datos de entrada sin ingeniería de características previa. Evaluamos y comparamos el rendimiento de este modelo con un método de referencia. El análisis preliminar de datos revela que las temperaturas extremas ejercen un efecto notable en los comportamientos de consumo de energía. Nuestro modelo propuesto alcanza un coeficiente de determinación de 0.77 al comparar la demanda predicha con las observaciones reales, subrayando su efectividad como una posible solución para optimizar las operaciones del sistema en Entre Ríos.

Palabras clave: predicción de demanda eléctrica, condiciones climáticas, aprendizaje profundo, red neuronal artificial.

Received September 2024; Accepted December 2024; Published March 2025

Modelling and forecasting electricity demand: comparing statistical and machine learning approaches

Mariela N. Uhrig[1], Leandro D. Vignolo[2], and Omar V. Müller[3]

[1] Centro de Investigación Científica y de Transferencia Tecnológica a la Producción, CICYTTP-CONICET-PROV. E. R.-UADER
[2] Research Institute for Signals, Systems and Computational Intelligence, sınc(i), FICH-UNL/CONICET
[3] Centro de Estudios de Variabilidad y Cambio Climático, FICH-UNL/CONICET
mnuhrig@conicet.gov.ar, ldvignolo@sinc.unl.edu.ar, ovmuller@unl.edu.ar

Abstract. Reliable forecasting of electricity demand is essential for enhancing distribution system management by optimizing resource utilization, streamlining operational planning, and reducing service disruptions. The fluctuation of electricity demand is influenced by various external factors, such as weather conditions, yet the intricate and non-linear associations between demand and these influences present significant prediction challenges.
In this study, we aim to forecast electricity demand by examining its relationship with weather variables in the province of Entre Ríos, Argentina. A recurrent neural network, specifically using long short-term memory (LSTM) architecture, is employed to model this complex relationship directly from raw input data without prior feature engineering. We assess and compare the performance of this model with a baseline method.
Preliminary data analysis reveals that extreme temperatures exert a notable effect on energy consumption behaviors. Our proposed model achieves a coefficient of determination of 0.77 when comparing predicted demand to actual observations, underscoring its effectiveness as a potential solution for optimizing system operations in Entre Ríos.

Keywords: electricity demand forecast · weather conditions · deep learning · artificial neural network.

1. Introduction

Electricity demand forecast is a key input for operational and strategic decision-making in electricity distribution system management [1]. Having prior information on the electricity needs for the coming days is crucial for companies providing this service, as it enables them to plan their daily operations effectively and determine necessary contingency measures. Thus, they can minimize operational issues, prevent service disruptions or equipment damage due to overloading, and avoid blackouts and financial losses.

The electricity consumption is influenced by different external factors such as meteorological conditions, seasons, holidays, service fee, among others, which make its prediction complex [3]. For instance, the impact of temperature is particularly evident, as people tend to use more electricity for heating or cooling when temperatures are uncomfortable, or they may engage in more indoor activities [4]. Additionally, on holidays, individuals' behavior patterns differ from their usual routines, making public holidays, especially long ones, an important factor influencing power demand [5].

Numerous studies have explored methodologies for forecasting energy demand, ranging from traditional statistical approaches to machine learning techniques such as artificial neural networks (ANNs). The main challenge is the difficulty in capturing non-linear relationships and the limited long-term forecasting capacity. In [6] they tested different ANNs to forecast the daily electricity demand in Greece, including ambient temperature, relative humidity, among others, as input variables. In [7] they applied an ANN model and trend extrapolation method to forecast energy demand. The information used as input was the primary, secondary, and tertiary value of the industry, energy consumption, the level of urbanization, among others, without taking into account weather data. They found that the precision of the neural network was much higher than trend extrapolation. Furthermore, the ANN backpropagation network model was used to forecast Turkey's electricity demand based on different socio-economic indicators [8]. The applications of traditional techniques such as econometric and time series models along with soft computing methods such as ANNs, fuzzy logic, and other models are reviewed in [9] and verified the remarkable performance of the neural network models.

A more recent approach involves using deep neural networks (DNN), which can automatically extract features from raw data to support predictions [14]. Machine learning, particularly deep learning, represents a significant advancement in data analysis, shifting the focus from model-driven to data-driven approaches. This shift allows for the development of models that learn from data patterns, making them well-suited for complex applications. In the context of time series data analysis, recurrent neural networks (RNNs) are more adept than traditional methods at capturing temporal dependencies. A notable variant of RNNs, long-short-term memory networks (LSTMs), incorporate specialized gates in their architecture to retain information over time and integrate both past and current data for forecasting. These features enable LSTMs to surpass conventional ARIMA-based models, especially for long-term prediction tasks [13]. While DNN techniques have shown promising results across various applications in recent years [10,11,12,13], their application to energy demand forecasting based on meteorological variables remains relatively unexplored.

This work focuses on the province of Entre Ríos, Argentina, where Energía de Entre Ríos S.A. (ENERSA) is the sole provider of energy services. Currently, ENERSA relies on rudimentary tools to make decisions regarding energy demand. For instance, they adjust the service cost according to an analysis of the seasonal variations in the demand from historical records. Other aspects they take into account for decision-making are the GPD (Gross Domestic Product) forecasts or available economic forecasts, the weather seasonal forecast by the SMN (from Spanish, Servicio Meteorológico Nacional), and the behavior of the activity of large customers and the residential sector [O. Bustamante, 2019, personal communication]. In this study, we contribute to the needs of ENERSA by developing an energy demand forecast model based on machine learning methods, aiming to improve decision-making processes and optimize electricity distribution in the province.

The proposed forecast is based on historical demand and meteorological data at daily time-step. We initially focused on developing a method to forecast the electricity demand in Entre Ríos just for one day using information from the previous 7 days. Though, the general goal is to extend the forecast to 7 days, so that the company may identify critical days of the coming week that may lead to failures in the regional system. The rest of the paper is structured as follows: Section 2 introduces a method frequently used as a benchmark, the proposed model, and the metrics used in the performance evaluation. Section 3 describes the dataset used and the dataset partitioning for the experiments. Section 4 analyses the main feature of the energy demand data and its relationship with meteorological conditions, and also evaluates the results of the performed experiments. Finally, Section 5 summarizes the main findings and guidelines for future work.

2. Methods

In time series analysis, it is crucial to decompose a series into latent components that represent different temporal patterns, such as trend, seasonality, cycles, and high-frequency variations. Selecting the appropriate forecasting model depends heavily on whether the underlying process is linear or nonlinear. Consequently, multiple models are in general tested, and the one that best predicts out-of-sample data is chosen. In this section, we focus on SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous variables) model, which is an extension of ARIMA (AutoRegressive Integrated Moving Average), and a recurrent ANN-based model

2.1. SARIMAX

The ARIMA model serves as a foundational method for time series forecasting and is frequently used as a benchmark due to its simplicity and effectiveness in capturing linear dependencies within a time series [15,16,17]. An ARIMA(p, d, q) model predicts future values as a linear function of past observations and residual errors. Its general form is given by:

$$(1-B)^d y_t = \theta_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \qquad (1)$$

where:

- B is the backshift operator such that $B * y_t = y_{t-1}$, i.e., it shifts the series one period back,
- d represents the order of differencing required to achieve stationarity of the series,
- y_t is the value of the series at time t,
- ϕ and θ are the coefficients of the autoregressive (AR) and moving average (MA) components, respectively,
- p and q are the orders of the AR and MA components, respectively,
- ε_t represents the error term, assumed to follow a white noise process.

The term $(1 - B)^d$ indicates that the series is differenced d times to eliminate trends and seasonality, allowing the ARIMA model to be applied to non-stationary series. The Seasonal ARIMA (SARIMA) model extends ARIMA to accommodate seasonality through additional parameters for seasonal autoregression (P), differencing (D), and moving average (Q), as well as a seasonal period (S). This leads to the SARIMA(p, d, q)(P, D, Q)$_S$ formulation:

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D y_t = \theta_q(B)\Theta_Q(B^S)\varepsilon_t \tag{2}$$

where:

- $\phi_p(B)$ is the AR polynomial of order p,
- $\Phi_P(B^S)$ is the seasonal autoregressive (SAR) polynomial of order P,
- $(1-B)^d$ is the non-seasonal differencing operator, applied d times to ensure stationarity,
- $(1 - B^S)^D$ is the seasonal differencing operator, applied D times with a seasonal lag S,
- y_t is the value of the time series at time t,
- $\theta_q(B)$ is the MA polynomial of order q,
- $\Theta_Q(B^S)$ is the seasonal moving average (SMA) polynomial of order Q,
- ε_t are the residuals (or errors) at time t, assumed to be white noise with a mean of zero and constant variance.

The SARIMAX model further extends SARIMA by incorporating external predictors, or exogenous variables, to model the influence of external factors on the target variable. The general structure of the SARIMAX(p, d, q)(P, D, Q)$_S$ model, with the inclusion of exogenous predictors X, is described as follows:

$$y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \cdots + \beta_k X_{k,t} + \omega_t \tag{3}$$

where:

- y_t is the value of the target variable at time t,
- β_0 is the intercept term, representing the baseline level of y_t when all predictors are zero,
- β_i is the coefficients associated with each exogenous variable $X_{i,t}$, quantifying the effect of $X_{i,t}$ on the target variable y_t,

- $X_{i,t}$ is the i-th exogenous variable at time t, representing external predictors that influence y_t (examples could include weather conditions, economic indicators, or policy changes),
- ω_t is the residual error term.

By integrating this regression equation into the SARIMA framework, SARIMAX models can effectively capture complex dynamics influenced by internal and external factors.

2.2. Long Short-Term Memories

ANNs are powerful functions that simulate how the human brain processes information. An ANN is composed of a network of processing nodes (or neurons), which perform numerical manipulations and are interconnected in a specific order. Historical data can be used by ANNs to predict the future values of noisy multivariate times series [19]. Feed-forward ANNs can be applied to sequential or time series data, however, there are several issues that render them unsuitable for these types of problems.

RNNs are able to capture the dynamics of sequences via recurrent connections, overcoming the limitations of feed-forward ANNs. LSTM is a special type of RNN, which has the capability to learn longer dependencies in the dataset [19,20]. LSTMs have been effectively used in several time series forecasting applications, e.g. in [21].

The simple LSTM architecture with forget gate [23] is depicted in Figure 1. The forget gate determines the unnecessary component from the previous cell state which can be computed as follows:

$$f_t = W_f X_t + U_f h_{t-1} + b_f \tag{4}$$

where f_t is the forget gate activation at time t. This vector determines which part of the past information to forget, W_f is the weight matrix applied to the input X_t at time t, U_f is the recurrent weight matrix applied to the previous hidden state h_{t-1} (the output of the LSTM cell in the previous time step) and b_f the ias term for the forget gate. The state update of the cell is determined by the input gate and *tanh* layer, which is calculated as

$$i_t = W_i X_t + U_i h_{t-1} + b_i, \tag{5}$$

where i_t is the input gate activation at time t. This vector decides how much new information to store in the cell state, W_i is the weight matrix applied to the input X_t, U_i recurrent weight matrix applied to the previous hidden state h_{t-1}, b_i the bias term for the input gate.

The candidate cell state at time t is a new candidate memory created from the input and the previous hidden state which is calculated as

$$C_t^* = \tanh(W_c X_t + U_c h_{t-1} + b_c) \tag{6}$$

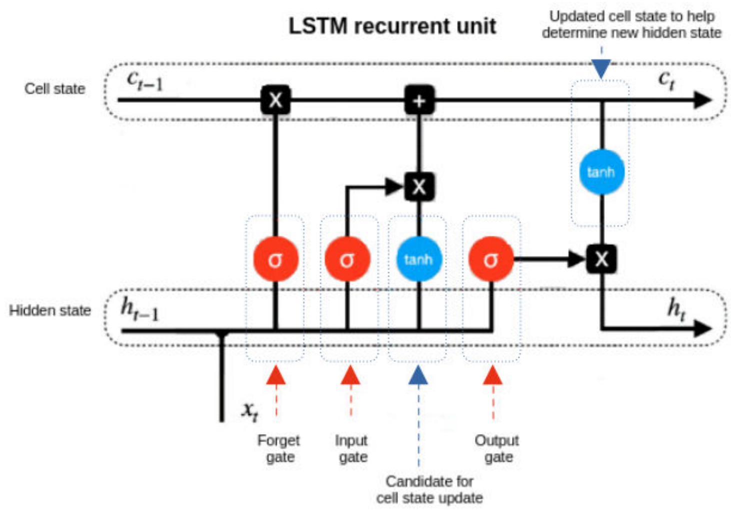

Fig. 1. LSTM unit architecture (Based from [22]).

where tanh hyperbolic tangent activation function, ensuring the output values are between *-1* and *1*, W_c is the weight matrix applied to the input X_t, U_c the recurrent weight matrix applied to the previous hidden state h_{t-1} and b_c the bias term for the candidate cell state.

The cell state at time t is a combination of the previous cell state, modulated by the forget gate, and the new candidate memory, modulated by the input gate which is calculated as

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \tag{7}$$

where $f_t C_{t-1}$ the element-wise (Hadamard) product between the forget gate activation and the previous cell state C_{t-1}, representing the part of the previous cell state that is retained and $i_t \tilde{C}_t$ the element-wise (Hadamard) product between the input gate activation and the candidate cell state, representing the new information being stored.

The output from the cell to the next cell is calculated by the output gate as

$$o_t = W_o X_t + U_o h_{t-1} + b_o \tag{8}$$

where o_t the output gate activation at time t. This determines which part of the cell state will be used for the output. W_o the weight matrix applied to the input X_t, U_o the recurrent weight matrix applied to the previous hidden state h_{t-1} and b_o bias term for the output gate.

$$h_t = o_t \tanh(C_t) \tag{9}$$

where h_t hidden state at time t. This is the output of the LSTM cell, based on the output gate activation and the current cell state, o_t output gate activation,

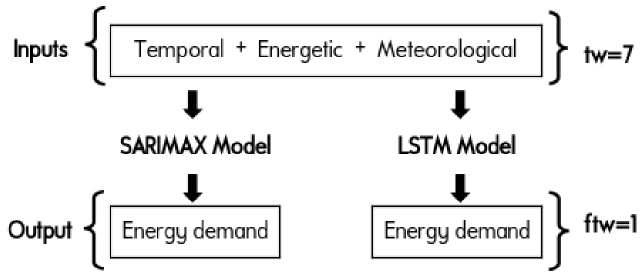

Fig. 2. Overview of the models setup.

modulating the influence of the cell state on the output and $\tanh(C_t)$ application of the hyperbolic tangent function to the cell state C_t, ensuring the output values are between *-1* and *1*.

2.3. Model setup

Figure 2 illustrates the basic flow of data of the two models employed in this study for energy demand forecasting. The models, SARIMAX and LSTM, are fed with a combination of temporal, energetic, and meteorological inputs. These inputs are processed independently by each model, resulting in separate forecasts of energy demand. The figure highlights both models sharing the same set of inputs but producing independent outputs. The notation 'tw=7' indicates a 7-day input window, while 'ftw=1' refers to a 1-day forecast window. This means the model is trained on 7 days of historical data to predict the energy demand for the next day (eighth day). Observed electricity demand is the target variable, as seen in Table 1.

2.4. Metrics for performance evaluation

For evaluating training and prediction performance, we considered the Root Mean Squared Error (RMSE), which measures the average magnitude of the errors between predicted and actual values. A lower RMSE indicates better model performance. The RMSE can vary from 0 (no error) to infinity, with higher values indicating larger errors and poorer model performance. Additionally, we use the coefficient of determination (R^2), which represents the proportion of variance in the dependent variable that can be predicted from the independent variables. R^2 values range from 0 to 1, with a value of 1 indicating perfect prediction, where the model explains all the variance in the target variable.

Table 1. Input and output variables.

Input			Output
Temporal	**Energetic**	**Meteorological**	
year month day encoded day	users	minimum, mean, maximum, and skin temperature downward and net thermal and solar radiation latent and sensible heat flux precipitation runoff surface pressure total cloud cover wind speed relative humidity	energy demand

3. Data description

3.1. Dataset

The electricity and meteorological data are daily time series for the period 2012 to 2023 obtained from two sources: ENERSA and the fifth-generation ECMWF atmospheric reanalysis of the global climate (ERA5). After preprocessing, the final dataset includes a total of 21 variables summarized in Table 2, all of them offered at a daily time-scale. The encoded representation of the weekday assigns 1 for Monday, 2 for Tuesday, ..., 7 for Sunday, and 8 for holidays (regardless of the actual weekday). This parameter is used to distinguish holidays from workdays and weekends, as holidays typically have a different energy demand compared to other days.

As depicted in Figure 7 the data was split as follows: 80% for training, 10% for validation, and 10% for testing. The split is performed using consecutive periods of daily data, for instance, January 2012 to February 2021 for training, March 2021 to April 2022 for validation, and May 2022 to May 2023 for testing. This approach ensures that each subset of data (training, validation, and testing) preserves the inherent temporal structure which includes a marked annual cycle. This is crucial for capturing seasonal fluctuations and other temporal patterns that could impact the predictive performance of the model.

3.2. General features of the demand series

An exploratory analysis was conducted to first understand the main features of the energy demand time series and then examine its relationship with meteorological variables. The electricity demand over the entire period shows a general upward trend in demand over the years (see black curve in Figure 3). The long-term positive trend includes two declines in 2017 and 2019, which can

Table 2. Description of input variables.

Variable	Description
Year	Year of the data (2012-2023).
Month	Month of the data (1-12).
Day	Day of the data (1-31).
Encoded day	Encoded representation of the weekday (1-8).
Users	Number of electricity service users.
Minimum temperature (tmin)	Minimum 2-m temperature.
Mean temperature (tmean)	Mean 2-m temperature.
Maximum temperature (tmax)	Maximum 2-m temperature.
Skin temperature (skt)	Mean skin surface temperature.
Solar radiation downwards (ssrd)	Solar radiation reaching the surface.
Thermal radiation downwards (strd)	Thermal radiation returning to the surface.
Net shortwave radiation (ssr)	Net balance of solar radiation at the surface.
Net longwave radiation (str)	Net balance of thermal radiation at the surface.
Latent heat flux (slhf)	Flux of latent heat at the surface.
Sensible heat flux (slhf)	Flux of sensible heat at the surface.
Total precipitation (tp)	Total amount of precipitation.
Surface runoff (sro)	Water flowing over the surface.
Surface pressure (sp)	Surface air pressure.
Total cloud cover (tcc)	Percentage of sky covered by clouds.
Wind speed (ws)	Speed of the wind.
Relative humidity (rh)	Water vapor content in air.

be attributed to the tariff increases applied in Argentina during those years. These fluctuations highlight the influence of external factors such as economic conditions and population growth on energy consumption patterns. From the annual cycle analysis represented in Figure 4, we observe that energy consumption follows a clear 'W' pattern, with high consumption during the summer and winter months, and lower consumption during the intermediate seasons. This pattern is also evident in Figure 3 (see blue curve). This seasonal evolution can be interpreted as being directly related to temperature-driven demand. In summer, electricity demand rises for cooling purposes, while in winter, electricity is used for heating. However, during winter, electricity demand for heating is supplemented by gas services, as is common in this region and most of Argentina, which explains why the peak in winter is not as high as in summer. Another factor is that summer temperatures in Entre Ríos are more extreme and sustained over longer periods than winter temperatures, which are generally milder. Focusing on the dynamics of energy demand during the week, the left panel of Figure 5 shows consumption peaks from Tuesday to Thursday, while Sunday records the lowest levels of consumption. When holidays are isolated in the analysis, as shown in the above panel of Figure 5, the boxplot suggests that energy consumption during holidays is comparable to that on weekends. These findings align with previous studies [24,25,26].

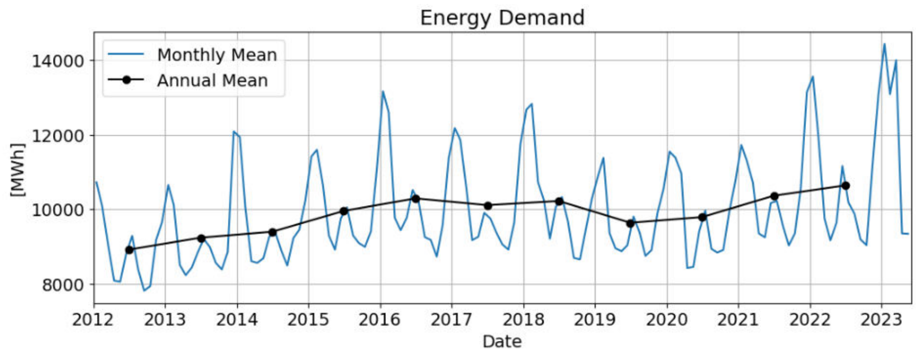

Fig. 3. Energy demand [MWh] time series at monthly and annual time-scale.

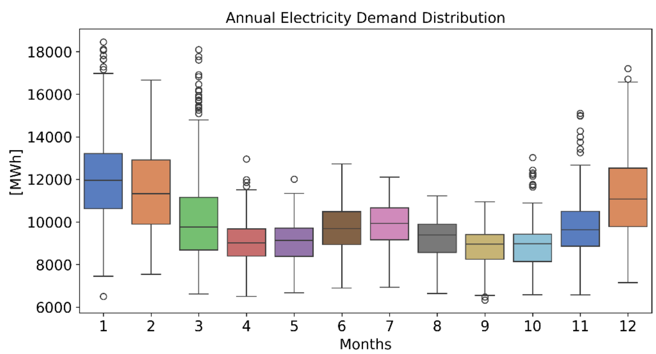

Fig. 4. Annual cycle of energy demand [MWh] shown with monthly boxplots.

Lastly, we evaluate the relationship between input variables and energy demand with scatter plots (Figure 6). The results underscore the dominant influence of temperature on energy consumption, with a clear pattern of demand that increases at both low and (mainly) high temperatures while decreasing at moderate temperatures, as illustrated in Figure 4. This temperature-energy demand relationship reveals a pronounced non-linear behavior. In contrast, the effects of other variables on energy consumption are less straightforward. While no direct correlation is immediately evident, it is plausible that non-linear interactions between these factors and energy demand exist, which are not apparent through simple observation. In this context, ANNs offer a valuable tool for identifying and modeling these complex, hidden relationships, allowing for a more comprehensive understanding of the various factors that influence energy consumption.

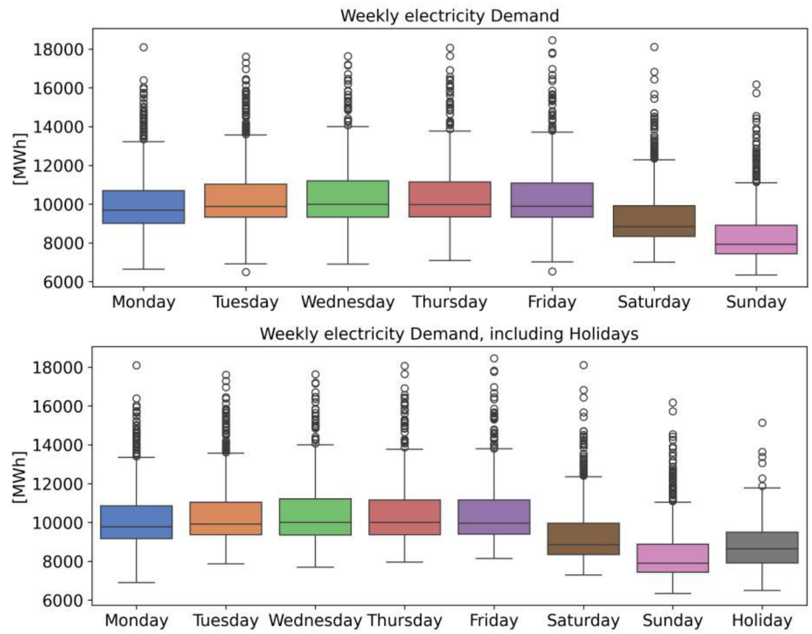

Fig. 5. Above: Week cycle of energy demand shown with daily boxplots. Below: As above, but separating holidays.

4. Experiments and results

4.1. Models evaluation

For the SARIMAX model, six different parameter configurations were tested to identify the optimal setup. The best performance was achieved with $p=7$, $d=0$, and $q=2$, where $p=7$ represents the order of the seasonal autoregressive component, $d=0$ indicates that no differencing was applied, and $q=2$ corresponds to the order of the moving average component. Additionally, exogenous variables were included in the model, taking advantage of SARIMAX's capability to account for external factors influencing the target variable. In Table 3, six different configurations used in the experiments along with their corresponding metric values for the test dataset can be observed. For LSTM models, a total of thirty-one different settings were explored using the options described in section II. 1-layered, 2-layered, and 3-layered LSTM structures are used for modeling with different amounts of neurons in each hidden layer. After the LSTM layers, a fully connected layer is incorporated to further process the learned features and capture complex patterns in the data. The number of layers and units in

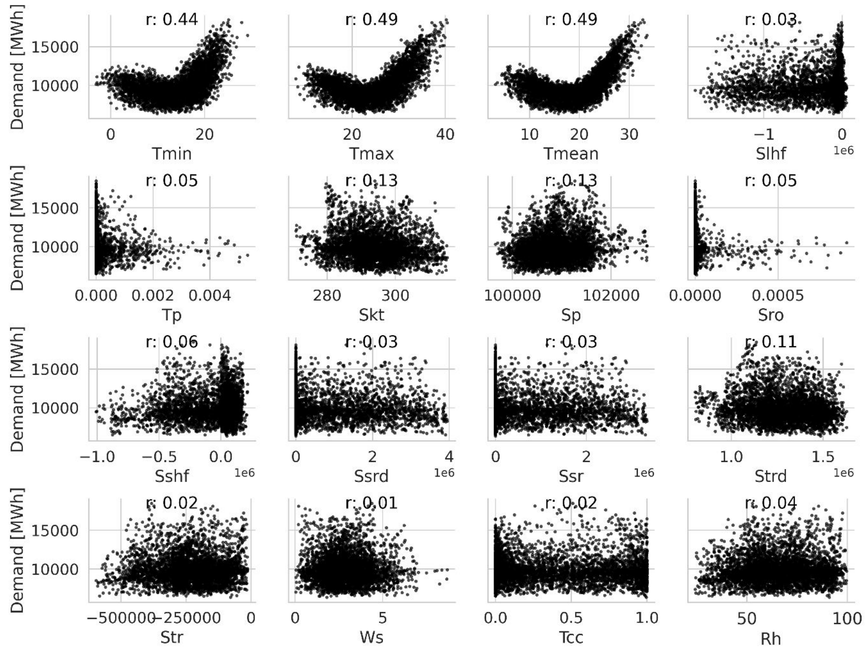

Fig. 6. Scatter plots of input meteorological variables versus energy demand at daily time-scale.

the LSTM was selected via trial-and-error, as well as the training parameters learning rate, and batch size. The performance of each architecture considered was evaluated using different configurations regarding learning rate, batch size, and number of epochs. In Table 4, some of the configurations used in the experiments along with their corresponding metric values for each dataset can be observed. For the experiments in the table, Adam optimizer [27] was used and a patience of 150 epochs for early stopping. The models were trained for 50 to 400 epochs. It was observed in the experiments that larger batch sizes and increasing the number of neurons in the layers led to a loss of generalization capacity and performance and required a greater number of epochs during training. The best configurations are those with, 2 layers and 5 neurons, and, 2 layers and 10 neurons, all yielding a test R^2 above 0.70, as highlighted in Table 1.

For the best LSTM model, which achieves an R^2 value of 0.77 on the test set, the corresponding R^2 value on the training set is 0.64 (see Table 4). This discrepancy suggests that the model is not overfitted and acquired generalization capabilities. On the other hand, the best approximation of SARIMAX models obtained R^2=0.40 and RMSE=0.19 (see Table 3), suggesting it may not be adequate to solve complex non-linear problems.

As shown in Figure 8, a forecasting comparison is presented between two models —SARIMAX and LSTM— and the observed demand. The results sug-

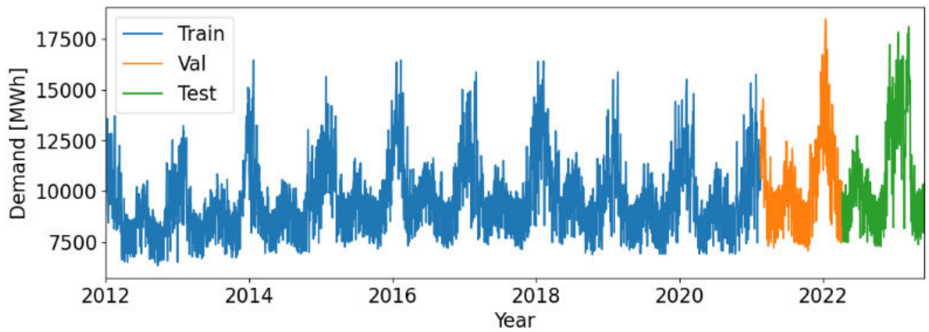

Fig. 7. Data split: 80% for training (blue), 10% for validation (orange), and 10% for testing (green).

Table 3. Configurations evaluated for the SARIMAX model and their metrics.

			Test	
p	**d**	**q**	**RMSE**	**R^2**
7	0	1	19.08	0.39
7	1	1	29.88	-0.49
7	2	1	2223.32	-82.63
7	**0**	**2**	**18.96**	**0.40**
7	0	3	19.17	0.39
7	2	7	2094.14	-73.30

gest that both models follow the general trend of the observed data, with the LSTM model significantly outperforming the SARIMAX model, particularly in capturing finer fluctuations.

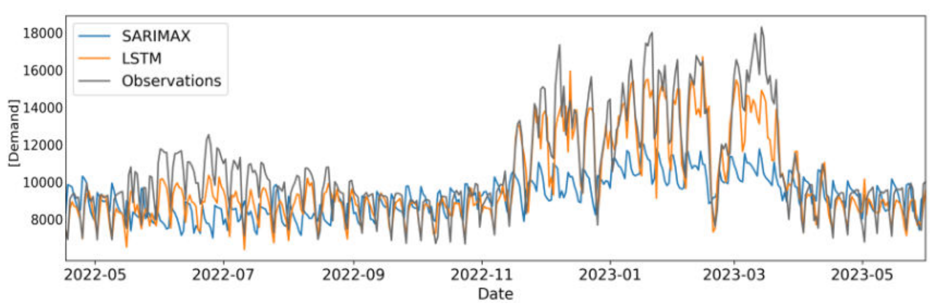

Fig. 8. Model performance comparison: SARIMAX vs. LSTM for demand forecasting

Table 4. Configurations evaluated for the LSTM model and their metrics.

				Train		Validation		Test	
Layers	**Units**	**Lr**	**Batch**	**RMSE**	R^2	**RMSE**	R^2	**RMSE**	R^2
1	10	0.0001	4	0.09	0.73	0.16	0.38	0.20	0.29
1	20	0.0001	4	0.08	0.78	0.11	0.68	0.17	0.51
1	20	0.0001	8	0.09	0.69	0.14	0.53	0.20	0.35
1	40	0.0001	4	0.07	0.84	0.12	0.62	0.19	0.41
1	40	0.0001	8	0.08	0.79	0.13	0.59	0.20	0.37
1	40	0.0001	10	0.08	0.78	0.13	0.60	0.18	0.43
1	40	0.001	4	0.05	0.92	0.12	0.63	0.15	0.59
1	40	0.001	8	0.05	0.90	0.13	0.56	0.17	0.50
1	40	0.001	10	0.05	0.90	0.13	0.56	0.19	0.41
2	5	0.001	10	0.10	0.67	0.12	0.64	0.17	0.50
2	5	0.001	16	0.09	0.73	0.13	0.58	0.18	0.42
2	**5**	**0.001**	**4**	**0.08**	**0.78**	**0.10**	**0.76**	**0.12**	**0.75**
2	5	0.001	8	0.08	0.77	0.11	0.69	0.17	0.48
2	**5**	**0.001**	**10**	**0.10**	**0.64**	**0.10**	**0.73**	**0.12**	**0.77**
2	5	0.001	16	0.10	0.62	0.11	0.68	0.14	0.68
2	5	0.001	32	0.10	0.64	0.12	0.65	0.18	0.48
2	5	0.001	8	0.10	0.63	0.13	0.59	0.17	0.53
2	5	0.001	16	0.12	0.45	0.14	0.49	0.19	0.40
2	5	0.001	32	0.12	0.51	0.14	0.52	0.18	0.47
2	**10**	**0.001**	**4**	**0.08**	**0.77**	**0.10**	**0.75**	**0.12**	**0.75**
2	10	0.001	8	0.08	0.76	0.10	0.73	0.14	0.67
2	**10**	**0.001**	**10**	**0.08**	**0.75**	**0.10**	**0.75**	**0.13**	**0.72**
2	10	0.001	16	0.10	0.67	0.12	0.63	0.17	0.52
2	20	0.001	4	0.08	0.78	0.12	0.64	0.17	0.52
2	20	0.001	8	0.08	0.75	0.11	0.07	0.15	0.62
2	20	0.001	10	0.09	0.68	0.13	0.60	0.17	0.49
2	20	0.001	16	0.09	0.71	0.12	0.64	0.17	0.50
2	20	0.001	32	0.09	0.72	0.13	0.60	0.19	0.37

5. Conclusions and future work

This paper presents a recurrent artificial neural network architecture for forecasting electricity demand using a case study from the province of Entre Ríos, Argentina. The study innovates by incorporating historical energy demand and meteorological variables as inputs to the forecasting model, providing a comprehensive approach to forecasting energy demand. The preliminary results show a consistent level of competence in accurate demand forecasting. The proposed model scores 0.77 in the coefficient of determination with 0.12 RMSE when comparing the predicted electricity demand with observations. To potentially improve the performance of the models, it could be beneficial to extend the exploration of the hyperparameters or modify the early stopping criteria.

Future work will focus on extending the proposed model to analyze energy demand patterns at a more granular level by applying it to individual transformer stations or localities, as they can present particular energy demand patterns. While some localities exhibit notable seasonal fluctuations, others exhibit steady demand throughout the year. Factors like regional economic activity and demographic characteristics can be accused of this variation. It would be possible to capture the specific features of the energy demand at every place by expanding the suggested model to every transformer station or locality. This could lead to improved energy efficiency and more precise demand predictions.

Lastly, exploring other deep learning approaches, such as combining convolutional neural networks with LSTM models, and the use of transformers, could further enhance the accuracy and robustness of electricity demand forecasting.

References

1. Pai, Hong, Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms, Electric Power Systems Research, Volume 74, 3, 417-425, ISSN 0378-7796 (2005)
2. Barak, S. and S.S. Sadegh, Forecasting energy consumption using ensemble ARIMA– ANFIS hybrid algorithm. International Journal of Electrical Power & Energy Systems (2016). 82: p. 92-104.
3. Zhao, H.-x. and F. Magoulès, A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews (2012). 16(6): p. 3586-3592.
4. Wang Y, Zhang N, Chen X, et al. A Short-Term Residential Load Forecasting Model Based on LSTM Recurrent Neural Network Considering Weather Features. Energies. 14(10):2737. (2021)
5. Wan He. Load Forecasting via Neural Networks. Procedia Computer Science. Volume 122. Pages: 308-314. ISSN 1877-0509. (2017) https://www.sciencedirect.com/science/article/pii/S1877050917326170 https://doi.org/https://doi.org/10.1016/j.procs.2017.11.374
6. Ekonomou, L., Oikonomou, D.S.: Application and comparison of several artificial neural networks for forecasting the Hellenic daily electricity demand load. In: Proceedings of the 7th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and data bases, pp. 67–71. World Scientific and Engineering Academy and Society (WSEAS) (2008).
7. Wang, J.M., Liang, X.H.: The forecast of energy demand on artificial neural network. In: Artificial Intelligence and Computational Intelligence. AICI'09. International Conference. vol. 3, pp. 31—35. IEEE Computer Society, Washington, DC, United States (2009)
8. Kavaklioglu, K.: Modeling and prediction of Turkey's electricity consumption using Support Vector Regression. Appl. Energy (2011). 88(1), 368–375
9. Ghalehkhondabi, I., Ardjmand, E., Weckman, G.R. et al. An overview of energy demand forecasting methods published in 2005–2015. Energy Syst (2017). 8, 411–447. https://doi.org/doi:10.1007/s12667-016-0203-y
10. Rivenson, Y., Zhang, Y., Günaydın, H. et al. Phase recovery and holographic image reconstruction using deep learning in neural networks. Light Sci Appl 7, 17141. (2018) https://doi.org/doi.org/10.1038/lsa.2017.141
11. Belthangady, C., Royer, L.A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. Nat Methods (2019). 16, 1215–1225. https://doi.org/doi.org/10.1038/s41592-019-0458-z

12. Pereira, T.D., Aldarondo, D.E., Willmore, L. et al. Fast animal pose estimation using deep neural networks. Nat Methods (2019). 16, 117–125. https://doi.org/doi.org/10.1038/s41592-018-0234-5
13. M. M. Sachin and Melvin Paily Baby and Abraham Sudharson Ponraj. "Analysis of energy consumption using RNN-LSTM and ARIMA Model", Journal of Physics: Conference Series. (Dec, 2020) vol.1716, n°1, sp. 012048. https://doi.org/dx.doi.org/10.1088/1742-6596/1716/1/012048
14. P. Vincent, H. Larochelle, Y. Bengio, and P. Manzagol, Extracting and composing robust features with denoising autoencoders, in Proc. 25th Int. Conf. Mach. Learn. (ICML'08) , pp. 1096—1103. Association for Computing Machinery, New York (2008)
15. R. Sharda, R.B. Patil, Neural networks as forecasting experts: an empirical test, in: Proceedings of the International Joint Conference on Neural Networks, Washington, D.C., Vol. 2 (1990), pp. 491– 494.
16. Z. Tang, C. Almeida, P.A. Fishwick, Time series forecasting using neural networks vs Box–Jenkins methodology, Simulation 57 (1991) 303–310.
17. Z. Tang, P.A. Fishwick, Feedforward neural nets as models for time series forecasting, ORSA J. Comput. 5 (1993) 374–385.
18. Brockwell, P. J., & Davis, R. A. Introduction to time series and forecasting (3a ed.). Springer, New York (2016) https://doi.org/10.1007/978-1-4757-2526-1
19. Adamowski, J., Fung Chan, H., Prasher, S.O., Ozga-Zielinski, B., Sliusarieva, A.: Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. Water Resources Research 48(1) (2012)
20. Q. Gao, Y. Liu, J. Yang, and Y. Hong, Çomparative Research on Electricity Consumption Forecast Based on Deep Learning" 2nd International Conference on Artificial Intelligence and Education (ICAIE), Dali, China. (2021) pp. 213-217.https://doi.org/doi:10.1109/ICAIE53562.2021.00052
21. Junior Clony, Gusmão Pedro, Moreira José and Tome Ana Maria M. Time Series Forecasting in Retail Sales Using LSTM and Prophet. Handbook of Research on Applied Data Science and Artificial Intelligence in Business and Industry Copyright. Pages: 22 ISBN13: 9781799869856.ISBN10: 1799869857.EISBN13: 9781799869863. (2021) https://doi.org/doi:10.4018/978-1-7998-6985-6.ch011
22. Magallanes-quintanar, R., Galván Tejada C., Galván Tejada, J., and Gamboa-Rosales, H., and Méndez-Gallegos, S. and Garcia, Antonio. Neural Hierarchical Interpolation for Standardized Precipitation Index Forecasting. Atmosphere. Vol. 15. Pages 912. (2024) https://doi.org/= doi:10.3390/atmos15080912
23. R. Jozefowicz, W. Zaremba and I. Sutskever, An empirical exploration of recurrent network architectures, in: Int. Conf. Mach. Learn, pp. 2342-–2350. Association for Computing Machinery, New York (2015) https://doi.org/doi:10.1109/CVPR.2015.729876
24. Gutiérrez, E. La demanda residencial de energía eléctrica en la Comunidad Autónoma de Andalucía; un análisis cuantitativo. Tesis de Doctorado. Facultad de Ciencias Enconómicas y Empresariales, Universidad de Sevilla (2003).
25. Medina, S.; García, J. Predicción de demanda de energía en Colombia mediante un sistema de inferencia difuso neuronal. Revista Energética 33,15-24. (2005).
26. Murillo, J; Trejos, A. Carvajal, P. Estudio del pronóstico de la demanda de energía eléctrica utilizando modelos de series de tiempo. Scientia et Technica (2003) 23, 37-42.

27. Kingma, D. P., Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. (2014) Retrieved from https://arxiv.org/abs/1412.6980