IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

ELECTRICITY DEMAND FORECAST IN INDUSTRY USING DEEP LEARNING

[1,2,3,4]**Mihir Harishbhai Rajayaguru, Kaviya.T, Gokul A, Harmanpreet.H**
[1,2,3,4]Assistant Professor
[1]Department of Computer Engineering
[2]Department of Electrical and Electronic Engineering,
[3]Department of Computer Science Engineering,
[4]Department of Civil Engineering
[1]Madhuben and Bhanubhai Patel Institute of Technology,
[2]Sri Ranganathar Institute of Engineering and Technology,
[3]Paavai College of Engineering,
[4]Ludhiana Group of Colleges
[1,2,3,4]Anand , Gujarat, Coimbatore, Pachal, Namakkal, Ludhiana, Punjab.

ABSTRACT: This article presents electricity demand forecasting models for industrial and residential facilities, developed using ensemble machine learning strategies. Short term electricity demand forecasting is beneficial for both consumers and suppliers, as it allows improving energy efficiency policies and the rational use of resources. Computational intelligence models are developed for day-ahead electricity demand forecasting. An ensemble strategy is applied to build the day-ahead forecasting model based on several one-hour models. Three steps of data preprocessing are carried out, including treating missing values, removing outliers, and standardization. Feature extraction is performed to reduce overfitting, reducing the training time and improving the accuracy. The best model is optimized using grid search strategies on hyperparameter space. Then, an ensemble of 24 instances is generated to build the complete day-ahead forecasting model. Considering the computational complexity of the applied techniques, they are developed and evaluated on the National Supercomputing Center (Cluster-UY), Uruguay. Three different real data sets are used for evaluation: an industrial park in Burgos (Spain), the total electricity demand for Uruguay, and demand from a distribution substation in Montevideo (Uruguay). Standard performance metrics are applied to evaluate the proposed models. The main results indicate that the best day ahead model based on Extra Trees Regressor has a mean absolute percentage error of 2.55% on industrial data, 5.17% on total consumption data and 9.09% on substation data.

KEYWORDS: Energy; forecasting; artificial intelligence

INTRODUCTION

Uncertainty is a specific characteristic of the energy sector. Although decisions in the energy sector are generally not based on predictable outcomes, some variables that affect decision making can be predicted, with certain degree of confidence, using information from different sources [1, 2]. Examples of useful information for decision making is that related to natural variables (temperature, wind speed, etc.). Information related to energy consumption and demand profiles of users is valuable too. Furthermore, new sources of renewable energy generation developed in the last 30 years are directly related to natural variables, and the corresponding information is often incorporated in prediction models for decision making [3].

Due to the aforementioned reasons, a large number of stochastic variables must be taken into account to improve operational decision making, but also to assure that the derived actions are feasible from an economic point of view. When considering a large number of variables, the complexity of the underlying models notoriously increases. However, the increase in complexity associated with the number of variables is partly compensated because the hardware infrastructure to perform computations on large volumes of data has developed strongly.

New challenges have emerged from the described reality. A very relevant one is related to the development of an intelligent system to take advantage of new sources of information and available data. Classic statistical models, that were useful for making predictions some decades ago, have limitations in this new context. Computational intelligence methods have demonstrated excellent forecasting accuracy in different areas, in recent years [4–6]. These methods are robust and tolerant to uncertainty, and they are able to learn the most relevant features of the considered data to provide a precise forecast, thus providing excellent results by excluding non-relevant information and focusing on the most useful data.

Overall, the major contributions of the research reported in this article are: *i*) the evaluation and comparison of computational intelligence models applied to forecasting the demand of an industrial park in Spain, the demand of a substation in Uruguay and the total demand of Uruguay. Also *ii*) the optimization of the proposed models using the high performance computing infrastructure of the National Supercomputing Center, in Uruguay.

- **Energy demand forecasting**

This section introduces the energy demand forecasting problem, describes forecasting techniques, and reviews related works.

- **General considerations**

The energy demand forecasting problem is usually solved applying mathematical methods using historical data for prediction. There is no a general method that can be used in all types of energy demand forecasting. Thus, an appropriate method must be found for each demand profile. Using historical data of a particular demand profile is common in practice to determine the most effective algorithm. The problem can be classified by the time horizon to forecast: ultra short-term demand forecasting (up to a few minutes ahead), short-term demand forecasting (up to few days ahead), medium-term demand forecasting (up to few month ahead), and long-term demand forecasting (years ahead). Different techniques are applied when considering each time horizon. This work focuses on short-term demand forecasting using historical data.

The energy management and operation of electric grids becomes highly difficult and uncertain, particularly when new technologies are incorporated. The power demand of end customers is versatile and changes on hourly, daily, weekly, and seasonally basis. Hence, there is a real need of developing models for accurately forecasting at different time horizons, depending on the management

is the observed value at time *t*.

$$pred_{(t+1)} = model_1(obs_t, obs_{(t-1)}, ..., obs_{(t-n)})$$
$$pred_{(t+2)} = model_2(obs_t, obs_{(t-1)}, ..., obs_{(t-n)})$$
$$\cdots$$

$$pred_{(t+24)} = model_{24}(obs_t, obs_{(t-1)}, ..., obs_{(t-n)}) \quad (1)$$

This work focuses on both industrial and residential power consumption. Residential power profiles are usually variable, mainly depend on the time of the day and the day of the week, but they also depend on occasional vacations and other factors [8]. On the other hand, industrial power profiles tend to be stable, due to the needs of industrial processes themselves.

- **Problem formulation and strategies**

This section describes the problem formulation and the studied strategies for electric demand forecasting.

Relation between one hour and 24 hour forecasting. This article focuses on applying computational intelligence methods to develop a model for forecasting electricity demand 24 hours ahead. When historical data are available with hourly frequency, is natural to develop a model that predicts next hour. From that model, a multi-step forecasting model can be constructed (i.e., 24 steps in the future).

- *Recursive strategies* apply a one-step model (recursively), multiple times. Predictions for previous time steps are used as input for the prediction on the following time step. The structure to develop for a recursive strategy is presented in Equation 2.

$$pred_{(t+1)} = model_1(obs_t, obs_{(t-1)}, ..., obs_{(t-n)})$$
$$pred_{(t+2)} = model_1(pred_{(t+1)}, obs_t, obs_{(t-1)}, ..., obs(t-n+1))$$
$$...$$
$$pred_{(t+24)} = model_1(pred_{(t+23)}, pred_{(t+22)}, ..., pred(t+1), obs(t-n+23))$$

- *Hybrid strategies* combine the previously described to get benefits form both methods. A separate model is constructed for each time step to be predicted. Each model may use the predictions made by models at prior time steps as input values. For example, using all known prediction, a hybrid strategy produces the structure in Equation 3.

$$pred_{(t+1)} = model_1(obs_t, obs_{(t-1)}, ..., obs(t-n))$$
$$pred_{(t+2)} = model_1(pred_{(t+1)}, obs_t, ..., obs(t-n))$$
$$.$$
$$pred_{(t+24)} = model_1(pred_{(t+23)}, pred_{(t+22)}, ..., obs_t, ..., obs(t-n)) \quad (2)$$

One hour forecasting model training. Section 2.3 reviews different approaches and methods for short term demand forecasting. This work explores the use of machine learning techniques, mainly those based on model ensembles. Feature selection is commonly applied in this kind of problems due to several reasons. Simpler models are easier to interpret, and have shorter training times. Also, the size of the model using less features is smaller, mitigating the *curse of dimensionality* [9]. But the main reason to apply feature selection is to reduce overfitting, enhancing generalization of the model to unseen data.

Complete model. After obtaining a one-hour model with optimized parameters, it is trained for the next hour taking all steps mentioned. Thus, 24 four different instances of this model are trained, one for each of the next 24 hours. Then, the hybrid strategy described in Equation 3 is applied to build a 24-hour forecasting model. The complete model is evaluated on testing data and results are reported.

- **Related works**

Several methods have been proposed for electricity demand forecasting, applying short, medium and long-term predictions. These methods are classified in statistical models and machine learning models. This work focuses on short-term demand forecasting using machine learning.

- **The proposed approach for day ahead demand forecasting**

This section describes the proposed approach to solve the day-ahead electricity demand forecasting for an industrial park in Spain, a substation in Uruguay, and for the total demand of Uruguay applying the strategies described in Section 2.2. In addition, implementation details of the proposed models are presented.

- **General approach**

This subsection describes the data and the proposed methodology for electricity demand forecasting.

DATA DESCRIPTION, PREPARATION, AND METRICS

Data description. Data for the three studied scenarios are described next. *Industrial park in Burgos, Spain.* The first scenario reported in this article considers historical hourly energy consumption data from an industrial park in Spain. Data were collected between January 2014 and December 2017. The dataset consists of industrial energy consumption measurements. Each measurement is composed of the following fields:

- *Year* (integer), representing the year on which the measure was taken.
- *Month* (integer), indicating the month on which the measure was taken.
- *Day* (integer), indicating the day on which the measure was taken.
- *Hour* (integer), indicating the hour on which the measure was taken.
- *Dayofweek* (integer), indicating the day on which the measure was taken.
- *Workingday* (boolean), indicating whether the measure was taken in a working day or not.
-
- *Useful* (boolean), indicating whether the measure is valid.
- *Demand* (float), indicating the real power measured.

Substation SB1872 in Montevideo, Uruguay. The second scenario studied in this article considers historical hourly energy consumption data from a substation in *Tres Cruces* neighborhood in Montevideo, Uruguay. *Tres Cruces* is a neighborhood located near the centre of Montevideo that serves 390 citizens distributed in 117 homes with medium socio-economic level [29]. The studied dataset contains residential energy consumption measurements collected between January 2017 and December 2018. Each measurement is composed of the following fields:

- *Year* (integer), representing the year on which the measure was taken.
- *Month* (integer), indicating the month on which the measure was taken.
- *Day* (integer), indicating the day on which the measure was taken.
- *Hour* (integer), indicating the hour on which the measure was taken.

- *Dayofweek* (integer), indicating the day on which the measure was taken.
- *Workingday* (boolean), indicating whether the measure was taken in a working day or not.
- *Useful* (boolean), indicating whether the measure is valid.
- *Temperature* (float), indicating the temperature.
- *Humidity* (float), indicating humidity.
- *Wind speed* (float), indicating the average wind of a specific hour.
- *Demand* (float), indicating the real power measured.

Data preparation. For the three studied scenarios, data preparation consists in eliminating useless measurements and replacing outliers. A few useless measurements were found (less than 0.0001%) in each dataset.

TRAINING THE ONE HOUR AHEAD FORECASTING MODELS

Feature standardization was applied to the three scenarios data to avoid typical scale issues. For instance, if a feature in the dataset has a different order of magnitude compared to others then in algorithms where a metric is involved this big scaled feature becomes dominating and needs to be standardized [31].

Several visualization analysis were performed to gain an intuitive insight of the information contained in each feature. The most relevant fact confirmed in this preliminary analysis was the daily periodicity of the demand value in all scenarios. The diagram shown in Figure 1 reports the high correlation between actual demand and the demand of the same hour of two days before in the case of the industrial park scenario. Weather data was not considered in the industrial scenario because

Once all data were prepared for model training, a four-step procedure was applied for training and evaluation in all the scenarios studied. The four steps are:

- Training and test sets were generated in a 3:1 proportion. In the industrial park scenario, the training set considered data from 2014 to 2016 and the test set considered data from 2017. In the substation scenario, the training set considered data from January 2017 to June 2018 and the test set from July 2018 to December 2018. In the total demand scenario, the training set uses data from 2010 to 2016 and the test set considered data from 2017 to 2018.
- A simple base model was trained for benchmarking. Using the trained model, a recursive feature elimination process was performed. The ten most important features are preserved.
- Several models were trained and compared with the benchmark model.
- The best model according to *MAPE* metric was chosen.

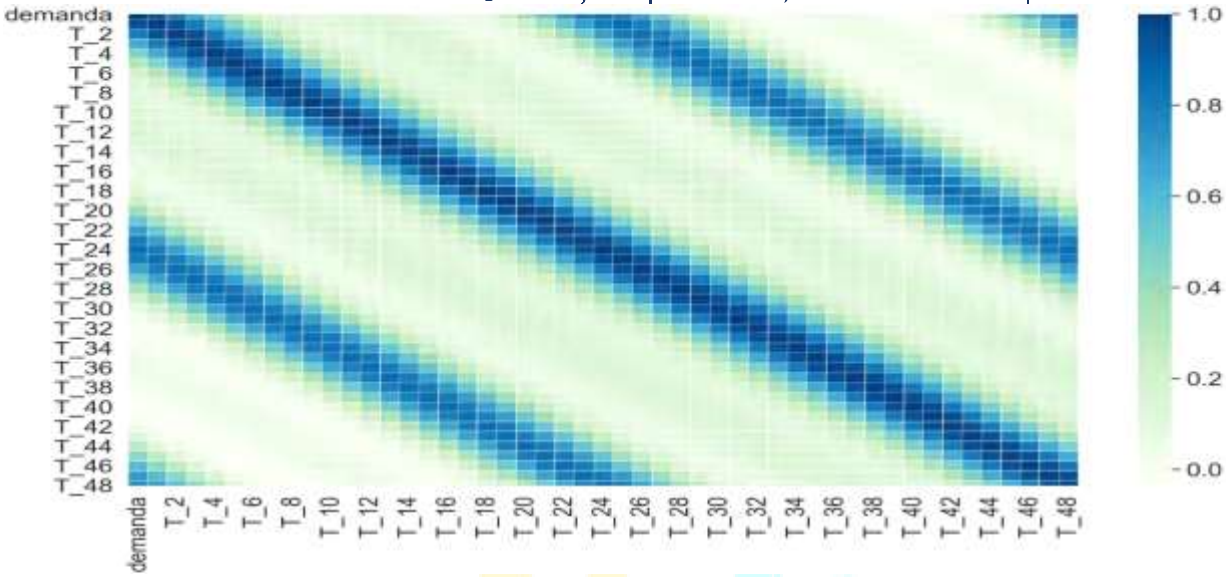


Figure 1 Correlation diagram between actual demand and 48 last demand measures

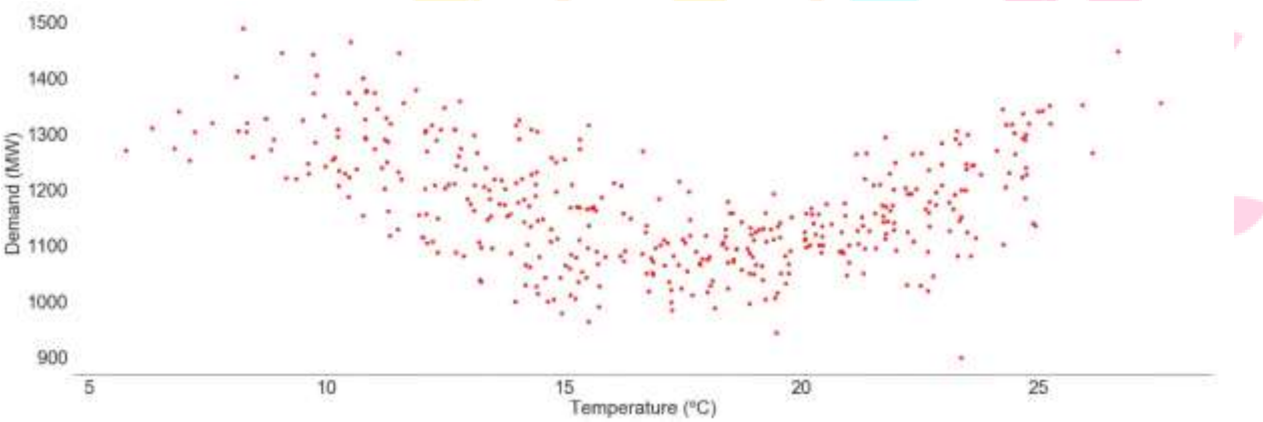


Figure 2 Relation between temperature and total demand of Uruguay

- An optimization of hyperparameters of the best model was performed using grid search techniques.

Finally, the best model found with the optimized hyperparameters was used as a reference to train the 24 hour forecasting model.

GENERATION OF THE 24 HOUR MODEL

The best model configured with the best hyperparameters obtained in the previous step, was used to generate twenty four models M_1, M_2, ..., M_{24} to forecast day ahead hours. The twenty four models were generated by applying the following procedure:

- Training and test sets were generated using the same procedure described in Section 3.1. In the industrial park scenario, the training set considered data from 2014 to 2016 and the test set considered data from 2017. In the substation scenario, the training set considered data from January 2017 to June 2018 and the test set from July 2018 to December 2018. In the total demand scenario, the training set uses data from 2010 to 2016 and the test set considered data from 2017 to 2018.

- Model M_i was trained using y_i as output, where y_i consists of the demand value corresponding to i hours ahead, and input X is enriched for models M_i, $i > 2$ with a new column consisting of the $i - 1$ prediction obtained by the trained model M_{i-1}

- Models M_i are assembled to get a complete model M to forecast the next 24 hours altogether.

- **Implementation**

This section describes the implementation of the approach described in Section 3.1.

COMPUTATIONAL PLATFORM AND SOFTWARE

The experimental evaluation was performed in an HP ProLiant DL380 G9 high end server with two Intel Xeon Gold 6138 processors (20 cores each) and 128 GB RAM, from the high performance computing infrastructure of National Supercomputing Center, Uruguay (Cluster-UY) [34].

IMPLEMENTATION OF ONE-HOUR MODEL

All one hour models described in this section use training and test sets and data preprocessing presented in Section 3.1.

Base model: Linear regression. A linear regression model was trained to be used as a baseline for the results comparison. A recursive feature selection strategy [35] was also applied on this model for each of the three

scenarios to determine the most important features. The rest of features were removed from the dataset.

Ten features were selected based on their relative importance in the industrial demand scenario:

- T_1, T_2, T_{24}, T_{25}: demand values lagged.
- *workingday*:flag indicating whether the day of measured value is a working day
- *month*: month on which the measure was taken.
- *hour*: hour of the day on which the measure was taken.
- *dayofweek*: day of the week on which the measure was taken.
- *day*: day of month on which the measure was taken.
- *year*: year on which the measure was taken.

Optimization of the best method. Parameter search techniques were applied for each scenario to optimize a model based on the best method obtained (M_{best}). The model M_{best} trained with default parameters was optimized using two standard tools available in scikit-learn:

- GridSearchCV: combines an estimator with a grid search preamble to tune hyper-parameters. The method picks the optimal parameter from the grid search and uses it with the estimator selected according to a predetermined metric.
- RandomizedSearchCV: sets up a grid of hyperparameter values and selects random combinations to train the model and score. After that, the method finds the best parameters setting according to a predetermined metric.

The best parameter set obtained for M_{best} are used in an optimal model M_{opt}. The main details of the implementation of the complete model are described in the next subsection.

- **Implementation of the complete model**

Model M_{opt} was optimized for predicting the next hour and used for predicting any of the following 24 hours to build the complete model. This decision was adopted assuming that the forecasting quality of the parameter setting obtained in the previous phase is independent of the hour used as output.

Experimental analysis

This section presents the results of the experimental analysis of the proposed computational intelligence methods for day ahead electricity demand forecasting in industrial and residential scenarios.

- **Recursive feature elimination**

A feature selection analysis was performed using the recursive feature elimination tool available in sklearn.

A model is specified and a number of features are selected, and the tool works by recursively removing features and building a new model (of the specified type) on those remaining features.

The accuracy of the new model is used to identify the features or combination of features that contribute the most to predicting the target attribute.

- **Experimental results on preliminary models**

Performance metrics defined in Section 3.1 were used to evaluate the implementation of the one hour models as Results reported in Table 1 for the industrial scenario indicate that three of the studied methods achieved the best results regarding the analyzed metrics. Focusing on *MAPE*, Extratreesregressor improved over MLP by 4.16%

- **Experimental results after parameter tuning**

Tables 4–6 report the results of the *ExtraTreesRegressor* model before and after parameter tuning for each scenario. The best results are highlighted (cells with green background).

Results computed by the tuned configuration of *ExtraTreesRegressor* considerably improved the baseline (non-tuned) version, regarding the three studied metrics.

The input for both grid search studied techniques in the three data scenarios was generated using the following values:

- *n_estimators*: [10, 50, 75, 100, 150];
- *max_features*: [auto, sqrt, log2];
- *max_depth*: [50, 100,150, 200, 250]

- **Experimental results of the complete model**

The forecast accuracy of the final model was validated by applying *MAPEtot* a metric that extends *MAPE*. Let

Table 1 Results for each regression method in the industrial scenario

Regression method	MAE	MAPE	RMSE	Score	Time (s)
LinearRegression	123.63	3.60	176.00	0.96	1.72
Ridge	125.63	3.60	176.00	0.97	0.09
KNeighbors	182.54	5.03	253.20	0.93	0.07
RandomForest	110.20	3.21	151.54	0.98	3.10
GradientBoosting	121.97	3.38	166.17	0.97	1.99
MLP	112.08	3.13	154.23	0.97	6.21
ExtraTrees	107.44	3.00	148.61	0.99	1.21

Table 2 Results for the studied regression method in the substation scenario

Regression method	MAE	MAPE	RMSE	Score	Time (s)
LinearRegression	472.33	14.60	511.00	0.86	1.71
Ridge	473.11	14.91	521.11	0.85	0.11
KNeighbors	533.41	17.31	593.10	0.79	0.08
RandomForest	453.50	12.90	583.15	0.91	5.01
GradientBoosting	466.73	13.83	599.72	0.88	2.19
MLP	448.18	12.74	576.35	0.93	6.91
ExtraTrees	441.14	11.24	558.33	0.95	1.43

$MAPEh$ be the *MAPE* value for a predicted horizon *h*, the extension of *MAPE* to the complete testing set is defined by Equation 10.

$$\sum k \quad MAPE_h$$

The expected behaviour is that the models trained for highly correlated hours in the future respect to the current hour, perform better than less correlated.

Table 3 Comparative results of ExtraTrees before and after parameter tuning

Regression method	MAE	MAPE	RMSE	Score	Time(s)
ExtraTrees before tuning	208.14	5.99	265.11	0.96	1.26
ExtraTrees after tuning	100.08	5.17	131.83	0.98	1.28

This fact is due to predictability, and it is enhanced when the correlation between input features and predicted values is higher.

According to Figure 1, highly correlated demand values correspond to the immediately preceding hours and from the same hours of the day before.

Analyzing the obtained results for the *MAPEtot* metric for each one of the 24 hourly models, the performance got worse from $i = 1$ to 17 and then improved from $i = 18$ to 24. These results show that highly correlated demand values performed better, as expected.

Finally, the complete model ET_{opt} was applied. A day-ahead hourly forecast demand curve was generated for each time window for the testing set and the *MAPEtot* value was calculated.

The final result for the complete model was *MAPEtot* = 2.55% in the industrial scenario, *MAPEtot* = 5.17% in the industrial scenario and *MAPEtot* = 9.09% in the substation scenario.

For the substation scenario, there are no known previous analysis in Uruguay to compare, but considering that the group of homes connected to the substation is small, an
The scenarios analyzed are representative of the studied industrial and residential demands. The results obtained with the model created were very good for all three cases.

- **Conclusions and future work**

This article presented an approach to address the problem of day ahead electricity demand forecasting.
Several machine learning models were presented and studied for next hour forecasting. Recursive feature selection was applied to select the most relevant features to train the studied models. After a comparative evaluation, the best model was optimized using random search and grid search techniques.

A hybrid strategy (combining direct and recursive approaches) was built based on the optimized model for single hour prediction. It was applied to build a complete day ahead electricity demand hourly forecasting model

tot = 9.09% is considered acceptable.
in three scenarios: an industrial demand forecasting scenario in Spain, a residential demand forecasting

For the total demand scenario, a relevant baseline for comparison is provided by the prediction models currently used by the National Administration of the Electric Market, Uruguay (ADME, adme.com.uy). According to public information reported in the ADME website, currently used prediction models have errors (*MAPEtot*) between 5.00% and 7.00%, with an average of *MAPEtot* = 5.52%. The model evaluated in this article reported an error of *MAPEtot* = 5.17%, which constitutes an excellent result, improving over baseline ADME methods by 6.34%. This is a very encouraging result for total demand prediction in Uruguay.

Table 4 *MAPEtot* score for each *ETopt,i* single hour model, industrial scenario

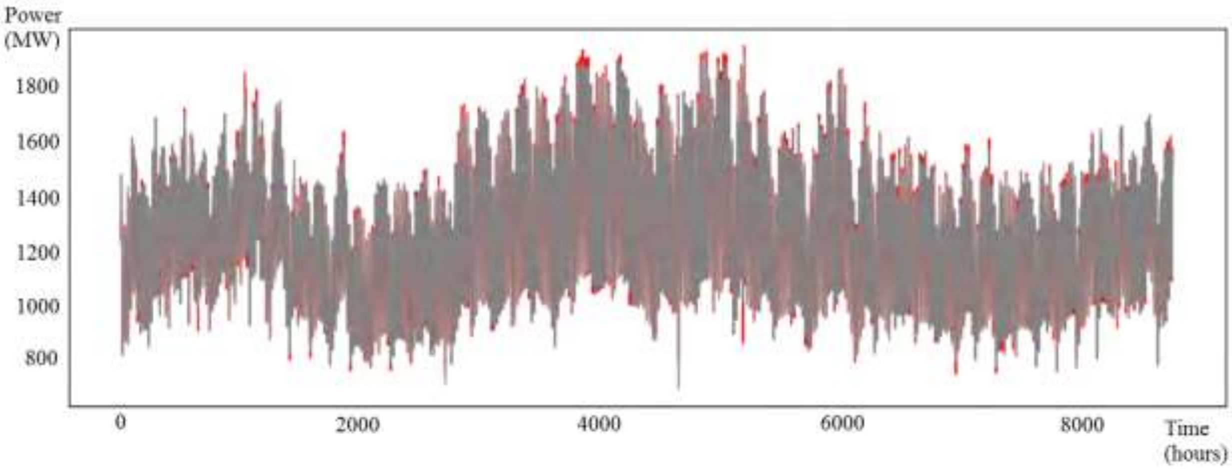
						hour						
	1	2	3	4	5	6	7	8	9	10	11	12
MAPE*tot*	1.79	1.84	1.90	1.97	2.09	2.19	2.39	2.52	2.68	2.75	2.80	2.86
						hour						
	13	14	15	16	17	18	19	20	21	22	23	24
MAPE*tot*	2.93	3.02	3.05	3.08	3.09	3.02	2.88	2.77	2.63	2.49	2.32	2.17

MAPE*tot*	5.94	6.13	6.18	6.25	6.26	6.13	5.84	5.61	5.33	5.05	4.70	4.40

For the substation scenario, the evaluation of the complete model reported a value of $MAPE_{tot} = 9.09\%$.

In this case, the proposed algorithm considered weather variables due to the high correlation detected between them and electricity demand. Results indicated that the complete model can predict the demand with an acceptable accuracy, in line with results from the literature, especially

Figure 1 Predicted demand and testing data curves of substation demand



considering that the variation of residential demand of a small group of houses is significantly higher than the variation of industrial demand.

Finally, the application of the complete model to the total demand scenario in Uruguay reported a value of $MAPE_{tot}$ = 5.17%. In this scenario, only one weather variable (temperature) was considered (humidity and wind speed were excluded due to low relative importance).

Another line of future work consists in enriching the studied models to generate mid-term and long-term synthetic demand scenarios that preserve the statistical structure of historical data. These kinds of models are very relevant to be included in planning and operation models based on new computational intelligence techniques such as reinforcement learning or approximate dynamic programming. Furthermore, prediction results can be applied in practice for household energy planning by using intelligent recommendation systems [37].

Declaration of competing interest

We declare that we have no significant competing interests including financial or non-financial, professional, or personal interests interfering with the full and objective presentation of the work described in this manuscript.

REFERENCES

- A. Diniz and *et al*, “Short/mid-term hydrothermal dispatch and spot pricing for large-scale systems-the case of Brazil,” in *2018 Power Systems Computation Conference (PSCC)*, Dublin, Ireland, 2018, pp. 1–7.
- L. Resende, M. Soares, and P. Ferreira, “Electric power load in Brazil:View on the long-term forecasting models,” *Production*, vol. 28, October 8 2018. [Online]. Available: https://doi.org/10.1590/ 0103- HYPERLINK "https://doi.org/10.1590/0103-6513.170081"6513.170081

- D. Lazos, A. Sproul, and M. Kay, "Optimisation of energy management in commercial buildings with weather forecasting inputs: A review," *Renewable and Sustainable Energy Reviews*, vol. 39, November 2014. [Online]. Available: https://doi.org/10.1016/ j.rser.2014.07.053
- S. Fan, L. Chen, and W. Lee, "Machine learning based switching model for electricity load forecasting," *Energy Conversion and Management*, vol. 49, no. 6, June 2008. [Online]. Available: https://doi.org/10.1016/j.enconman.2015.07.041
- A. Lahouar and J. Slama, "Day-ahead load forecast using random forest and expert input selection," *Energy Conversion and Management*, vol. 103, October 2015. [Online]. Available: https://doi.org/10.1016/j.enconman.2015.07.041
- S. S. Ahmed, R. Thiruvengadam, S. Karrthikeyaa, and V. Vijayaraghavan, "A two-fold machine learning approach for efficient day-ahead load prediction at hourly granularity for NYC," in *FICC 2019: Advances in Information and Communication*, 2019, pp. 84–97.
- R. Porteiro, S. Nesmachnow, and L. Hernández, "Short term load forecasting of industrial electricity using machine learning," in *Ibero-American Congress on Information Management and Big Data (ICSC-CITIES 2019)*, 2019, pp. 146–161.
- J. Chavat, J. Graneri, and S. Nesmachnow, "Household energy disaggregation based on pattern consumption similarities," in *Ibero-American Congress on Information Management and Big Data (ICSC-CITIES 2019)*, 2019, pp. 54–69.