Universidad Central - Universidad Jorge Tadeo Lozano

Neural network model for building electric energy consumption forecasting in densely populated tropical areas

MSc Thesis

presented to obtain the degree of

Master of Sciences in Modeling and Simulation

presented by

JOSE ANTONIO BELLO ACOSTA

Supervisors: Jimeno Fonseca, PhD Hugo Franco, PhD

Bogotá D.C. 2020

A mi madre, quién me ha apoyado en todo el camino que he recorrido y me ha impulsado a llegar hasta donde estoy y de quién aprendí los mejores valores humanos. A mi padre, de quién aprendí la fortaleza y el seguir adelante a pesar de las adversidades y a quien la vida este año le dio una segunda oportunidad para continuar compartiendo con nosotros. A mi familia y amigos, con quienes compartí momentos enriquecedores y que van a perdurar por siempre.

Acknowledgments

I would like to thank my tutors for their patience and trust along this learning path, the High Performance Computing LAB HPCLAB at Universidad Central for allowing me their infrastructure in the development of my project, the Future Cities Laboratory FCL-ETH in Singapore for welcoming me and introducing me to a whole new world of culture and science, the City Energy Analyst CEA research team for letting me use their data. The Universidad Central for the support offered to me in multiple steps of my career, the Universidad Jorge Tadeo Lozano for their time, patience and comprehension. The Master's professors for their dedication and effort poured into our formation.

Contents

1 Introduction **1**

2 Literature review **3**
2.1 Time series analysis and forecasting . . . 3
2.1.1 ARIMA family models . . . 4
2.1.2 Deep learning techniques . . . 6
2.2 Building energy consumption modeling . . . 9
2.2.1 The City Energy Analyst . . . 10
2.3 Related works . . . 11

3 Objectives **13**
3.1 Objectives . . . 13
3.1.1 General objective . . . 13
3.1.2 Specific objectives . . . 13

4 Methods **15**
4.1 Data description and pre-processing . . . 15
4.1.1 Data description . . . 15
4.1.2 Data pre-processing . . . 16
4.2 Implemented models . . . 22
4.2.1 Univariate linear model -SARIMA- . . . 22
4.2.2 Univariate non-linear model -NAR- . . . 23
4.2.3 Multivariate model with DNN -NARX- . . . 24
4.2.4 RNN model implementation . . . 25
4.2.5 Models evaluation and comparison . . . 25

5 Results **27**
5.1 Dataset . . . 27
5.2 Univariate linear model -SARIMA- . . . 27
5.3 Univariate non-linear model -NAR- . . . 31

5.4 Multivariate non-linear model -NARX- 36
5.5 Recurrent neural network -RNN- 41

6 Discussion and conclusions **45**

Chapter 1

Introduction

According to the United Nations [32], dense urban zones could offer important opportunities for economic development and support a better supply of public services. Public transportation, electricity, water and sanitation services tend to be cheaper and less environmentally damaging in these regions than in rural areas. Furthermore urban planning oriented to the construction of new cities (or densification of current ones) requires a robust scenario analysis. Such certainty levels should accurately depict the requirements of future city dwellers [84].

Electricity demand in urban regions has increased proportionally to its population growth [91]; nevertheless, this demand experiences a high short-term variation, as a consequence of the tendencies of industry, commerce and residential activity in each specific case [11, 19]. Moreover, factors such as weather, building geometry, time of day and even others like building occupancy or day of the week can affect the variation in patterns of accumulated energy consumption across the buildings of an urban region [7, 96, 13]. Based on these considerations, it is desirable to propose forecasting models able to predict the energy demand of one or more structures, intended to support public policies for generation and distribution of electric energy.

Both the future supply of electric energy and the real-time management of electricity generation and distribution in certain conditions, highly require the best forecasting tools to depict different scenarios and to support decision making and planning in the urban planning scope [30]. In order to identify the model that best fits the available data to provide reliable forecasts, it is necessary to compare different approaches and describe the advantages and disadvantages of each one and with the purpose of objectively choose the best option.

Stochastic models, s.a. the "ARIMA family", are a common choice in the specialized literature since they provide uni- and multivariate frameworks to perform prediction tasks. In the same fashion, recent works adopt Deep Learning based models (i.e. deep neural networks) s.a. recurrent neural networks (RNN), to address

the same kind of forecasting problems, also addressing the issue of collective energy demand forecasting using a single common model for the entire urban region under study.

This project approaches the comparison of energy demand forecasting models in tropical cities. The forecasting models under evaluation are based on stochastic methods (SARIMA) and neural networks (autoregressive and RNN models). The design, training and evaluation of the hybrid models use both metered (energy consumption) and simulated (building features) data from the city of Singapore, on a single-building basis for different building types.

Chapter 2

Literature review

The literature review is divided in three sections. The first one is related to the forecast tools for stochastic time series including ARIMA family models and machine learning based models. The second section is related to different approaches of buildings energy consumption modeling. The final section shows the most related works exposing relevant information for this study.

2.1 Time series analysis and forecasting

A wide variety of techniques has been used for forecasting time series in many fields of knowledge. Moscoso-López et al. [76] used RNN and support vector regression to predict the air pollution. Wang et al. [119] used RNN and ARIMA models to forecast hydrological time series. Cogollo et al. [23] used NN to forecast non-linear time series with moving average components. Zheng et al. and Nyaramneni et al. [129, 80] used some of those machine learning techniques for predicting the traffic congestion using deep learning models and stochastical approach. All of the above are a small sample of the possible applications of forecasting techniques in diverse areas of study.

Usually, forecasting models are trained and validated by splitting the dataset into a *training set* and a *testing set*. The training set allows the application of a "learning" process, changing the values of the model internal parameters. Once the model is trained so its output is consistent with an expected behaviour defined by the training samples, the testing stage starts and the model prediction is compared to the real values for each testing sample, allowing a quantitative evaluation of the model performance in training and testing phases.

Some of the metrics used to compare model performances are based on error, such as the root mean squared error, *RMSE* (2.1), the mean absolute error, *MAE*

(2.2), the mean percentage error, *MPE* (2.3), and mean absolute percentage error, *MAPE* (2.4). All of these metrics only tend to reduce the dissimilarity presented in the predicted and the real values.

$$RMSE = \sqrt{\text{mean}((y_t - \hat{y}_t)^2)} \tag{2.1}$$

$$MAE = \text{mean}(|y_t - \hat{y}_t|) \tag{2.2}$$

$$MPE = 100 * \text{mean}((y_t - \hat{y}_t)/y_t) \tag{2.3}$$

$$MAPE = 100 * \text{mean}(|(y_t - \hat{y}_t)/y_t|) \tag{2.4}$$

2.1.1 ARIMA family models

The basic ARIMA model proposed by Box and Jenkins [15] considers a linear combination of autoregressive (AR) and moving average (MA) terms to make a prediction of future values for a time series. The complete model also includes a differencing (I) process to remove tendencies and make the time series more similar to a stationary process. The aim of this model is to find the amount and values of the factors for each component AR (denoted with p), MA (denoted with q) and the number of differences to apply (denoted with d). The equations related to the ARIMA model are as follows [55]:

$$y_t = c + \phi_1 * y_{t-1} + \phi_2 * y_{t-2} ... + \phi_p * y_{t-p} + \epsilon_t \tag{2.5}$$

$$y_t = c + \epsilon_t + \theta_1 * \epsilon_{t-1} + \theta_2 * \epsilon_{t-2} + ... + \theta_q * \epsilon_{t-q} \tag{2.6}$$

$$y'_t = y_t - y_{t-1} \tag{2.7}$$

$$y'_t = c + \phi_1 * y_{t-1} + ... + \phi_p * y_{t-p} + \theta_1 * \epsilon_{t-1} + \theta_q * \epsilon_{t-q} + \epsilon_t \tag{2.8}$$

with y_t as the time series prediction at time t and ϵ as the forecast error. Equation (2.5) corresponds to a stationary AR model, (2.6) is a stationary MA model and (2.7) expresses the differentiation of the time series. The expression (2.8) represents the entire ARIMA model. Using the backshift operator notation in (2.9) and (2.10), the ARIMA model can be expressed by (2.11) as follows:

$$(B)y_t = y_{t-1} \tag{2.9}$$

$$y'_t = y_t - (B)y_t = (1 - B)y_t \tag{2.10}$$

$$(1 - \phi_1(B) - ... - \phi_p(B^p))(1 - B^d)y_t = c + (1 + \theta_1(B) + ... + \theta_q(B^q))\epsilon_t \tag{2.11}$$

The Seasonal ARIMA (SARIMA) model was proposed by Box and Jenkins as a stochastic approach for regression and forecasting of time series exhibiting seasonal periodicity. This model includes seasonal lag terms, denoted with P for the AR component, D for the differences and Q for the MA component. Thus the factors for

seasonal and non seasonal phenomena can be expressed as follows in the equation (2.12) that shows the SARIMA equation:

$$\begin{aligned}(1 - ... - \phi_p(B^p))(1 - ... - \Phi_P(B^{PS}))(1 - B^d)(1 - B^S)y_t = \\ (1 + ... + \theta_q(B^q))(1 + ... + \Theta_Q(B^{QS}))\epsilon_t\end{aligned} \tag{2.12}$$

SARIMA models include a representation of the time series seasonal behavior by adding an extra linear combination of coefficients related to the seasonal period (denoted with S) of the time series under consideration. The number of terms for each of the components -stationary and seasonal- is estimated by examining the autocorrelation (ACF) and partial autocorrelation (PACF) functions.

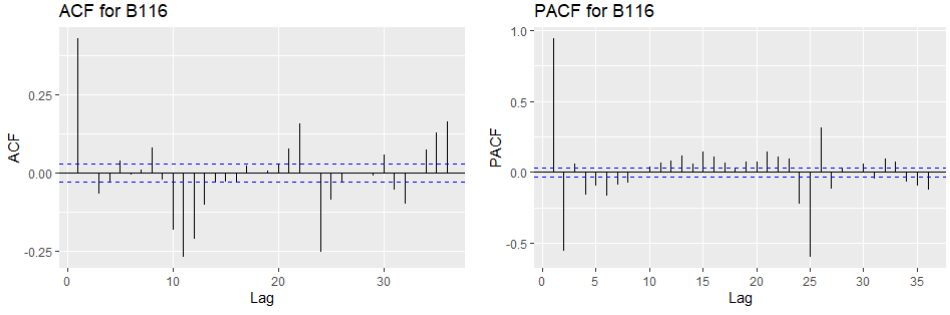


Figure 2.1: ACF and PACF for energy consumed in one building. Source: author

Thus, the model is defined by six parameters and the seasonal length value: $(p, d, q)(P, D, Q)_S$, where the lowercase variables correspond to the stationary component and the uppercase variables to the seasonal component. In this model the values for the parameters can be calculated following maximum likelihood. Using the Newton's method for maximizing the log-likelihood solving the equation (2.13):

$$\delta_1(\psi) = \frac{\delta \log L(Y_n|\psi)}{\delta\psi} = 0 \tag{2.13}$$

Considering the likelihood as the value that holds the largest probability of occurrence for a set of given parameters, the equation that rules that concept is (2.14):

$$L(Y_n) = p[y(t_1)]p[y(t_2)|y(t_1)]...p[y(t_n)|y(t_1), ..., y(t_{n-1})] \tag{2.14}$$

The information criteria is another common tool used to determine the simplest forecasting model i.e., the model with the lowest amount of terms but the lowest prediction error. Akaike's Information Criterion (AIC) and its corrected version

(AICc), suggested for ARIMA models [55], can be written as follows in equations (2.15) and (2.16):

$$AIC = -2log(L) + 2(p + q + k + 1) \tag{2.15}$$

$$AICc = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2} \tag{2.16}$$

With T as the size of the sample used for estimating the model. Another criteria can be used for choosing the simplest model: the Bayesian information criterion (BIC), also based on the AIC and its equation (2.17) is as follows:

$$BIC = AIC + [\log(T) - 2](p + q + k + 1) \tag{2.17}$$

Statistical models have been used for predicting certain kind of phenomena. Zhang [127] used ARIMA models in Canadian lynx data, sunspot data and British pound to US dollar exchange rate data, comparing the performance with a combination with NN. Smith et al. [114] implemented SARIMA models to forecast the traffic flow in London. Kavasseri et al. [63] executed an adaptation of the ARIMA model to forecast the wind speed. Contreras et al. [24] applied ARIMA models to predict the electricity prices in Spain and California. And Serrano-Guerrero et al. [109] used ARIMA model to predict electric demand in buildings in Ecuador and Spain.

2.1.2 Deep learning techniques

Two of the main goals of machine learning algorithms are the classification and/or the regression of real data [43]. There are several tools for each task: nearest neighbors, logistic regression, support vector machines (SVM), decision trees and NN are used for classification. Linear regression, non-linear regression, random forest (RF) and also NN are models used for prediction and allow the forecasting of univariate and/or multivariate time series. However, NN have shown good forecast performances for linear and non-linear time series in different contexts [34, 76, 119].

The idea of a NN emerged from a model of the human neuron, the "perceptron": proposed by Rosenblatt [98]. This model uses a simple equation (2.18), where z is the result of a linear combination of inputs X multiplied by a set of weights W and b a bias added. The combination of a set of perceptrons (also called neurons, nodes or cells) to work in parallel is called a "layer". The result of a single neuron or a whole layer feeds the following layer or node using a non-linear activation function such as the rectified linear unit, ReLU (2.19), the sigmoid function (σ) (2.20), the hyperbolic tangent (tanh) (2.21), among others, obtaining the output (2.22) of the NN.

$$z = (WX + b) \tag{2.18}$$

$$ReLU(x) = \max(x, 0) \tag{2.19}$$

$$\text{sigmoid}(x) = \frac{1}{1 + \exp(-x)} \tag{2.20}$$

$$\tanh(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)} \tag{2.21}$$

$$\text{output} = \text{activation}(W^{(h)} z + b^{(h)}) \tag{2.22}$$

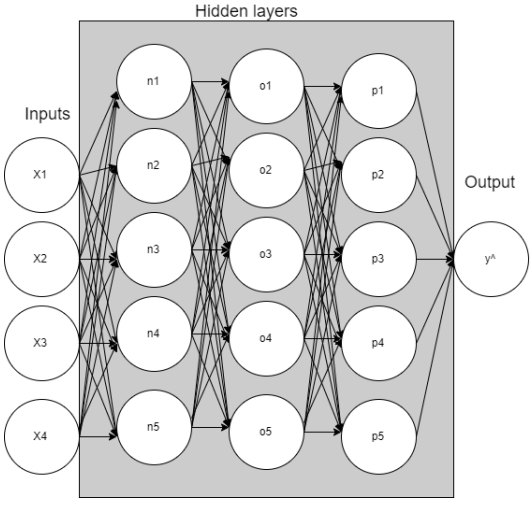


Figure 2.2: Multilayer perceptron structure. A NN with 3 hidden layers, 4 inputs and 1 output. Source: own ellaboration.

As shown in figure 2.2 the output can feed a next layer. Applying this process subsequently to the further layers until the last one is reached, the final output of the NN is obtained. This is the structure of a general deep NN also called a *multilayer perceptron* (MLP). Then the final output is compared with the real values using the loss function defined (MSE, MAE, etc.) finishing the forward propagation process. After that, the back propagation process updates the weights of each layer. The updating starts from the last output of the NN, using the gradient descent (2.23) to optimize the loss function or some others adaptations of the optimization algorithms like the Adam [66] or Adagrad [31] among others.

$$w_{i,j} = w_{i,j} - \alpha \frac{\delta e}{\delta w_{i,j}} \tag{2.23}$$

Some of the descriptive characteristics of the NN models, also called hyperparameters, for adjusting a proper model were mentioned above (optimizer, activation

function, number of layers, number of nodes), however there are some others features that can be tuned to help the model to obtain better performance or shorter training times. That is the case of normalization, which helps the layer weights to remain small and avoids the gradient exploding the weight values, also helping to speed the training stage. Dropout is other tool for generalizing the model in training stage. It consists on randomly eliminating a certain percentage of nodes in order to make the model choose better values for the weights, helping avoid over fitting [126].

One of the variations of the NN that has been used in time series forecasting is the RNN. This kind of model uses the previous values of the time series and recall the effect of past data on the prediction. Two of the RNN adaptations are the gated recurrent unit, GRU [20], and the long-short term memory, LSTM [51]. The first one as a simpler version of the last one. The structure of a single LSTM cell is shown in figure 2.3 and the equations related to the input (2.24), forget (2.25) and output (2.26) gates, candidate memory (2.27) and memory cell (2.28) are as follows:

$$I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \tag{2.24}$$

$$F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f) \tag{2.25}$$

$$O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \tag{2.26}$$

$$\tilde{C}_t = \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \tag{2.27}$$

$$C_t = F_t C_{t-1} + I_t \tilde{C}_t \tag{2.28}$$

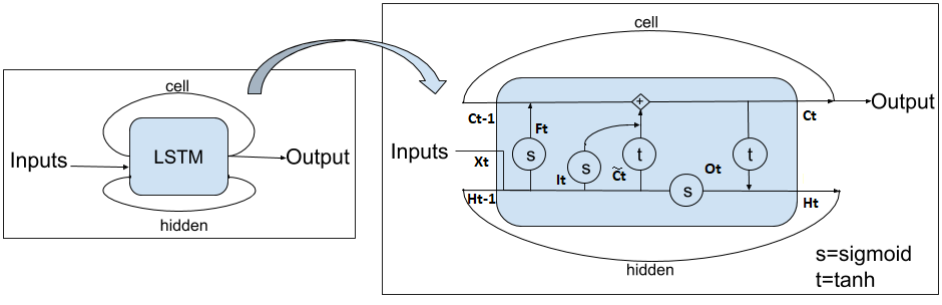


Figure 2.3: LSTM cell inner gates and operations. Source: Zhang (2020), adapted with permission.

Just like other forecasting tools, NN had been used in a wide variety of fields. Elsheikh et al. [34] used LSTM configuration for forecasting the freshwater yield from a solar still. Moscoso-Lopez et al. [76] and Hepziba et al. [49] used NN, SVM and LSTM to predict the hourly air quality index. Zheng et al. [129] used DNN to forecast traffic flow in large and medium cities. Kumar et al. [67] used RNN to predict Uber demand using daily data and weather information.

2.2 Building energy consumption modeling

Urban planning in highly populated areas is one of the main issues for decision makers, and it gets even more relevance in places with constrained land availability. In such context, the resource demand for different goods and services should be accurately estimated. Indeed, electric generation and distribution is critical for decision making at urban scale, since a reliable electric power supply must be guaranteed to keep the city economic activity running without discontinuities [36]. This service is in the heart of every sector of modern economies [56].

Besides the demand boost related to end-user applications, resulting from the massification of specific services (e.g., the Internet) and electronic devices like cell phones, laptops, tablets, etc. [26], the economic growth related to industrial, commercial and public sector activities creates a new challenge: satisfying the increasing demand while reducing the energy production costs and environmental impacts.

The electric energy demand exhibits different temporal patterns according to the specific usage associated for each activity and the place where it is carried out. Usually, morning hours present higher consumption levels than night hours, yet residential buildings experience demand peaks before and after the workday, while commercial, industrial and public sector structures are the opposite. Furthermore, the usage, architectural and geographical features of the buildings are tightly related to their electric energy requirements [79, 96, 120]. Energy consumption in buildings has gained great importance in the last decades due to its considerable share in global energy demand [88, 72, 69].

These consumption patterns are a key research issue in the related literature. Mohsenian-Rad et al. [74] considered a time varying energy pricing for encouraging people to reduce the energy demand in specific hours of the day usually associated to the peak demand hours. Another related approach was proposed by Paatero [85] who concluded that the energy demand varies according to human behavior and appliance usage. Similar conclusion was made by Richardson [96] who proposed a model for determining electricity use patterns, all related to the activity profiles and active occupancy of people in their homes. That proposition stands in a close line to Shen [113] who considered the intervention on the human behavior for reducing the energy need in buildings. Even the detection of anomalies in the consumption patterns is studied [108] in order to guarantee that the time series can be used for forecasting.

Several methods have been explored to increase the energy efficiency in buildings. One particular strategy to reduce the required electricity was addressed by Akbari et al. [5] who considered the use of vegetation for shading properties, evaluating the impact of urban trees in the temperature perceived and the energy demanded for heating, ventilation and air conditioning (HVAC) technologies. A related approach was presented by Sailor [100] who concluded that green roofs can reduce the energy demanded in some buildings in the United States, according to simulations performed

to reproduce the behaviour of office buildings in Chicago and Houston.

Other methods study building HVAC technologies, along with other systems, e.g., lighting. Zhang [128] recommended the use of heating and cooling storage (LHTES) to reduce the energy used in buildings. Chua et al. [21] found that almost 50% of the energy consumed in structures located in tropical areas was related to the HVAC appliances. More recently, Azaza et al. [9] have stated the use of a combination of low cost photovoltaic and heat-pump for reducing the electricity demanded in a single house. Wagiman et al. [118] describes how to reduce energy demanded in buildings with innovations in lighting control systems.Under this light, the analysis of electric energy consumption has been also focused on people's behavior inside buildings, as shown by Schoetter [104], Wei [122] and Fonseca [37].

Reinhart et al. [95] considered two main techniques for buildings energy consumption forecasting: data driven and physics based models. In that sense, Bourdeau et al. [13] have shown a detailed explanation of the characteristics of the methods used in energy consumption forecasting. Including white, gray and black box models and some of them for buildings classification.

Additionally, several works in the literature have addressed the problem of electric energy consumption forecasting in buildings at the urban scale from different perspectives, including building performance simulation [89, 47] and thermodynamics based models [73, 121, 40, 124]. Some approaches apply modeling and simulation techniques to reproduce the physical interactions between building structure, technology, users and the environment [102]. Also classifying the consumption in each kind of buildings, an unsupervised clustering model was tested by Fonseca et al. [38] suggested that a better prediction can be obtained thanks to a proper clustering in buildings groups.

Physics based modeling uses the thermodynamics laws and the effect of radiation and conduction on buildings floor, roof and walls and the variation in inner temperature [28]. Some others used surrounding characteristics of buildings such as shading and local wind [90]. Partial differential equations solved with finite-difference methods were used to model the heat-transfer inside buildings [17]. Moreover a recent common approach imply the use of GIS data for analyzing and showing the energy demand information in buildings stocks [53, 106].

2.2.1 The City Energy Analyst

The City Energy Analyst (CEA) [40] is an open-source urban simulation engine created to assess multiple energy efficiency strategies at different urban scales, i.e., building, neighborhood or city districts. This versatile tool allows city planners to simulate the energy demand of multiple buildings and find the optimal energy conversion systems for a district.

The CEA integrates a GIS extension for visualizing the spatiotemporal data and the distribution of the optimized systems. In addition, it also includes a way to see

the time series in buildings or districts and comparative graphics for the scenarios performance. In order to achieve this the CEA uses physical thermodynamical models and stochastical approaches to calculate the energy demand in a building or in a city.

The CEA is aimed to use exogenous and endogenous variables of one particular building or multiple buildings configuration and predict the individual or global power demand in different scenarios. This tool also considers the whole electricity distribution system showing the effects of future implementations in the regions environment.

The data that can be obtained from the CEA is based on the combination of all of the above mentioned considerations and even including negative effects of small disturbances such as air infiltration [44] or microclimate effect on the energy demand [78]. As a result, the data generated by the CEA is aimed to improve a forecasting model thanks to its precision.

2.3 Related works

Many researchers have analyzed methods for building energy demand prediction. Serrrano-Guerrero et al. [109] compared stochastic models like Holt-Winter with the ARIMA model appropriating a better performance for the ARIMA model in the three metrics used (MAE, MAPE and RMSE) and concluding that the prediction of the ARIMA models could be improved incorporating clustering techniques, seasonality analysis and artificial intelligence methods. Kim et al. [65] used linear regression and NN to combine the exogenous features included: temperature, humidity, solar irradiance, cloud type, wind speed and building occupancy (the last one measured with thermal sensors). For model comparison, the normalized mean bias error and the coefficient of variation of the RMSE were calculated, obtaining better performance for linear regression on the non-working days data but better results during working days using NN. It was found that the occupancy strongly dominates the electricity consumption while temperature and humidity also affect the results, to a lesser extent.

Amber et al. [7] implemented multiple regression, NN, DNN, and SVM with exogenous variables on a daily basis: temperature, solar irradiance, humidity, wind speed and a binary value for weekday or weekend. According to the performance, in the metrics considered (RMSE, MAE, MAPE and normalized RMSE), better results were provided by NN model, considering that this study was run on a single building. To summarise, the results of DNN were not as promising as expected due to the limited amount of data used and a slightly higher prediction for non-working days was also encountered.In the same way, Ruiz et al. [99] implemented a non-linear combination of AR terms using a NN architecture and also combining that model with exogenous variables (e.g., temperature). MSE was used as the performance

metric for training and testing sets (divided in 70% and 30% respectively). For the whole 8 buildings configurations from 2 to 20 neurons were tested and concluded that NARX networks furnish a better performance thanks to the exogenous data that also helps to decrease the network complexity and the amount of data used for forecasting.

Liao et al. [68] compared linear regression, RF, SVM,NN and RNN using variables like air pressure, temperature, humidity, wind speed, rainfall, sunshine hours, month, building orientation, floor and area because of the use of rooms consumption instead of whole building approach. Several hyperparameters for the NN based models were tested, including variations in nodes (32 to 512), layers (1 to 5), epochs (30 to 50), activation functions (tanh, ReLU), optimizer (nadam, adam, rmsprop), batch length (32 to 128) and dropout (80% to 100%). The metrics used for models comparison were RMSE, MAE and the R^2 concluding that the best performance was achieved by the RF followed by the RNN and suggesting that a more robust model can be obtained with data from diverse buildings. Rahman et al. [93] used NN and RNN with variables like temperature, humidity, hour of the day, day of the month, day of the week, weekend alert and a fast Fourier transformation of the HVAC consumption profiles. Mixed NN architecture were carried out, combining LSTM cells and a couple of MLP at the end of the model and RMSE was used as the performance metric. Better results were obtained using the RNN cells than a 3 layers MLP in commercial buildings.

Chapter 3

Objectives

3.1 Objectives

3.1.1 General objective

To compare the performance of stochastic and deep learning models for electric energy consumption forecasting in tropical city buildings.

3.1.2 Specific objectives

- To select the relevant variables representing building electric energy demand for metered and simulated data from Singapore for the year 2016.
- To design a stochastic model based on time series analysis for energy consumption forecasting.
- To design neural network models based on Autoregressive and Recurrent Neural Networks for energy consumption forecasting.

Chapter 4

Methods

4.1 Data description and pre-processing

4.1.1 Data description

All data used in this work belongs to the CEA research team [22]. The metered data was acquired in the facilities of the Nanyang Technological University (NTU) of Singapore, with a total of 40 buildings, while the simulated data was generated for the same buildings by the CEA team. Consequently, there were four different sources of data:

- Electricity consumption: hourly measured in the 40 buildings in the NTU. The total of variables was 1 for this database. This is the variable to forecast. An example of its behavior is shown in figure 4.1.
- Weather data: obtained from a weather station at the NTU. The dataset contains features like solar radiation, dry bulb temperature, relative humidity, wind speed among others. The total of variables was 36 for this database.
- Geometric data: features like buildings height, length, width and footprint were included for the 40 buildings. Additionally a set of proportions of the uses encountered in each building were considered. Some buildings have mixed uses like schools, offices and laboratories while some others have adopted just one use (e.g., residential, restaurant, retail, etc.). The total of variables was 13 for this database.
- Simulated data: generated from the CEA simulation tool for the 40 buildings in the NTU. These features included people inside the building, inner relative humidity, energy consumed by buildings appliances, indoor temperature, among others. The total of variables was 120 for this database.

Table 4.1: Time features used.

Name	Description	Values
Holiday Alert	Varies for Saturdays and Sundays	Weekday=1; Saturday=2; Sunday=3
Weekdays	One value for each day of the week	Monday=1; Tuesday=2; Wednesday=3 ...
Hour of the day	Starting from 1:00 am	From 1 to 24

Time features were used as linking elements. Variables like hour of the day and day of the week helped to define the exact moment of the measured or simulated variables; in sense, a weekend mark for Saturdays and Sundays were also considered and included in the time database. A brief summary of the time variables included is shown in Table 4.1. Thus, all data in the study correspond to hourly acquisitions or simulations for a calendar year. All data begins in January 1st and ends in December 31st of 2015.

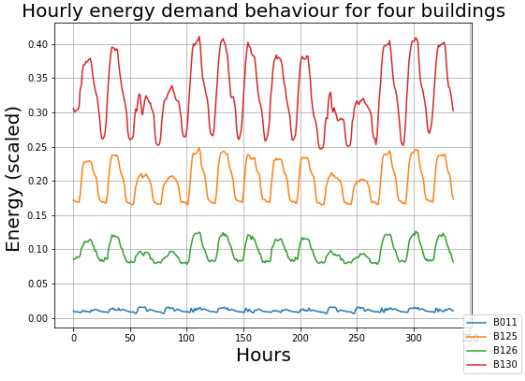


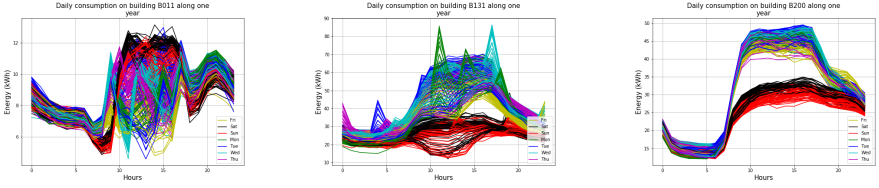
Figure 4.1: Scaled values for two-week energy consumption in a set of buildings.

4.1.2 Data pre-processing

As shown in figure 4.1 the demanded energy exhibits both daily and weekly seasonal behavior. The lower consumption values in most of the time series correspond to weekends, and the five peak values, to the maximum energy consumed during weekdays. Due to the fact that Singapore is located in a tropical region, there are no noticeable seasonality at other time scales. Following this behavior, the inclusion of a variable that represents weekends alert helps the model to obtain a better prediction.

To analyze the uni-variate linear and non-linear models it is not necessary to

perform any pre-processing stage with the energy consumed in the 40 buildings. Some examples of energy consumption patterns can be found in the seasonal plots of figure 4.2 that shows 24 hour of consumption along the year. Notorious differences can be found between the three buildings related; as for building B200, it could be appreciated the difference between weekdays high consumption and weekends lower values (in red and black). Similar patterns are found in building B131 that, in spite of taking into account variations among the year, they allow the conclusion of evident consumption patterns. However, there were also found buildings which do not determine a differential behavior during the days of the week; for instance building B011.



(a) Building B011 (b) Building B131 (c) Building B200

Figure 4.2: Seasonal plots for energy consumption in three buildings

On the other hand, in the multivariate NN based models there is necessary to reduce the amount of variables to be included for forecasting. As mentioned above, the amount of variables to be considered are more than 150, but not all of them are useful to get a proper prediction. In that order, a selection criteria was applied to each dataset to use just the most relevant variables. The higher correlation values for the input variables with the energy consumed determined the inclusion of the feature in the input list.

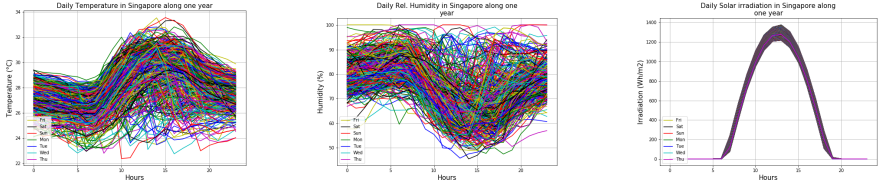
From the weather dataset 8 features were removed, since they had NULL values. With the remaining features, Pearson correlation tests were performed between them to avoid highly similar inputs that could introduce some noise to the prediction model. The threshold for high correlation between variables was defined in 0.8. The variables with higher correlation than the defined threshold are considered as highly crossed correlated. Most of the weather variables with high crossed correlation values were represented by a closely related physical phenomena (e.g., radiation and illuminance), exhibiting small variations between them (e.g., horizontal, direct or diffuse), so that only the most representative variable for each group was included.

Nevertheless, two particular variables: dry bulb temperature and relative humidity, that also had high crossed correlation value, were included in the list of inputs of the model, since they have different underlying physical phenomena. The weather features that accomplish both correlation conditions (high with energy consumed and low among them) were used to formulate the model and are listed in Table 4.2.

Table 4.2: Weather features used.

Name	Description	Unit	Values
Dry.Bulb. Temp	Temperature measured in dry conditions	° C	18 to 40
Relative Humidity	Relative Humidity	%	30 to 100
Extraterrestrial Horizontal Radiation	Solar radiation received on a surface normal to the rays of the sun at the top of the atmosphere	$\frac{Wh}{m^2}$	From 0

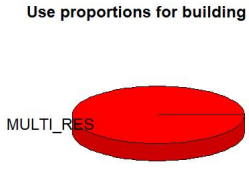
Figure 4.3 shows the daily behavior of the weather data included along the year. It can be seen that the temperature and the humidity present notorious variations over time, however a pattern can be identified for each feature. On the other hand the Extraterrestrial Horizontal Radiation concur with the tropical (and very close to Ecuator line) location of Singapore, showing almost no variation on sunrise and sundown hour over the year.



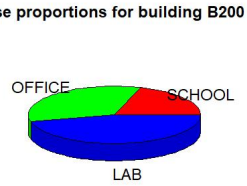
(a) Dry bulb temperature (b) Relative Humidity (c) Horizontal Radiation

Figure 4.3: Seasonal plots for weather variables included in the forecasting models

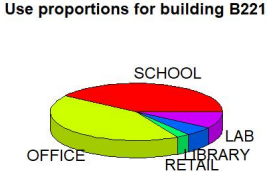
For the geometric dataset, the footprint of each building was used as a single representation of their dimensions (i.e., area, floors, width, etc.). To simplify the geometric data, those features were summarized in seven variables. These variables were created from the buildings usage proportions (see figure 4.4) multiplied by its footprint. In that order, these features represent the footprint proportion for every use in each building. This follows the idea that different uses have different electricity consumption behavior. The variables related to the geometric features used for the forecasting models can be found in Table 4.3.



(a) Building B097



(b) Building B200



(c) Building B221

Figure 4.4: Use proportions for three buildings.

Table 4.3: Geometric features used.

Name	**Description: Footprint**
fp.school	Proportion dedicated to school
fp.office	Proportion dedicated to office
fp.retail	Proportion dedicated to retail
fp.restaurant	Dedicated to restaurant
fp.multi_res	Dedicated to multiple residential
fp.library	Proportion dedicated to library
fp.lab	Proportion dedicated to lab

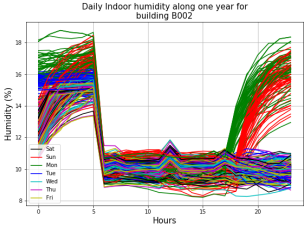
With the idea of guaranteeing that all buildings had the same input data, the simulated dataset was analyzed and found that a total of 28 variables were available for the whole 40 buildings set. Using the same procedure exposed previously, Pearson correlation tests were performed between these variables; some of them had higher crossed correlation values and, as in the weather dataset, only the most representative variables were included. A total of eight variables remained to be included in the forecasting models. These features are shown in table 4.4.

Table 4.4: Simulated features used. Variables names are the same as in CEA.

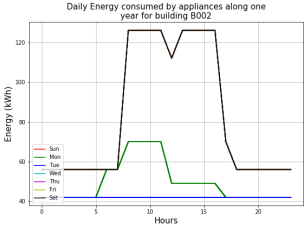
Name	**Description**	**Unit**	**Values**
People	People inside the building		From 0
x_int	Indoor relative humidity	%	From 0
Eal_kWh	Electricity consumption of appliances and lights	kWh	From 0
Qcs_sen_ sys_kWh	Total sensible cooling demand for all systems	kWh	From 0
Qcs_lat_ sys_kWh	Total latent cooling demand for all systems	kWh	From 0
Q_gain_sen_ vent_kWh	Sensible heat gain from ventilation and infiltration	kWh	From 0
I_sol_and _I_rad_kWh	Net radiative heat gain	kWh	From 0
T_int_C	Indoor temperature	°C	From 0

Figure 4.5 shows some of the simulated variables enclosed in the forecast models. It can be identified how some variables work with discrete values due to the own nature of the variable; whereas for example, in the case of indoor temperature, HVAC systems ensure steadiness. As a consequence, it is pretended to interpret outdoors and indoors temperature variations and relate them to crescent energetic demand.

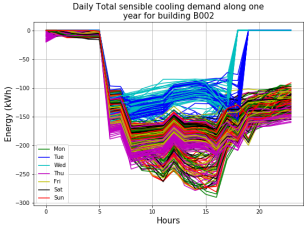
(a) Simulated people inside building



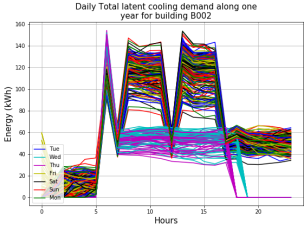
(b) Indoor relative humidity



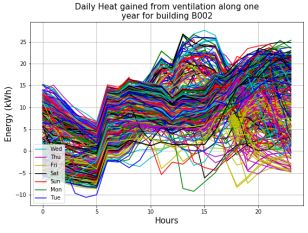
(c) Energy demand of appliances



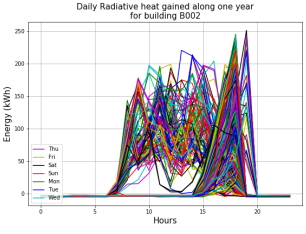
(d) Sensible cooling demand



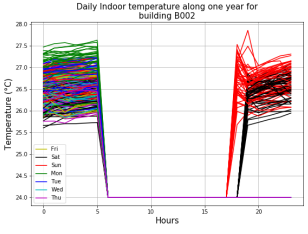
(e) Latent cooling demand



(f) Sensible heat from ventilation



(g) Radiative heat gained



(h) Indoor temperature

Figure 4.5: Simulated variables for building B002.

All models under evaluation used a scaling process. The Box-Cox transformation [14] was applied for the ARIMA model using the *BoxCox* function from the *forecast* package [54] in *R*. For all NN based models, the scaling process was implemented by using the *MinMaxScaler* function of the *Scikit-learn* package [86] in *python*.

An additional step is required for the RNN model implementation, these methods require time series as inputs. Using the main periodicity of the energy time series (24 hours), the input and output variables were sorted to obtain 24 hours of input data for each feature and also the following 24 hours of energy demand as output.

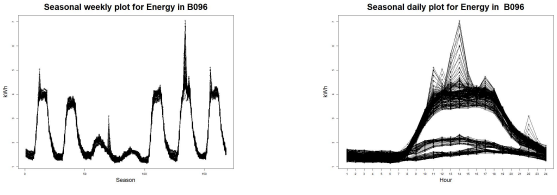
4.2 Implemented models

In this section, the models used to obtain the prediction of energy consumed are explained considering the previous pre-processing step. Two kind of models were implemented. The uni-variate models that used the energy demanded time series as input were divided in two: the linear, based on the ARIMA family model; and the non-linear, based on a NN configuration with AR terms as input. The multivariate models implemented were based on DNN models using the exogenous variables as input: the first one based on a MLP and the last one based on LSTM cells architecture.

4.2.1 Univariate linear model -SARIMA-

The Seasonal ARIMA (SARIMA) helps the time series forecasting when there are evident seasonal patterns. In this case it is important to identify the main periodicity of the time series. There are cases when the time series has more than one periodicity. Sometimes numerous seasons are closely related, under this light it could be said that one season effect is multiple of another effect e.g., a weekly seasonal behavior can be considered as seven daily seasonal effects. In this case, the main periodicity is the daily one.

To determine the main periodicities, the energy consumption time series were plotted, and two seasonal periods were found: daily and weekly (see figure 4.6). The daily periodicity was considered as the main seasonal effect for the time series and according to this, the SARIMA model used 24 hours as the S component.



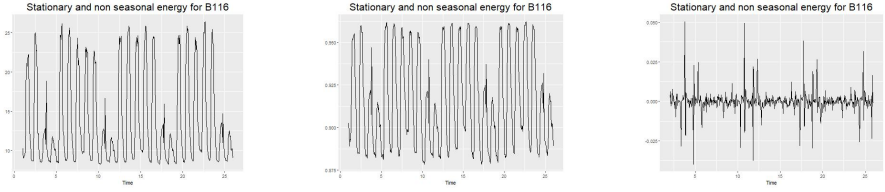
(a) Weekly behavior (b) Daily behavior

Figure 4.6: Energy consumption patterns for building B096

The data was transformed into time series objects with a frequency of 24 hours

by the *ts* function from *base* package in *R*. The Box-Cox transformation requires a lambda parameter. This value was auto-determined by the *BoxCox.lamba* function of the *forecast* package.

A discrete differencing step was used to transform the data into stationary process. To define the number of required differences, the *ndiffs* and *nsdiffs* functions (from the *forecast* package) were implemented for non-seasonal and seasonal differences respectively. Once the differences were applied, the time series became stationary and non-seasonal. Then, the ACF and the PACF were plotted (see figure 2.1) to determine, by visual inspection, the number of autoregressive and moving average terms for both non-seasonal, first few lags, and seasonal components, as the multiples of the season lags [15]. Figure 4.7 shows the transformation stages of the time series for a single building.



(a) Original data (b) Box-Cox Transformed (c) Differenced data

Figure 4.7: Energy consumption time series transformation for building B116.

Finally, with all the parameters needed to calculate the coefficients, the *arima* function from the *forecast* package in *R* was used to obtain the value of the model coefficients, using the same 70% of the data as training set. Once these values were obtained, the predicted values were calculated with the *forecast* function.

4.2.2 Univariate non-linear model -NAR-

For the non-linear univariate model implementation, a basic configuration of NN was deployed. Within the purpose of comparing used hyperparameters, grid research methods were applied. Thereby, in order to avoid underfitting or overfitting due to extremely low or high nodes usage, the number of inputs was used to define the number of nodes for each layer in the NN: in one scenario the number of nodes per layer was equal to the number of inputs. Two more variations were made: one includes 10 extra nodes. Another one excludes 10 nodes from the initial number of inputs.

The number of layers in the model was defined according to the architecture proposed by Ringwood et. al. [97] and Rahman et. a.l. [93]. Then, 5 layers were chosen as the base model, and NN architectures scenarios were defined by adding 3 layers to the base model, and by removing 3 layers from it.

All layers were normalized by using the *batch_normalization* function from the *TensorFlow* API [1] to increase the training speed. The ReLU activation function was selected for all hidden layers. Linear activation function was used for the output layer. Following the work of Goodfellow [43], a random mini-batch training was implemented to improve the training process convergence, with a batch size of 256 train samples. The number of epochs for all the scenarios was 300. The learning rate for the Adam optimizer was set to 0.003 and Mean Squared Error (MSE) was selected as loss function. Table 4.5 summarizes the architecture parameters under evaluation.

Table 4.5: Hyperparameters variation for univariate non-linear model

Layers	**Nodes**	**Output**
2	14	1
5	24	12
8	34	24

Hidden layers weights were initialized using a hundredth of random values with standard normal distribution, assuring proximity to 0. For each layer, a normalization stage was included, having as parameters the mean and standard deviation of mentioned layer values.

The data were split in 70% for the training dataset and 30% for the testing dataset and shuffled with the *train_test_split* function, to guarantee that the training set has examples from different hours of the year.

4.2.3 Multivariate model with DNN -NARX-

Previously selected variables were introduced to multivariate models and a couple scenarios were tested: whether to include simulated variables or not, since the last ones had not been contemplated in the literature review and are believed to grant a notorious advantage at the moment of achieving an appropriate forecast.

To include the physical characteristics of the buildings that are constant along the year (such as the geometric features), they were repeated for every hour of the year in each building dataset. Thus, each element from the training and testing sets must have the input parameters.

Another group of variations was also taken into consideration. With the purpose of establishing if AR terms contribute to enhance the models development, as shown by Ruiz et. al. [99], with the NARX model, the following variations were added: zero AR terms, twelve AR terms and twenty-four AR terms.

Once inputs, outputs and train and test sets were defined for each building, all these sets were stacked to obtain a general dataset for all the buildings in the study.

After dividing the data into training and testing sets, they were shuffled allowing training samples from all buildings homogeneously.

Table 4.6: NARX model architecture parameters in this study.

AR input	CEA vars	Layers	Nodes variation	Output
0	Yes	2	-10	1
12	No	5	0	12
24		8	+10	24

4.2.4 RNN model implementation

For the RNN model implementation, the *TensorFlow* API was used. The number of time steps ahead selected for the forecasting evaluation was set to 24 hours; subsequently, every feature was reshaped to include 24-time steps in the time dimension. With this choice, there was the chance of making the output overlap on the 24-time step; then, to make comparable the RNN forecasts with their NARX counterparts (where we generated the energy demand of every building in time-windows of 1 and 24-hours), the time steps of the output were overlapped 23 and 0 hours, generating 1 and 24 new hours ahead as outputs. The datasets were sorted properly for each building with the same proportions for training and testing sets (70% for train, 30% for test), and stacked to include all buildings. The LSTM cell was selected for the RNN architectures. The number of cells or nodes was set based also on the number of input features, with the same variations as in the previous model exploration (by adding or removing 10 nodes), as well as the number of layers (2, 5 and 8 layers). The Adam optimizer and the MSE loss function were used to train the model with random mini-batches. Table 4.7 shows the architecture parameters used to build RNN models in this work.

Table 4.7: RNN model architecture parameters in this study.

CEA vars	Layers	Nodes	Output
Yes	2	-10	1
No	5	0	12
	8	+10	24

4.2.5 Models evaluation and comparison

To evaluate the performance of all models in this study, four metrics were calculated for each forecasting window length: Root Mean Squared Error (RMSE) 2.1 and Mean Absolute Error (MAE) 2.2, Mean Percentage Error (MPE) 2.3, Mean Absolute Percentage Error (MAPE) 2.4 .

Chapter 5

Results

5.1 Dataset

Depending on the NN model to evaluate, a given number of entries should be removed from the dataset: in the 1 hour ahead forecasting, the inputs belong to time t and the energy demanded to time $t + 1$. In that way, one register is not considered into the dataset. This procedure applies also for AR terms, where the number added to the forecasting window, in hours, implies removing the same number of entries. This consideration also is employed for the RNN models. At the end, the number of training and testing samples (entries) varies in a range between 8712 and 8759 for each building (since the hours in a 365-day year are 8760).

5.2 Univariate linear model -SARIMA-

The SARIMA model was run for each building individually because of the high variability among the building time series values between buildings prevented the training convergence of a generalized model. The number of stationary and seasonal terms for a sample of buildings are shown in Table 5.1.

Table 5.1: SARIMA parameters for one subset of buildings.

Component	**Stationary**			**Seasonal**		
Building	p	d	q	P	D	Q
B084	2	1	3	2	1	4
B097	2	1	1	2	1	5
B098	2	1	1	2	1	5
B099	3	1	1	4	1	5
B100	3	1	1	3	1	2
B125	7	1	3	2	1	2
B126	8	1	2	2	1	5
B133	0	1	0	1	1	4
B198	8	1	7	2	1	2
B202	4	1	0	5	1	3

It can be seen that some buildings need five or more stationary AR or seasonal MA terms. The last remark provides the evidences to understand that to accomplish the forecast using this method, there is an important effect of what happened several time periods ago. Some groups of buildings need the same number of stationary and seasonal coefficients; however, the values of each term are not equal and the model performance exhibits a high variation between buildings with the same number of terms.

Thanks to the time versatility that this model offer, longer periods enabled not only forecast for a day, but to a week. The values of performance on selected metrics in a day (24 hours) forecast are shown in Table 5.2, for a week (168 hours) in Table 5.3. Performance values from the best forecast models show the same errors between day and week period, meaning that extending the time does not increase uncertainty and it remains low. It is noteworthy that only one out of the five buildings did not show a proper compliance during the week as it was during the day. Figure 5.1 acknowledges the models recognition of the buildings own consumption patterns. On the other side, the worst models proceed with results that show percentage errors that go over 50% or even some cases reach over 100%, stating how individual consumption adjustment for these buildings is not exceeded with an univariate lineal approximation. These behaviors are illustrated in Figure 5.2 where daily and weekly adjustments do not reflect the given base of the consumption pattern.

Table 5.2: SARIMA forecast for one day ahead.

Building	Time window	RMSE	MAE	MPE	MAPE
B125	24 h	2.9E-4	2.3E-4	0.02	0.02
B126	24 h	5.5E-4	3.5E-4	-0.01	0.04
B133	24 h	5.3E-4	3.4E-4	0.02	0.03
B198	24 h	4.7E-4	3.9E-4	0.03	0.04
B202	24 h	1.1E-4	8.3E-5	-0.00	0.01

(a) Best forecasting results.

Building	Time window	RMSE	MAE	MPE	MAPE
B084	24 h	1.1E-1	9.3E-2	-248.30	315.09
B097	24 h	1.3E-1	8.8E-2	12.71	32.98
B098	24 h	1.1E-1	7.3E-2	-57.41	61.11
B099	24 h	2.5E-1	1.9E-1	-69.93	73.24
B100	24 h	3.8E-1	2.9E-1	-105.64	113.02

(b) Worst forecasting results.

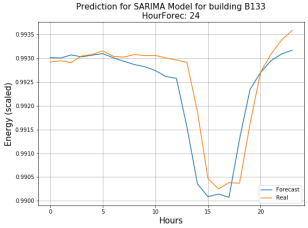
Table 5.3: SARIMA forecast for one week ahead.

Building	Time window	RMSE	MAE	MPE	MAPE
B125	168 h	4.7E-4	4.1E-4	0.032	0.041
B126	168 h	6.0E-4	4.5E-4	0.002	0.045
B133	168 h	5.5E-4	4.2E-4	0.025	0.042
B201	168 h	5.7E-4	5.1E-4	0.05	0.051
B202	168 h	1.9E-4	1.4E-4	0.001	0.014

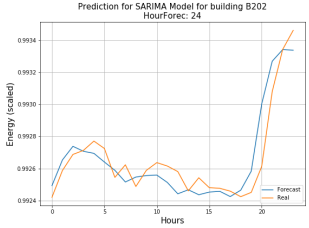
(a) Best forecasting results.

Building	Time window	RMSE	MAE	MPE	MAPE
B084	168 h	1.2E-1	9.2E-2	-27.38	111.02
B097	168 h	2.6E-1	2.0E-1	37.37	44.54
B099	168 h	2.6E-1	2.0E-1	-31.30	56.64
B100	168 h	3.7E-1	2.7E-1	-52.56	73.96
B101	168 h	3.0E-1	2.4E-1	53.37	56.02

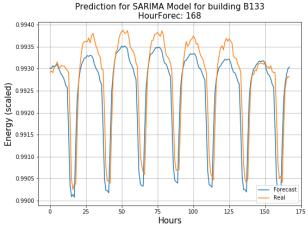
(b) Worst forecasting results.



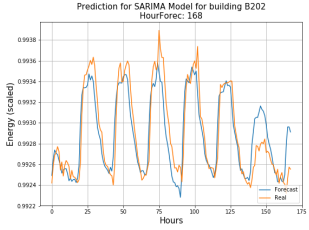
(a) One day ahead for building B133.



(b) One day ahead for building B202.

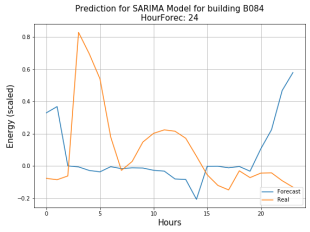


(c) One week ahead for building B133.

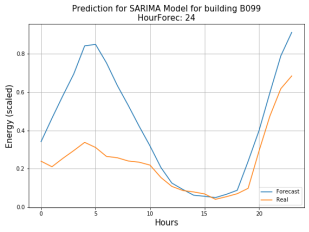


(d) One week ahead for building B202.

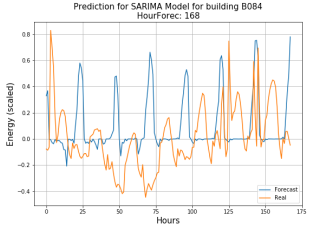
Figure 5.1: Best performance in SARIMA forecast.



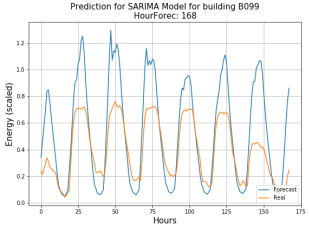
(a) One day ahead for building B084.



(b) One day ahead for building B099.



(c) One week ahead for building B084.

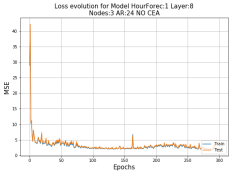


(d) One week ahead for building B099.

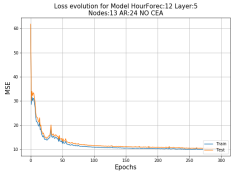
Figure 5.2: Worst performance in SARIMA forecast.

5.3 Univariate non-linear model -NAR-

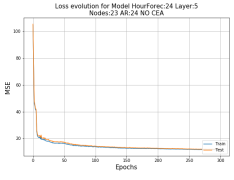
A MLP based model was run within all buildings as entry data. The 27 different architecture settings (varying nodes, layers and outputs) were tested. A model was sought aiming to allow the generalization of forecasts for multiple buildings using only energy consumption time series as entry. The evolution of loss function for some models during the training epochs defined are found in Figure 5.4 where there is expressed the way those models did not have a smooth behavior in the MSE value descent, and also the Figure 5.3 show models with a smooth and constant one.



(a) Forecast: 1 hour

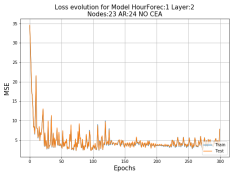


(b) Forecast: 12 hours

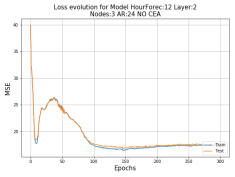


(c) Forecast: 24 hours

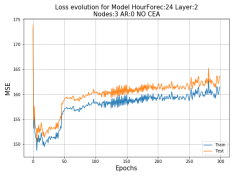
Figure 5.3: Best models loss evolution with NARX.



(a) Forecast: 1 hour



(b) Forecast: 12 hours



(c) Forecast: 24 hours

Figure 5.4: Worst models loss evolution with NARX.

Forecast results for 1, 12 and 24 hours are found in Tables 5.4, 5.5 and 5.6. The average error along with the standard deviation for the training and testing sets were calculated for the all models under evaluation. Moreover, the performance for the 2 layers setting presented a high variability due to architecture changes (added nodes), showing one of the best for no additional nodes but the least for 10 additional. Alongside for the 12 hours, the best model appears to be 5 layers with no additional nodes; and the worst is seen over 2 layers with 10 nodes less. Furthermore, error values tend to increase due to the difficulties forecasting more hours. Lastly, on the 24 hours side, the previous assertion about forecasting difficulty is confirmed when error values increase more with the time-lapse. For this case, no 2-layer model achieved the top performances; signifying the necessity to deepen NN in order to obtain a better result. The best model was 5 layers and 10 extra nodes; the worst, 2 layers with no additional nodes.

Table 5.4: Forecasting performance for 1 hour ahead in NAR model.

Lay	Nod	AR		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	0	24	mean	2.6E-3	2.6E-3	-16.07	17.15	2.4E-3	2.4E-3	-16.27	17.44
			s.d.	4.6E-3	4.6E-3	27.61	26.95	3.1E-3	3.1E-3	30.20	29.54
8	-10	24	mean	2.2E-3	2.2E-3	6.42	8.80	2.3E-3	2.3E-3	6.83	9.12
			s.d.	3.8E-3	3.8E-3	11.95	10.33	3.7E-3	3.7E-3	11.95	10.31
8	0	24	mean	2.2E-3	2.2E-3	0.01	6.86	2.2E-3	2.2E-3	-0.09	6.60
			s.d.	3.6E-3	3.6E-3	10.95	8.54	3.6E-3	3.6E-3	10.37	7.99

(a) Best forecasting performance.

Lay	Nod	AR		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	-10	24	mean	3.7E-3	3.7E-3	22.43	23.68	3.5E-3	3.5E-3	20.68	21.98
			s.d.	4.8E-3	4.8E-3	33.27	32.40	4.6E-3	4.6E-3	31.55	30.66
2	10	24	mean	6.8E-3	6.8E-3	-114.9	115.25	6.7E-3	6.7E-3	-115.1	115.3
			s.d.	6.2E-3	6.2E-3	212.66	212.46	5.5E-3	5.5E-3	205.98	205.85
5	10	24	mean	3.2E-3	3.2E-3	-44.93	46.19	3.2E-3	3.2E-3	-43.77	45.40
			s.d.	4.1E-3	4.1E-3	95.26	94.66	4.2E-3	4.2E-3	90.20	89.39

(b) Worst forecasting performance.

Table 5.5: Forecasting performance for 12 hours ahead in NAR model.

Lay	Nod	AR		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	10	24	mean	6.7E-3	5.7E-3	-47.95	52.19	6.2E-3	5.3E-3	-37.45	42.42
			s.d.	9.5E-3	8.2E-3	110.28	108.74	8.2E-3	7.1E-3	88.62	86.91
5	-10	24	mean	6.4E-3	5.4E-3	1.46	25.05	6.2E-3	5.2E-3	2.13	27.04
			s.d.	9.6E-3	8.4E-3	23.31	30.39	8.6E-3	7.5E-3	25.23	32.60
5	0	24	mean	5.8E-3	4.9E-3	0.36	17.55	5.8E-3	4.8E-3	0.45	16.64
			s.d.	8.5E-3	7.4E-3	15.39	17.99	8.1E-3	7.1E-3	13.91	17.48

(a) Best forecasting performance.

Lay	Nod	AR		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	-10	24	mean	8.8E-3	7.7E-3	80.75	84.32	8.3E-3	7.4E-3	87.34	90.32
			s.d.	1.1E-2	9.1E-3	151.88	150.00	8.5E-3	7.4E-3	162.70	161.12
2	0	24	mean	7.4E-3	6.5E-3	74.98	78.99	7.0E-3	6.1E-3	77.78	80.93
			s.d.	9.5E-3	8.6E-3	127.85	125.48	8.4E-3	7.5E-3	131.15	129.30
8	10	24	mean	7.0E-3	6.0E-3	46.53	51.42	7.2E-3	6.2E-3	49.86	53.63
			s.d.	7.5E-3	6.4E-3	77.76	75.80	8.1E-3	7.0E-3	78.39	76.55

(b) Worst forecasting performance.

Table 5.6: Forecasting performance for 24 hours ahead in NAR model.

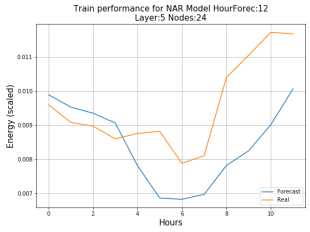
Lay	Nod	AR		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
5	10	24	mean	6.9E-3	5.5E-3	10.29	29.08	7.0E-3	5.6E-3	10.82	29.10
			s.d.	8.0E-3	6.3E-3	25.16	38.34	8.9E-3	7.2E-3	25.24	33.64
8	-10	24	mean	8.5E-3	6.9E-3	-2.70	34.57	7.9E-3	6.4E-3	-1.42	32.49
			s.d.	1.1E-2	8.5E-3	37.01	41.30	9.4E-3	7.6E-3	39.19	42.19
8	10	24	mean	7.2E-3	5.8E-3	11.84	34.12	7.2E-3	5.8E-3	13.44	32.84
			s.d.	7.5E-3	6.0E-3	26.35	53.02	8.1E-3	6.5E-3	28.86	47.25

(a) Best forecasting performance NAR.

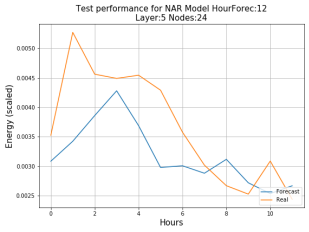
Lay	Nod	AR		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	0	24	mean	8.4E-3	7.0E-3	-71.60	75.96	8.7E-3	7.2E-3	-70.13	74.77
			s.d.	8.7E-3	6.9E-3	111.38	109.32	1.0E-2	8.2E-3	118.72	116.42
2	10	24	mean	8.1E-3	6.8E-3	-23.56	32.85	8.4E-3	7.1E-3	-23.33	32.90
			s.d.	1.0E-2	8.5E-3	41.53	42.16	1.0E-2	8.5E-3	53.22	54.00
5	-10	24	mean	7.9E-3	6.3E-3	-2.15	30.76	8.8E-3	7.1E-3	0.22	28.25
			s.d.	9.5E-3	7.5E-3	27.32	35.23	1.2E-2	9.7E-3	25.19	32.50

(b) Worst forecasting performance.

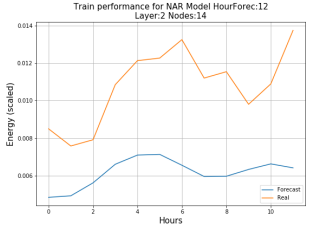
Figures 5.5 and 5.6 exhibit some consumption forecasts for the different time windows. Each forecasts performance is depicted. In some cases, forecasts results take serious approach to the real building behavior; however, the worst models found do not permit a consumption pattern prediction.



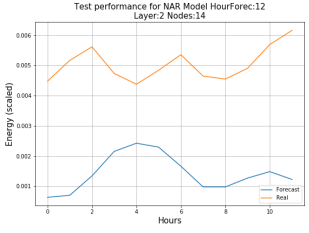
(a) Best 12 hours forecast in train set



(b) Best 12 hours forecast in test set

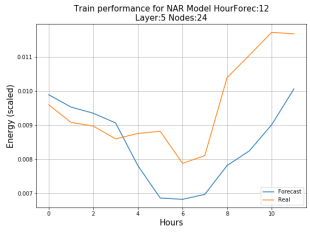


(c) Worst 12 hours forecast in train set

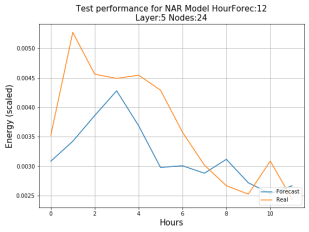


(d) Worst 12 hours forecast in test set

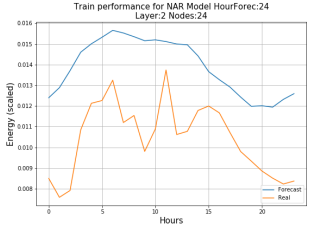
Figure 5.5: Forecasting performance on NAR model for 12 hours prediction.



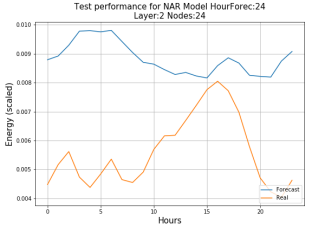
(a) Best 24 hours forecast in train set



(b) Best 24 hours forecast in test set



(c) Worst 24 hours forecast in train set

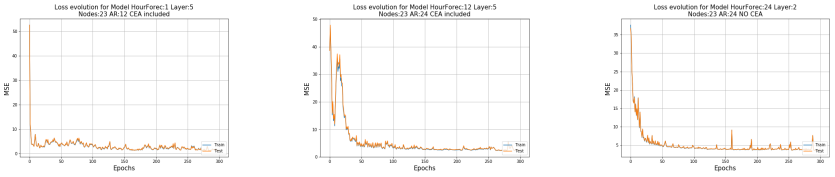


(d) Worst 24 hours forecast in test set

Figure 5.6: Forecasting performance on NAR model for 24 hours prediction.

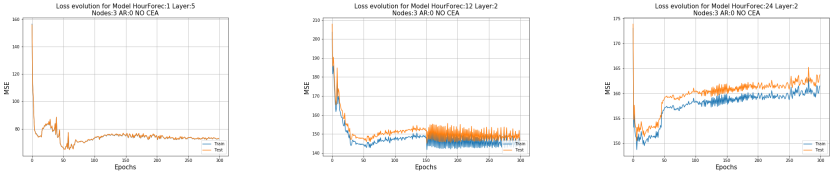
5.4 Multivariate non-linear model -NARX-

There were a total of 108 NARX models tested. All buildings were used as entry data divided into train and test sets (70% and 30% respectively). These kind of models work with the purpose of achieving the energy demand forecast using different variables from the energy consumption time series. AR values inclusion was verified to evaluate the forecast performances improvement. In the same way, it is intended to enhance that improvement with the CEA Simulation tool variables. Variations in layers, nodes and inclusion or exclusion of mentioned variable groups were also tested. Some models training evolution can be perceived in Figures 5.7 and 5.8. In addition, Figure 5.7b shows an anomalous behavior as the loss function descent halts to ascend quickly, and then, descend further again even more. In spite of that, as well as the images that accompany this last one, it is shown a constant descent in the loss function, allowing to understand that the models are in a learning process for interpreting the energy consumption patterns. It can also be argued that Figure 5.8 depicts erratic behavior on the loss function descent curve until it reaches growth in a function that was minimizing (see Figure 5.8c). This particular situation is presented on the model with the least amount of layers, nodes and entry variables.



(a) Forecast: 1 hour (b) Forecast: 12 hours (c) Forecast: 24 hours

Figure 5.7: Best models loss evolution with NARX.



(a) Forecast: 1 hour (b) Forecast: 12 hours (c) Forecast: 24 hours

Figure 5.8: Worst models loss evolution with NARX.

Error results are registered in Tables 5.7, 5.8 and 5.9. A frequent presence of two kind of variables is recognized in the best performance models: AR terms and CEA variables. It has been also detected an increase over the forecast errors values since the lengthened time lapses, in spite of this, in comparison to the NAR model, the variations are still lower. For all the forecast time windows, models of 2 layers, 10 nodes less and no CEA variables are settled into the worst performance models.

Table 5.7: Forecasting performance for 1 hour ahead in NARX model.

Lay	Nod	AR	CEA		Train				Test			
					RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	0	24	1	mean	1.9E-3	1.9E-3	8.12	11.96	1.7E-3	1.7E-3	9.40	13.59
				s.d.	2.9E-3	2.9E-3	23.87	22.19	2.4E-3	2.4E-3	26.57	24.69
2	10	24	1	mean	1.8E-3	1.8E-3	-10.9	13.76	1.8E-3	1.8E-3	-8.27	11.43
				s.d.	2.4E-3	2.4E-3	23.76	22.20	2.6E-3	2.6E-3	20.47	18.88
5	10	12	1	mean	1.7E-3	1.7E-3	-0.24	10.05	1.9E-3	1.9E-3	0.28	9.19
				s.d.	2.4E-3	2.4E-3	23.09	20.78	2.5E-3	2.5E-3	17.55	14.95
8	10	24	1	mean	2.0E-3	2.0E-3	-11.7	15.14	2.1E-3	2.1E-3	-11.5	14.75
				s.d.	2.4E-3	2.4E-3	29.56	27.93	3.2E-3	3.2E-3	28.58	27.06

(a) Best forecasting performance.

Lay	Nod	AR	CEA		Train				Test			
					RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	-10	0	0	mean	2.0E-2	2.0E-2	-154.7	177.52	1.9E-2	1.9E-2	-117.1	137.28
				s.d.	2.7E-2	2.7E-2	581.21	574.64	2.6E-2	2.6E-2	312.65	304.31
2	0	0	0	mean	8.0E-3	8.0E-3	77.26	81.80	8.1E-3	8.1E-3	85.87	89.09
				s.d.	9.7E-3	9.7E-3	128.33	125.48	9.1E-3	9.1E-3	160.50	158.73
2	-10	0	1	mean	8.8E-3	8.8E-3	100.57	108.44	9.5E-3	9.5E-3	86.75	94.82
				s.d.	8.5E-3	8.5E-3	225.73	222.05	9.6E-3	9.6E-3	171.79	167.47
5	-10	0	0	mean	1.7E-2	1.7E-2	-150.43	162.08	1.8E-2	1.8E-2	-137.47	149.40
				s.d.	2.1E-2	2.1E-2	315.22	309.39	2.2E-2	2.2E-2	363.56	358.82
8	-10	0	0	mean	1.5E-2	1.5E-2	12.14	120.04	1.5E-2	1.5E-2	26.54	111.17
				s.d.	1.6E-2	1.6E-2	292.74	267.24	1.8E-2	1.8E-2	303.79	283.94

(b) Worst forecasting performance.

Table 5.8: Forecasting performance for 12 hour ahead in NARX model.

Lay	Nod	AR	CEA		Train				Test			
					RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	0	24	0	mean	4.2E-3	3.5E-3	-8.97	29.09	3.9E-3	3.2E-3	-7.41	28.81
				s.d.	4.4E-3	3.7E-3	35.14	45.38	3.9E-3	3.2E-3	35.24	46.04
2	10	24	1	mean	3.5E-3	2.9E-3	1.76	21.11	3.6E-3	3.0E-3	2.07	18.94
				s.d.	3.5E-3	2.9E-3	24.18	30.34	3.7E-3	2.9E-3	21.37	27.51
5	10	24	1	mean	3.5E-3	2.9E-3	-9.87	18.68	3.3E-3	2.7E-3	-8.37	18.09
				s.d.	3.5E-3	2.8E-3	26.86	28.81	3.0E-3	2.5E-3	23.56	26.78
8	10	24	1	mean	3.8E-3	3.1E-3	-24.9	29.76	3.8E-3	3.1E-3	-23.6	28.10
				s.d.	4.2E-3	3.6E-3	47.67	47.88	3.9E-3	3.2E-3	50.00	49.88

(a) Best forecasting performance.

Lay	Nod	AR	CEA		Train				Test			
					RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	-10	0	0	mean	2.3E-2	2.1E-2	60.33	101.72	2.4E-2	2.2E-2	61.01	102.13
				s.d.	2.7E-2	2.6E-2	173.27	153.63	2.8E-2	2.7E-2	176.14	156.89
2	0	0	0	mean	1.1E-2	9.9E-3	64.68	77.15	1.1E-2	1.0E-2	51.48	63.84
				s.d.	9.9E-3	9.0E-3	204.44	201.46	1.0E-2	9.2E-3	126.16	122.79
2	-10	0	1	mean	1.2E-2	1.1E-2	82.71	90.59	1.2E-2	1.1E-2	84.17	90.86
				s.d.	9.4E-3	8.4E-3	194.62	192.43	1.0E-2	9.0E-3	204.60	202.09
5	-10	0	0	mean	2.1E-2	1.9E-2	-40.2	69.11	2.1E-2	1.9E-2	-30.6	60.22
				s.d.	2.5E-2	2.3E-2	134.44	124.08	2.5E-2	2.4E-2	102.44	90.81

(b) Worst forecasting performance.

Table 5.9: Forecasting performance for 24 hour ahead in NARX model.

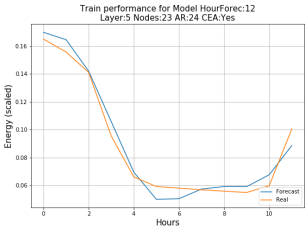
Lay	Nod	AR	CEA		Train				Test			
					RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	10	24	0	mean	4.6E-3	3.7E-3	10.38	22.46	4.5E-3	3.6E-3	10.36	22.53
				s.d.	4.6E-3	3.8E-3	31.04	35.05	4.6E-3	3.8E-3	23.84	29.72
5	10	12	1	mean	4.7E-3	3.8E-3	-20.8	30.58	4.7E-3	3.8E-3	-23.8	34.70
				s.d.	4.1E-3	3.3E-3	41.27	42.94	3.9E-3	3.1E-3	46.86	51.61
5	0	24	1	mean	4.6E-3	3.7E-3	2.37	25.79	4.6E-3	3.8E-3	1.77	27.69
				s.d.	4.2E-3	3.5E-3	20.11	35.17	4.6E-3	3.7E-3	30.01	41.57
8	0	24	1	mean	4.7E-3	3.8E-3	-16.1	27.40	4.4E-3	3.5E-3	-15.8	26.72
				s.d.	4.7E-3	3.9E-3	32.15	37.90	4.4E-3	3.5E-3	33.05	37.58

(a) Best forecasting performance

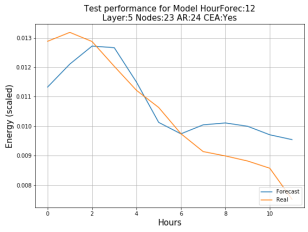
Lay	Nod	AR	CEA		Train				Test			
					RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	-10	0	0	mean	2.5E-2	2.3E-2	-44.6	78.75	2.7E-2	2.5E-2	-45.8	80.10
				s.d.	3.0E-2	2.8E-2	134.42	119.90	3.0E-2	2.8E-2	136.68	121.88
2	0	0	0	mean	1.2E-2	1.0E-2	172.40	180.75	1.2E-2	1.0E-2	160.37	169.16
				s.d.	8.1E-3	6.9E-3	310.47	305.84	7.8E-3	6.6E-3	296.33	291.58
5	-10	0	0	mean	2.0E-2	1.8E-2	4.77	49.87	2.2E-2	2.0E-2	8.00	50.01
				s.d.	2.3E-2	2.1E-2	52.53	37.52	2.6E-2	2.4E-2	56.11	44.22
8	-10	0	0	mean	1.9E-2	1.7E-2	41.04	57.74	1.8E-2	1.6E-2	43.46	57.31
				s.d.	2.0E-2	1.7E-2	56.57	45.93	1.9E-2	1.7E-2	55.91	47.48

(b) Worst forecasting performance

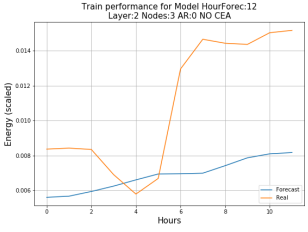
The prediction shown in Figures 5.9 and 5.10 allows to observe the energy consumption models adjustment degree. The upper figures reached a good approximation to the energy consumption in both training and testing sets without evidence of over fitting; whereas the ones below the image fail to achieve some basic patterns of energy demand adjustment



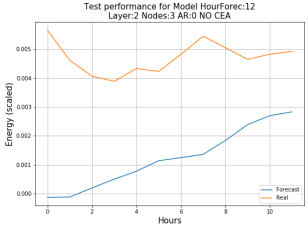
(a) Best forecast in train set



(b) Best forecast in test set

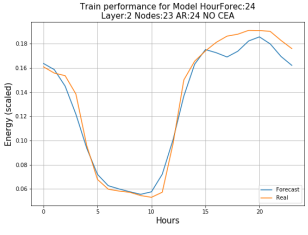


(c) Worst forecast in train set

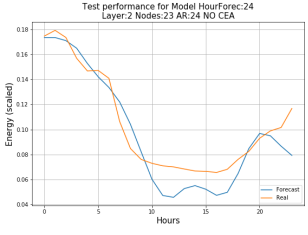


(d) Worst forecast in test set

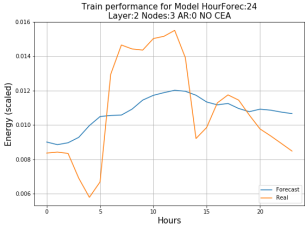
Figure 5.9: Forecasting performance on NARX model for 12 hours prediction.



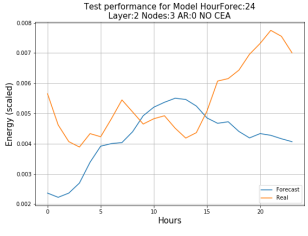
(a) Best forecast in train set



(b) Best forecast in test set



(c) Worst forecast in train set

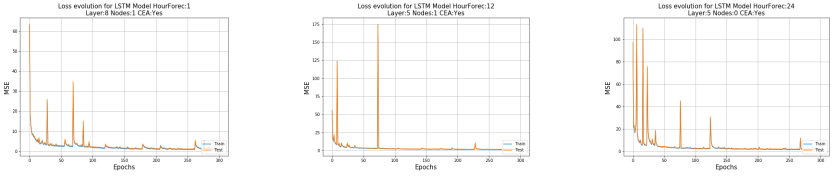


(d) Worst forecast in test set

Figure 5.10: Forecasting performance on NARX model for 24 hours prediction.

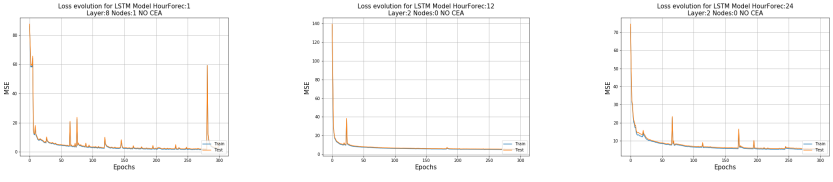
5.5 Recurrent neural network -RNN-

The RNN with LSTM cell configuration was implemented modifying the parameters mentioned above (layers, nodes, outputs). Recalling the main periodicity found for the SARIMA model, 24 hours gap was used for feeding the RNN. All data was sorted to accomplish the time dimension required for the LSTM configuration. The loss evolution over the epochs is shown in Figures 5.11 and 5.12. Allowing to observe the energy consumption models adjustment degree. The images in Figure 5.11 reached a good approximation to the energy consumption in both training and testing sets without evidence of over fitting; whereas the ones in Figure 5.12 fail to achieve some basic patterns of energy demand adjustment. The descendent behaviour shown in these images, in spite of having a smoother process of forecasting, do not stand out.



(a) Forecast: 1 hour (b) Forecast: 1 hours (c) Forecast: 24 hours

Figure 5.11: Best models loss evolution with RNN.



(a) Forecast: 1 hour (b) Forecast: 12 hours (c) Forecast: 24 hours

Figure 5.12: Worst models loss evolution with RNN.

The results of calculated errors for each time window forecast are recorded in Tables 5.10, 5.11 and 5.12. In general, all time windows have at least one of the best models with MAPE under 10%. Yet, once again CEA variables are remarkably outstanding due to, on one side, occupying places all over the best models tables; and on the other, displaying absence on the tables marking the worst. It is also remarkable that no 2-layered settled model is recognized as effective, thus, signifying that the architecture complexity of the network and given data requires a deeper approach. Moreover, regarding the best models, a small variation is easily found on

the test set error values that goes below in comparison to the corresponding training sets, although some of them do not present great differences.

Table 5.10: Forecasting performance for 1 hour ahead in RNN model.

Lay	Nod	CEA		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
5	10	1	mean	2.7E-3	2.1E-3	-1.35	8.95	2.5E-3	2.0E-3	-1.60	9.14
			s.d.	3.3E-3	2.6E-3	11.31	10.65	2.6E-3	2.1E-3	12.01	11.46
8	0	1	mean	2.7E-3	2.1E-3	-1.71	9.13	2.5E-3	2.0E-3	-2.39	9.67
			s.d.	4.8E-3	3.6E-3	10.94	10.13	2.7E-3	2.2E-3	12.97	11.70
8	10	1	mean	2.4E-3	1.9E-3	-4.06	8.10	2.2E-3	1.7E-3	-3.86	8.97
			s.d.	4.6E-3	2.9E-3	12.00	11.55	2.1E-3	1.7E-3	14.06	13.72

(a) Best forecasting performance.

Lay	Nod	CEA		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	0	0	mean	4.5E-3	3.7E-3	-9.65	20.19	4.7E-3	3.9E-3	-11.7	22.60
			s.d.	5.1E-3	4.3E-3	29.59	29.73	5.5E-3	4.6E-3	34.39	34.19
2	10	0	mean	4.5E-3	3.7E-3	-11.9	22.04	4.9E-3	4.0E-3	-13.4	23.09
			s.d.	5.5E-3	4.7E-3	31.52	32.61	7.8E-3	5.9E-3	33.29	34.26
8	10	0	mean	5.6E-3	4.7E-3	-25.2	32.38	5.5E-3	4.6E-3	-27.7	34.82
			s.d.	5.5E-3	4.7E-3	59.69	57.61	6.2E-3	5.1E-3	60.55	58.20

(b) Worst forecasting performance.

Table 5.11: Forecasting performance for 12 hour ahead in RNN model.

Lay	Nod	CEA		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
5	10	1	mean	2.6E-3	2.0E-3	-1.69	8.37	2.5E-3	2.0E-3	-1.19	7.93
			s.d.	5.1E-3	3.2E-3	11.06	10.30	2.8E-3	2.0E-3	10.35	9.76
8	0	1	mean	2.6E-3	2.1E-3	-1.65	8.91	2.6E-3	2.0E-3	-1.74	8.74
			s.d.	2.9E-3	2.3E-3	12.46	11.55	2.4E-3	1.9E-3	11.60	10.89
8	10	1	mean	2.3E-3	1.9E-3	-0.90	8.34	2.3E-3	1.9E-3	-0.48	8.13
			s.d.	2.4E-3	1.9E-3	11.71	11.11	2.4E-3	1.9E-3	10.48	9.50

(a) Best forecasting performance.

Lay	Nod	CEA		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	0	0	mean	4.9E-3	4.0E-3	1.78	22.39	4.5E-3	3.7E-3	4.74	21.82
			s.d.	5.3E-3	4.4E-3	33.24	33.77	4.7E-3	3.9E-3	30.62	31.86
5	-10	0	mean	4.6E-3	3.8E-3	-13.2	22.25	4.7E-3	3.8E-3	-10.1	19.06
			sd	5.2E-3	4.4E-3	36.10	35.03	5.2E-3	4.4E-3	28.81	27.56
8	-10	0	mean	4.8E-3	4.0E-3	-0.44	17.75	4.7E-3	3.8E-3	-1.79	19.53
			sd	6.1E-3	5.3E-3	24.60	25.76	6.0E-3	5.0E-3	28.39	29.21

(b) Worst forecasting performance.

Table 5.12: Forecasting performance for 24 hour ahead in RNN model.

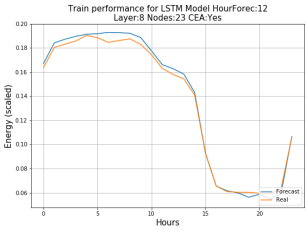
Lay	Nod	CEA		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
5	0	1	mean	2.8E-3	2.3E-3	-1.52	9.81	2.8E-3	2.2E-3	-2.00	9.78
			s.d.	2.9E-3	2.3E-3	16.93	16.03	2.7E-3	2.2E-3	13.70	12.47
5	10	1	mean	2.6E-3	2.0E-3	-1.69	8.37	2.5E-3	2.0E-3	-1.19	7.93
			s.d.	5.1E-3	3.2E-3	11.06	10.30	2.8E-3	2.0E-3	10.35	9.76
8	10	1	mean	2.3E-3	1.9E-3	-0.90	8.34	2.3E-3	1.9E-3	-0.48	8.13
			s.d.	2.4E-3	1.9E-3	11.71	11.11	2.4E-3	1.9E-3	10.48	9.50

(a) Best forecasting performance.

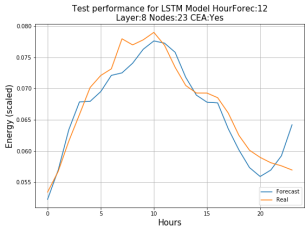
Lay	Nod	CEA		Train				Test			
				RMSE	MAE	MPE	MAPE	RMSE	MAE	MPE	MAPE
2	-10	0	mean	4.7E-3	3.8E-3	-5.47	18.57	4.5E-3	3.7E-3	-8.49	21.04
			s.d.	5.3E-3	4.6E-3	25.51	26.70	4.9E-3	4.1E-3	30.55	31.54
2	0	0	mean	4.9E-3	4.0E-3	1.78	22.39	4.5E-3	3.7E-3	4.74	21.82
			s.d.	5.3E-3	4.4E-3	33.24	33.77	4.7E-3	3.9E-3	30.62	31.86
8	-10	0	mean	4.8E-3	4.0E-3	-0.44	17.75	4.7E-3	3.8E-3	-1.79	19.53
			s.d.	6.1E-3	5.3E-3	24.60	25.76	6.0E-3	5.0E-3	28.39	29.21

(b) Worst forecasting performance.

The performance obtained for the 8-layer model with 10 extra nodes and CEA variables is outlining, owing to the low error found in the three time windows analyzed. Forecasts of the best and worst model are registered on Figures 5.13 and 5.14. From the graphics and errors, it is possible to assert that the setting of LSTM cells allows a better adjustment to the consumption patterns with less error, even a 30% for the worst models; the closeness to varied of the same patterns is also noticed too.

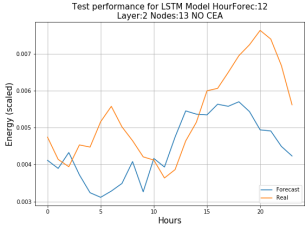


(a) Best forecast in train set



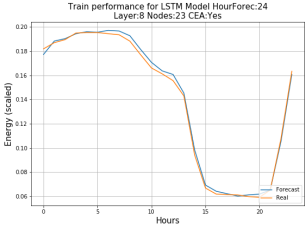
(b) Best forecast in test set

(c) Worst forecast in train set

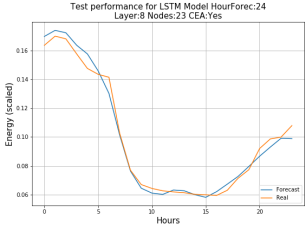


(d) Worst forecast in test set

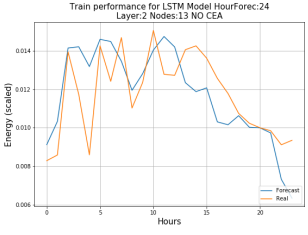
Figure 5.13: Forecasting performance on NARX model for 12 hours prediction.



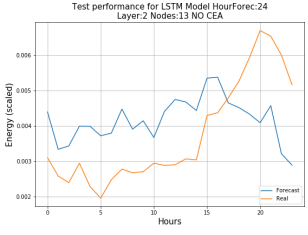
(a) Best forecast in train set



(b) Best forecast in test set



(c) Worst forecast in train set



(d) Worst forecast in test set

Figure 5.14: Forecasting performance on RNN model for 24 hours prediction.

Chapter 6

Discussion and conclusions

The electric energy consumption time series were characterized by strong periodicity patterns related to the specific use of each building. There is a remarkable difference between weekdays and weekends electricity consumption sequences for most of buildings. Due to its geographical location, Singapore climatic variations are low and sudden or large changes were not presented, producing homogeneous behavior on power requirement. However, the hybrid model proposed in this work (real and simulated data as input of a regression-based forecasting model) was able to provide a reliable energy consumption forecasting method by using both simulated and metered data besides a proper regression approach.

The SARIMA model exhibited a poor performance and did not reach an acceptable fit of the time series for most of the buildings in the dataset, likewise, sometimes presented large forecast deviations from the actual metered energy consumption data. Moreover, the resulting SARIMA model structure exhibited a high variation between buildings and required five or more AR and/or MA terms, sometimes reaching almost a half of the season length. Indeed, there are considerable variations in the number of stationary and seasonal component coefficients, causing that every new building will require a specific training process, including ACF and PACF analysis. Given such drawback, this model appears as an expensive alternative for a large number of buildings, since it lacks a unique, generalized structure to produce reliable forecasts for every new instance (building). The possible addition of exogenous variables to propose a SARIMAX approach could be an interesting research path to confirm such limitations.

The NAR method is designed to achieve a more generalized model for several buildings using the non-linear combinations obtained from NN. However, the prediction performance decays quickly with longer time forecasting. The best models for the time-windows evaluated reached MAPE values of 7% for 1-hour ahead forecast to 17% for 12-hour ahead and 30% for 24-hour ahead prediction. In 1-hour

prediction the number of layers in the NN ("depth") becomes relevant to the model performance:, better performances are based on an 8-layer architecture. Despite the increase in nodes, it did not lead to a better prediction in tested scenarios.

Despite of the generalization achievement for multiple buildings at once, the performance is not outstanding. As a result, this type of models should be considered for short length periods when there is no external data.

On the other hand, the NARX models explored in this study provide sound forecasts for all buildings. The inclusion of AR terms in the input features enhanced both the model training convergence and the forecasting performance. Architectures without AR terms cannot reach the performance levels achieved by their counterparts with 24 AR terms. The non-linear combination of explanatory variables and AR terms in the NN architecture could generalize the energy forecasting for the buildings in the dataset, with a reliable fitting of the real data. Including simulated variables from the CEA toolbox improved the NARX performance for most of the scenarios, with a MAPE reduction of about 10% in the 24-hour forecasting, 16% in the 12-hour forecasting and 22% in the 1-hour forecasting. Furthermore, the number of epochs to achieve a good performance were much lower when the CEA variables were included. However, the amount of input variables affects model training, making only the deeper and wider ones to achieve better performances. Opposite to NAR, External variables inclusion prevent from error rising when longer forecasts are calculated.

The RNN LSTM-based models performed even better than the NARX models for all forecasting scenarios, regardless the time-window selected for each test. A reduction of 55% in MAPE value of the best 24-hour forecasting model was achieved using LSTM instead NARX as regression approach. Most of the models in this study provide better predictions by adding extra nodes to each layer (up to 10 in the proposed experimentation) and if extra layers are included (up to 3 for this study). A combination of LSTM cells with CEA variables yields the best performance for all time-windows. The best model for 24-hour energy demand forecasting was an RNN with LSTM cells, 8 layers and 10 extra nodes including the CEA simulation variables. This architecture reduced the MAPE value in about 30% compared to the model without CEA variables. The error decay is smoother than it is for the NARX models, and a proper fitting performance is rapidly achieved before 50 epochs. The forecast comparison proves that the training and testing processes exhibit an appropriate performance, accurately predicting the variations in the energy consumption real data.

A noisy behavior in the evolution of training error was presented, possibly caused by the random mini-batch training. The models fit the consumption patterns with no strong evidence of overfitting even with the large epochs value considered. The inclusion of dropout methods can improve training convergence and the forecasting performance of the model.

The wider models (more nodes per layer) but shallower struggled in handling large amount of variables getting most of the worst performances for all models considered.

Nevertheless narrow and deeper models did not achieve the best performance in all models. According to the results of this study, 2-layer models should be avoided and it can be recommended to include at least the same amount of variables than nodes per each layer.

Most of NN based models significantly reduced the loss function in less than 100 epochs, the use of normalization and mini-batch training helped in the training speed. Early stopping can be implemented for saving training time.

Using clustering techniques to group buildings according similar attributes can be considered as an interesting approach for getting even better performances in the proposed models.

Bibliography

[1] Martın Abadi et al. “TensorFlow: A system for large-scale machine learning”. en. In: (2016), p. 21.

[2] Yuvraj Agarwal et al. “Occupancy-driven energy management for smart building automation”. In: *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*. BuildSys ’10. Zurich, Switzerland: Association for Computing Machinery, Nov. 2010, pp. 1–6. ISBN: 978-1-4503-0458-0. DOI: 10.1145/1878431.1878433. URL: https://doi.org/10.1145/1878431.1878433 (visited on 07/22/2020).

[3] Muhammad Waseem Ahmad, Monjur Mourshed, and Yacine Rezgui. “Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption”. en. In: *Energy and Buildings* 147 (July 2017), pp. 77–89. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2017.04.038. URL: http://www.sciencedirect.com/science/article/pii/S0378778816313937 (visited on 12/08/2020).

[4] Muhammad Waseem Ahmad et al. “Random Forests and Artificial Neural Network for Predicting Daylight Illuminance and Energy Consumption”. en. In: (2017), p. 7.

[5] H Akbari, M Pomerantz, and H Taha. “Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas”. en. In: *Solar Energy*. Urban Environment 70.3 (Jan. 2001), pp. 295–310. ISSN: 0038-092X. DOI: 10.1016/S0038-092X(00)00089-X. URL: http://www.sciencedirect.com/science/article/pii/S0038092X0000089X (visited on 06/02/2020).

[6] K. P. Amber, M. W. Aslam, and S. K. Hussain. “Electricity consumption forecasting models for administration buildings of the UK higher education sector”. en. In: *Energy and Buildings* 90 (Mar. 2015), pp. 127–136. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2015.01.008. URL: http://www.sciencedirect.com/science/article/pii/S0378778815000110 (visited on 07/22/2020).

[7] K. P. Amber et al. “Intelligent techniques for forecasting electricity consumption of buildings”. en. In: *Energy* 157 (Aug. 2018), pp. 886–893. ISSN: 0360-5442. DOI: 10.1016/j.energy.2018.05.155. URL: http://www.sciencedirect.com/science/article/pii/S036054421830999X (visited on 06/02/2020).

[8] A. Angelova, A. Krizhevsky, and V. Vanhoucke. “Pedestrian detection with a Large-Field-Of-View deep network”. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. May 2015, pp. 704–711. DOI: 10.1109/ICRA.2015.7139256.

[9] Maher Azaza, Douglas Eriksson, and Fredrik Wallin. “A study on the viability of an on-site combined heat- and power supply system with and without electricity storage for office building”. en. In: *Energy Conversion and Management* 213 (June 2020), p. 112807. ISSN: 0196-8904. DOI: 10.1016/j.enconman.2020.112807. URL: http://www.sciencedirect.com/science/article/pii/S0196890420303459 (visited on 06/02/2020).

[10] Jean-Marie Bahu et al. “Towards a 3D Spatial Urban Energy Modelling Approach”. en. In: *International Journal of 3-D Information Modeling (IJ3DIM)* (July 2014). ISSN: 2156-1710 Issue: 3 Journal Abbreviation: IJ3DIM Pages: 1-16 Publisher: IGI Global Volume: 3. DOI: 10.4018/ij3dim.2014070101. URL: www.igi-global.com/article/towards-a-3d-spatial-urban-energy-modelling-approach/122864 (visited on 12/09/2020).

[11] World Bank. *Electric power consumption (kWh per capita) | Data*. URL: https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC?end=2018&start=1960&view=chart (visited on 01/15/2020).

[12] S. Bhavya and Anitha S. Pillai. “Prediction Models in Healthcare Using Deep Learning”. en. In: *Proceedings of the 11th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2019)*. Ed. by Ajith Abraham et al. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2021, pp. 195–204. ISBN: 978-3-030-49345-5. DOI: 10.1007/978-3-030-49345-5_21.

[13] Mathieu Bourdeau et al. “Modeling and forecasting building energy consumption: A review of data-driven techniques”. en. In: *Sustainable Cities and Society* 48 (July 2019), p. 101533. ISSN: 2210-6707. DOI: 10.1016/j.scs.2019.101533. URL: http://www.sciencedirect.com/science/article/pii/S2210670718323862 (visited on 06/02/2020).

[14] G. E. P. Box and D. R. Cox. “An Analysis of Transformations”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 26.2 (1964). Publisher: [Royal Statistical Society, Wiley], pp. 211–252. ISSN: 0035-9246. URL: https://www.jstor.org/stable/2984418 (visited on 12/10/2020).

[15] George Edward Pelham Box and Gwilym Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, Incorporated, 1976. ISBN: 978-0-8162-1104-3.

[16] BP. *Statistical Review of World Energy | Energy economics | Home*. en. URL: https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html (visited on 06/02/2020).

[17] James E. Braun and Nitin Chaturvedi. "An Inverse Gray-Box Model for Transient Building Load Prediction". In: *HVAC&R Research* 8.1 (Jan. 2002). Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/10789669.2 pp. 73–99. ISSN: 1078-9669. DOI: 10.1080/10789669.2002.10391290. URL: https://www.tandfonline.com/doi/abs/10.1080/10789669.2002.10391290 (visited on 11/20/2020).

[18] Lars Buitinck et al. "API design for machine learning software: experiences from the scikit-learn project". In: *arXiv:1309.0238 [cs]* (Sept. 2013). arXiv: 1309.0238. URL: http://arxiv.org/abs/1309.0238 (visited on 02/01/2019).

[19] Cao, Xiaodong, Xilei, Dai, and Liu, Junjie. "Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade". en. In: *Energy and Buildings* 128 (Sept. 2016), pp. 198–213. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2016.06.089. URL: https://ezproxy.ucentral.edu.co:2154/science/article/pii/S0378778816305783 (visited on 01/15/2020).

[20] Kyunghyun Cho et al. "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation". In: *arXiv:1406.1078 [cs, stat]* (Sept. 2014). arXiv: 1406.1078. URL: http://arxiv.org/abs/1406.1078 (visited on 11/20/2020).

[21] K. J. Chua et al. "Achieving better energy-efficient air conditioning – A review of technologies and strategies". en. In: *Applied Energy* 104 (Apr. 2013), pp. 87–104. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2012.10.037. URL: http://www.sciencedirect.com/science/article/pii/S030626191200743X (visited on 06/02/2020).

[22] *City Energy Analyst (CEA)*. en-US. URL: https://cityenergyanalyst.com (visited on 12/09/2020).

[23] M. R. Cogollo and J. D. Velasquez. "Are neural networks able to forecast nonlinear time series with moving average components?" In: *IEEE Latin America Transactions* 13.7 (July 2015). Conference Name: IEEE Latin America Transactions, pp. 2292–2300. ISSN: 1548-0992. DOI: 10.1109/TLA.2015.7273790.

[24] J. Contreras et al. “ARIMA models to predict next-day electricity prices”. In: *IEEE Transactions on Power Systems* 18.3 (Aug. 2003). Conference Name: IEEE Transactions on Power Systems, pp. 1014–1020. ISSN: 1558-0679. DOI: 10.1109/TPWRS.2002.804943.

[25] Chirag Deb et al. “Forecasting Energy Consumption of Institutional Buildings in Singapore”. en. In: *Procedia Engineering*. The 9th International Symposium on Heating, Ventilation and Air Conditioning (ISHVAC) joint with the 3rd International Conference on Building Energy and Environment (COBEE), 12-15 July 2015, Tianjin, China 121 (Jan. 2015), pp. 1734–1740. ISSN: 1877-7058. DOI: 10.1016/j.proeng.2015.09.144. URL: http://www.sciencedirect.com/science/article/pii/S1877705815029720 (visited on 07/22/2020).

[26] Statista Research Department. *Devices IT spending worldwide 2012-2020*. en. URL: https://www.statista.com/statistics/314584/total-devices-spending-worldwide-forecast/ (visited on 06/02/2020).

[27] N. Ding et al. “Neural Network-Based Model Design for Short-Term Load Forecast in Distribution Systems”. In: *IEEE Transactions on Power Systems* 31.1 (Jan. 2016), pp. 72–81. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2015.2390132.

[28] Bing Dong et al. “A hybrid model approach for forecasting future residential electricity consumption”. en. In: *Energy and Buildings* 117 (Apr. 2016), pp. 341–351. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2015.09.033. URL: http://www.sciencedirect.com/science/article/pii/S0378778815302735 (visited on 11/17/2020).

[29] Bing Dong et al. “A hybrid model approach for forecasting future residential electricity consumption”. en. In: *Energy and Buildings* 117 (Apr. 2016), pp. 341–351. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2015.09.033. URL: http://www.sciencedirect.com/science/article/pii/S0378778815302735 (visited on 11/20/2020).

[30] Peter Droege. *Urban Energy Transition: From Fossil Fuels to Renewable Power*. en. Google-Books-ID: 0bEecvixZRMC. Elsevier, Sept. 2011. ISBN: 978-0-08-056046-5.

[31] John Duchi, Elad Hazan, and Yoram Singer. “Adaptive Subgradient Methods for Online Learning and Stochastic Optimization”. In: *The Journal of Machine Learning Research* 12.null (July 2011), pp. 2121–2159. ISSN: 1532-4435.

[32] Department of Economics and Social Affairs. *2014 revision*. en-US. Report. New York: United Nations, 2014. URL: 2014-revision-world-urbanization-prospects.html (visited on 01/15/2020).

[33] Ursula Eicker et al. "On the design of an urban data and modeling platform and its application to urban district analyses". en. In: *Energy and Buildings* 217 (June 2020), p. 109954. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2020.109954. URL: http://www.sciencedirect.com/science/article/pii/S0378778819322005 (visited on 06/02/2020).

[34] Ammar H. Elsheikh et al. "Utilization of LSTM neural network for water production forecasting of a stepped solar still with a corrugated absorber plate". en. In: *Process Safety and Environmental Protection* 148 (Apr. 2021), pp. 273–282. ISSN: 0957-5820. DOI: 10.1016/j.psep.2020.09.068. URL: http://www.sciencedirect.com/science/article/pii/S0957582020318061 (visited on 11/09/2020).

[35] Zannatul Ferdoush et al. "A short-term hybrid forecasting model for time series electrical-load data using random forest and bidirectional long short-term memory". en. In: *International Journal of Electrical and Computer Engineering (IJECE)* 11.1 (Feb. 2021). Number: 1, pp. 763–771. ISSN: 2722-2578. DOI: 10.11591/ijece.v11i1.pp763-771. URL: http://ijece.iaescore.com/index.php/IJECE/article/view/22929 (visited on 11/09/2020).

[36] Karen Fisher-Vanden, Erin T Mansur, and Qiong (Juliana) Wang. *Costly Blackouts? Measuring Productivity and Environmental Effects of Electricity Shortages*. Working Paper 17741. National Bureau of Economic Research, Jan. 2012. DOI: 10.3386/w17741. URL: http://www.nber.org/papers/w17741 (visited on 01/29/2019).

[37] Jimeno A. Fonseca. *BSTS-SG: An ensemble timeseries model for the evaluation of energy conservation strategies in households*. en. Report. Accepted: 2020-11-02T14:19:51Z Publication Title: Deliverable Technical Report Volume: D1.2.2.6. ETH Zurich, Oct. 2020. DOI: 10.3929/ethz-b-000441382. URL: https://www.research-collection.ethz.ch/handle/20.500.11850/441382 (visited on 12/09/2020).

[38] Jimeno A. Fonseca, Clayton Miller, and Arno Schlueter. "Unsupervised load shape clustering for urban building performance assessment". en. In: *Energy Procedia*. CISBAT 2017 International ConferenceFuture Buildings & Districts – Energy Efficiency from Nano to Urban Scale 122 (Sept. 2017), pp. 229–234. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2017.07.350. URL: http://www.sciencedirect.com/science/article/pii/S1876610217329545 (visited on 11/20/2020).

[39] Jimeno A. Fonseca and Arno Schlueter. "Integrated model for characterization of spatiotemporal building energy consumption patterns in neighborhoods and city districts". In: *Applied Energy* 142 (Mar. 2015), pp. 247–265. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2014.12.068. URL: http://www.

sciencedirect.com/science/article/pii/S0306261914013257 (visited on 08/03/2017).

[40] Jimeno A. Fonseca et al. "City Energy Analyst (CEA): Integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts". In: *Energy and Buildings* 113 (Feb. 2016), pp. 202–226. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2015.11.055. URL: http://www.sciencedirect.com/science/article/pii/S0378778815304199 (visited on 07/31/2017).

[41] Felix A. Gers, Jürgen Schmidhuber, and Fred Cummins. "Learning to Forget: Continual Prediction with LSTM". In: *Neural Computation* 12.10 (Oct. 2000), pp. 2451–2471. ISSN: 0899-7667. DOI: 10.1162/089976600300015015. URL: https://www.mitpressjournals.org/doi/10.1162/089976600300015015 (visited on 01/30/2019).

[42] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. "Deep Sparse Rectifier Neural Networks". en. In: *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (June 2011), pp. 315–323.

[43] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. en. Google-Books-ID: Np9SDQAAQBAJ. MIT Press, Nov. 2016. ISBN: 978-0-262-03561-3.

[44] Gabriel Happle, Jimeno A. Fonseca, and Arno Schlueter. "Effects of air infiltration modeling approaches in urban building energy demand forecasts". en. In: *Energy Procedia*. CISBAT 2017 International ConferenceFuture Buildings & Districts – Energy Efficiency from Nano to Urban Scale 122 (Sept. 2017), pp. 283–288. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2017.07.323. URL: http://www.sciencedirect.com/science/article/pii/S1876610217329260 (visited on 11/20/2020).

[45] Gabriel Happle, Jimeno A. Fonseca, and Arno Schlueter. "A review on occupant behavior in urban building energy models". en. In: *Energy and Buildings* 174 (Sept. 2018), pp. 276–292. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2018.06.030. URL: http://www.sciencedirect.com/science/article/pii/S0378778817333583 (visited on 07/22/2020).

[46] V. S. K. V. Harish and Arun Kumar. "A review on modeling and simulation of building energy systems". en. In: *Renewable and Sustainable Energy Reviews* 56 (Apr. 2016), pp. 1272–1292. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.12.040. URL: http://www.sciencedirect.com/science/article/pii/S1364032115014239 (visited on 11/30/2020).

[47] Z. Haydari et al. "Time-series load modelling and load forecasting using neuro-fuzzy techniques". In: *2007 9th International Conference on Electrical Power Quality and Utilisation*. Oct. 2007, pp. 1–6. DOI: 10.1109/EPQU.2007.4424201.

[48] G. Heigold et al. "Multilingual acoustic models using distributed deep neural networks". In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. May 2013, pp. 8619–8623. DOI: 10.1109/ICASSP.2013.6639348.

[49] S. Hepziba Lizzie and B. Senthil Kumar. "Air Quality Prediction Using Time Series Analysis". en. In: *Intelligent Data Engineering and Analytics*. Ed. by Suresh Chandra Satapathy et al. Advances in Intelligent Systems and Computing. Singapore: Springer, 2021, pp. 741–748. ISBN: 9789811556791. DOI: 10.1007/978-981-15-5679-1_72.

[50] H.S. Hippert, C.E. Pedreira, and R.C. Souza. "Neural networks for short-term load forecasting: a review and evaluation". In: *IEEE Transactions on Power Systems* 16.1 (Feb. 2001), pp. 44–55. ISSN: 1558-0679. DOI: 10.1109/59.910780.

[51] Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-Term Memory". In: *Neural Computation* 9.8 (Nov. 1997). Publisher: MIT Press, pp. 1735–1780. ISSN: 0899-7667. DOI: 10.1162/neco.1997.9.8.1735. URL: https://doi.org/10.1162/neco.1997.9.8.1735 (visited on 11/20/2020).

[52] Tianzhen Hong Ph. D. "Data and Analytics to Inform Energy Retrofit of High Performance Buildings". en. In: (June 2014). URL: https://escholarship.org/uc/item/0k32878x (visited on 06/02/2020).

[53] Jörg Huber and Christoph Nytsch-geusen. *Development of Modeling and Simulation Strategies 1 for Large-Scale Urban Districts 2 3*.

[54] Rob Hyndman et al. *forecast: Forecasting Functions for Time Series and Linear Models*. June 2018. URL: https://CRAN.R-project.org/package=forecast (visited on 07/08/2018).

[55] Hyndman R.J.Athanasopoulos G. *Forecasting: Principles and Practice*. 2nd edition. OTexts: Melbourne, Australia, 2018. URL: https://Otexts.com/fpp2/ (visited on 11/18/2020).

[56] IEA. *World Energy Outlook 2019 – Analysis - IEA*. URL: https://www.iea.org/reports/world-energy-outlook-2019 (visited on 06/02/2020).

[57] Sergey Ioffe and Christian Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: *arXiv:1502.03167 [cs]* (Feb. 2015). arXiv: 1502.03167. URL: http://arxiv.org/abs/1502.03167 (visited on 01/30/2019).

[58] Tim Januschowski et al. "Criteria for classifying forecasting methods". en. In: *International Journal of Forecasting*. M4 Competition 36.1 (Jan. 2020), pp. 167–177. ISSN: 0169-2070. DOI: 10.1016/j.ijforecast.2019.05.008. URL: http://www.sciencedirect.com/science/article/pii/S0169207019301529 (visited on 12/04/2020).

[59] Xiaolong Jin et al. "Integrated optimal scheduling and predictive control for energy management of an urban complex considering building thermal dynamics". en. In: *International Journal of Electrical Power & Energy Systems* 123 (Dec. 2020), p. 106273. ISSN: 0142-0615. DOI: 10.1016/j.ijepes.2020.106273. URL: http://www.sciencedirect.com/science/article/pii/S0142061520303094 (visited on 07/22/2020).

[60] Rune Grønborg Junker et al. "Stochastic nonlinear modelling and application of price-based energy flexibility". en. In: *Applied Energy* 275 (Oct. 2020), p. 115096. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2020.115096. URL: http://www.sciencedirect.com/science/article/pii/S0306261920306085 (visited on 07/22/2020).

[61] Prashant V. Kamat. "Meeting the Clean Energy Demand: Nanostructure Architectures for Solar Energy Conversion". In: *The Journal of Physical Chemistry C* 111.7 (Feb. 2007), pp. 2834–2860. ISSN: 1932-7447. DOI: 10.1021/jp066952u. URL: https://doi.org/10.1021/jp066952u (visited on 06/02/2020).

[62] Andrej Karpathy, Justin Johnson, and Li Fei-Fei. "Visualizing and Understanding Recurrent Networks". In: *arXiv:1506.02078 [cs]* (June 2015). arXiv: 1506.02078. URL: http://arxiv.org/abs/1506.02078 (visited on 02/01/2019).

[63] Rajesh G. Kavasseri and Krithika Seetharaman. "Day-ahead wind speed forecasting using f-ARIMA models". en. In: *Renewable Energy* 34.5 (May 2009), pp. 1388–1393. ISSN: 0960-1481. DOI: 10.1016/j.renene.2008.09.006. URL: http://www.sciencedirect.com/science/article/pii/S0960148108003327 (visited on 12/09/2020).

[64] Jin-Young Kim and Sung-Bae Cho. "Interpretable Deep Learning with Hybrid Autoencoders to Predict Electric Energy Consumption". en. In: *15th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2020)*. Ed. by Álvaro Herrero et al. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2021, pp. 133–143. ISBN: 978-3-030-57802-2. DOI: 10.1007/978-3-030-57802-2_13.

[65] Moon Keun Kim, Yang-Seon Kim, and Jelena Srebric. "Predictions of electricity consumption in a campus building using occupant rates and weather elements with sensitivity analysis: Artificial neural network vs. linear regression". en. In: *Sustainable Cities and Society* 62 (Nov. 2020), p. 102385. ISSN: 2210-6707. DOI: 10.1016/j.scs.2020.102385. URL: http://www.sciencedirect.com/science/article/pii/S2210670720306065 (visited on 07/22/2020).

[66] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". In: *arXiv:1412.6980 [cs]* (Dec. 2014). arXiv: 1412.6980. URL: http://arxiv.org/abs/1412.6980 (visited on 01/30/2019).

[67] Manish Kumar, Deepak Kumar Gupta, and Samayveer Singh. "Extreme Event Forecasting Using Machine Learning Models". en. In: *Advances in Communication and Computational Technology*. Ed. by Gurdeep Singh Hura, Ashutosh Kumar Singh, and Lau Siong Hoe. Lecture Notes in Electrical Engineering. Singapore: Springer, 2021, pp. 1503–1514. ISBN: 9789811553417. DOI: 10.1007/978-981-15-5341-7_115.

[68] Jun-Mao Liao, Ming-Jui Chang, and Luh-Maan Chang. "Prediction of Air-Conditioning Energy Consumption in R&D Building Using Multiple Machine Learning Techniques". en. In: *Energies* 13.7 (Jan. 2020), p. 1847. DOI: 10.3390/en13071847. URL: https://www.mdpi.com/1996-1073/13/7/1847 (visited on 06/02/2020).

[69] R Lindberg, A Binamu, and M Teikari. "Five-year data of measured weather, energy consumption, and time-dependent temperature variations within different exterior wall structures". en. In: *Energy and Buildings* 36.6 (June 2004), pp. 495–501. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2003.12.009. URL: http://www.sciencedirect.com/science/article/pii/S0378778804000040 (visited on 07/22/2020).

[70] Zachary C. Lipton, John Berkowitz, and Charles Elkan. "A Critical Review of Recurrent Neural Networks for Sequence Learning". In: *arXiv:1506.00019 [cs]* (May 2015). arXiv: 1506.00019. URL: http://arxiv.org/abs/1506.00019 (visited on 01/30/2019).

[71] Shaohui Ma. "A hybrid deep meta-ensemble networks with application in electric utility industry load forecasting". en. In: *Information Sciences* 544 (Jan. 2021), pp. 183–196. ISSN: 0020-0255. DOI: 10.1016/j.ins.2020.07.054. URL: http://www.sciencedirect.com/science/article/pii/S0020025520307179 (visited on 11/09/2020).

[72] Claudio Martani et al. "ENERNET: Studying the dynamic relationship between building occupancy and energy consumption". en. In: *Energy and Buildings* 47 (Apr. 2012), pp. 584–591. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2011.12.037. URL: http://www.sciencedirect.com/science/article/pii/S0378778811006566 (visited on 07/22/2020).

[73] Alessio Mastrucci et al. "A GIS-BASED APPROACH TO ESTIMATE ENERGY SAVINGS AND INDOOR THERMAL COMFORT FOR URBAN HOUSING STOCK RETROFITTING". en. In: (), p. 8.

[74] Amir-Hamed Mohsenian-Rad and Alberto Leon-Garcia. "Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments". In: *IEEE Transactions on Smart Grid* 1.2 (Sept. 2010), pp. 120–133. ISSN: 1949-3061. DOI: 10.1109/TSG.2010.2055903.

[75] Moret, Stefano, Bierlaire, Michel, and Marechal, Francois. "Strategic Energy Planning under Uncertainty: a Mixed-Integer Linear Programming Modeling Framework for Large-Scale Energy Systems". en. In: *Computer Aided Chemical Engineering* 38 (Jan. 2016), pp. 1899–1904. ISSN: 1570-7946. DOI: 10.1016/B978-0-444-63428-3.50321-0. URL: https://ezproxy.ucentral.edu.co:2154/science/article/pii/B9780444634283503210 (visited on 01/15/2020).

[76] Jose Antonio Moscoso-López et al. "Hourly Air Quality Index (AQI) Forecasting Using Machine Learning Methods". en. In: *15th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2020)*. Ed. by Álvaro Herrero et al. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2021, pp. 123–132. ISBN: 978-3-030-57802-2. DOI: 10.1007/978-3-030-57802-2_12.

[77] Martín Mosteiro-Romero, Jimeno A. Fonseca, and Arno Schlueter. "Seasonal effects of input parameters in urban-scale building energy simulation". en. In: *Energy Procedia*. CISBAT 2017 International ConferenceFuture Buildings & Districts – Energy Efficiency from Nano to Urban Scale 122 (Sept. 2017), pp. 433–438. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2017.07.459. URL: http://www.sciencedirect.com/science/article/pii/S1876610217334161 (visited on 06/02/2020).

[78] Martín Mosteiro-Romero et al. "An Integrated Microclimate-Energy Demand Simulation Method for the Assessment of Urban Districts". English. In: *Frontiers in Built Environment* 6 (2020). Publisher: Frontiers. ISSN: 2297-3362. DOI: 10.3389/fbuil.2020.553946. URL: https://www.frontiersin.org/articles/10.3389/fbuil.2020.553946/full (visited on 11/20/2020).

[79] Payam Nejat et al. "A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries)". en. In: *Renewable and Sustainable Energy Reviews* 43 (Mar. 2015), pp. 843–862. ISSN: 1364-0321. DOI: 10.1016/j.rser.2014.11.066. URL: http://www.sciencedirect.com/science/article/pii/S1364032114010053 (visited on 06/02/2020).

[80] Sarika Nyaramneni, Md Abdul Saifulla, and Shaik Mahboob Shareef. "ARIMA for Traffic Load Prediction in Software Defined Networks". en. In: *Evolutionary Computing and Mobile Sustainable Networks*. Ed. by V. Suma, Noureddine Bouhmala, and Haoxiang Wang. Lecture Notes on Data Engineering and Communications Technologies. Singapore: Springer, 2021, pp. 815–824. ISBN: 9789811552588. DOI: 10.1007/978-981-15-5258-8_75.

[81] Sarika Nyaramneni, Md Abdul Saifulla, and Shaik Mahboob Shareef. "ARIMA for Traffic Load Prediction in Software Defined Networks". en. In: *Evolutionary Computing and Mobile Sustainable Networks*. Ed. by V. Suma, Noured-

dine Bouhmala, and Haoxiang Wang. Lecture Notes on Data Engineering and Communications Technologies. Singapore: Springer, 2021, pp. 815–824. ISBN: 9789811552588. DOI: 10.1007/978-981-15-5258-8_75.

[82] OECD. *Transition to Sustainable Buildings: Strategies and Opportunities to 2050*. en. Text. Library Catalog: www.oecd-ilibrary.org. URL: https://www.oecd-ilibrary.org/energy/transition-to-sustainable-buildings_9789264202955-en (visited on 07/22/2020).

[83] A. S. O. Ogunjuyigbe, T. R. Ayodele, and O. A. Akinola. "User satisfaction-induced demand side load management in residential buildings with user budget constraint". en. In: *Applied Energy* 187 (Feb. 2017), pp. 352–366. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2016.11.071. URL: http://www.sciencedirect.com/science/article/pii/S0306261916317020 (visited on 06/02/2020).

[84] Vítor Oliveira. *Urban Morphology: An Introduction to the Study of the Physical Form of Cities*. en. Google-Books-ID: WlveCwAAQBAJ. Springer, Mar. 2016. ISBN: 978-3-319-32083-0.

[85] Jukka V. Paatero and Peter D. Lund. "A model for generating household electricity load profiles". en. In: *International Journal of Energy Research* 30.5 (2006), pp. 273–290. ISSN: 1099-114X. DOI: 10.1002/er.1136. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/er.1136 (visited on 06/02/2020).

[86] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[87] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[88] Trine Dyrstad Pettersen. "Variation of energy consumption in dwellings due to climate, building and inhabitants". en. In: *Energy and Buildings* 21.3 (Jan. 1994), pp. 209–218. ISSN: 0378-7788. DOI: 10.1016/0378-7788(94)90036-1. URL: http://www.sciencedirect.com/science/article/pii/0378778894900361 (visited on 07/22/2020).

[89] Ethan M. Pickering et al. "Building electricity consumption: Data analytics of building operations with classical time series decomposition and case based subsetting". In: *Energy and Buildings* 177 (Oct. 2018), pp. 184–196. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2018.07.056. URL: http://www.sciencedirect.com/science/article/pii/S0378778818306029 (visited on 02/01/2019).

[90] Dirk Pietruschka and Romain Nouvel. "CityGML-based 3D City Model for Energy Diagnostics and Urban Energy Policy Support". en. In: (). URL: https://www.academia.edu/26128958/CityGML_based_3D_City_Model_for_Energy_Diagnostics_and_Urban_Energy_Policy_Support (visited on 12/09/2020).

[91] Luis Pérez-Lombard, José Ortiz, and Christine Pout. "A review on buildings energy consumption information". In: *Energy and Buildings* 40.3 (Jan. 2008), pp. 394–398. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2007.03.007. URL: http://www.sciencedirect.com/science/article/pii/S0378778807001016 (visited on 01/29/2019).

[92] Steven Quan et al. "Urban Data and Building Energy Modeling: A GIS-Based Urban Building Energy Modeling System Using the Urban-EPC Engine". In: *Lecture Notes in Geoinformation and Cartography*. Vol. 213. Journal Abbreviation: Lecture Notes in Geoinformation and Cartography. May 2015, pp. 447–469. DOI: 10.1007/978-3-319-18368-8_24.

[93] Aowabin Rahman, Vivek Srikumar, and Amanda D. Smith. "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks". In: *Applied Energy* 212 (Feb. 2018), pp. 372–385. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2017.12.051. URL: http://www.sciencedirect.com/science/article/pii/S0306261917317658 (visited on 02/17/2018).

[94] T. Ramesh, Ravi Prakash, and K.K. Shukla. "Life cycle energy analysis of buildings: An overview". In: *Energy and Buildings* 42 (Oct. 2010), pp. 1592–1600. DOI: 10.1016/j.enbuild.2010.05.007.

[95] C.F. Reinhart and Davila Cerezo. "Urban building energy modeling - A review of a nascent field". In: *Building and Environment* 97 (2016), pp. 196–202. DOI: 10.1016/j.buildenv.2015.12.001.

[96] Ian Richardson et al. "Domestic electricity use: A high-resolution energy demand model". en. In: *Energy and Buildings* 42.10 (Oct. 2010), pp. 1878–1887. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2010.05.023. URL: http://www.sciencedirect.com/science/article/pii/S0378778810001854 (visited on 06/02/2020).

[97] J. V. Ringwood, D. Bofelli, and F. T. Murray. "Forecasting Electricity Demand on Short, Medium and Long Time Scales Using Neural Networks". en. In: *Journal of Intelligent and Robotic Systems* 31.1 (May 2001), pp. 129–147. ISSN: 1573-0409. DOI: 10.1023/A:1012046824237. URL: https://doi.org/10.1023/A:1012046824237 (visited on 01/29/2019).

[98] Frank Rosenblatt. *Principles of neurodynamics. Perceptrons and the theory of brain mechanisms.* en. Tech. rep. VG-1196-G-8. CORNELL AERONAUTICAL LAB INC BUFFALO NY, Mar. 1961. URL: https://apps.dtic.mil/docs/citations/AD0256582 (visited on 01/16/2020).

[99] Luis Ruiz et al. "An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings". en. In: *Energies* 9.9 (Aug. 2016), p. 684. DOI: 10.3390/en9090684. URL: http://www.mdpi.com/1996-1073/9/9/684 (visited on 08/23/2018).

[100] D. J. Sailor. "A green roof model for building energy simulation programs". en. In: *Energy and Buildings* 40.8 (Jan. 2008), pp. 1466–1478. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2008.02.001. URL: http://www.sciencedirect.com/science/article/pii/S0378778808000339 (visited on 06/02/2020).

[101] Abdelhay A. Sallam and Om P. Malik. *Electric Distribution Systems*. en. John Wiley & Sons, Nov. 2018. ISBN: 978-1-119-50931-8.

[102] M Santamouris et al. "On the impact of urban climate on the energy consumption of buildings". In: *Solar Energy*. Urban Environment 70.3 (Jan. 2001), pp. 201–216. ISSN: 0038-092X. DOI: 10.1016/S0038-092X(00)00095-5. URL: http://www.sciencedirect.com/science/article/pii/S0038092X00000955 (visited on 01/29/2019).

[103] M Santamouris et al. "On the impact of urban climate on the energy consumption of buildings". en. In: *Solar Energy*. Urban Environment 70.3 (Jan. 2001), pp. 201–216. ISSN: 0038-092X. DOI: 10.1016/S0038-092X(00)00095-5. URL: http://www.sciencedirect.com/science/article/pii/S0038092X00000955 (visited on 06/02/2020).

[104] R. Schoetter et al. "Parametrisation of the variety of human behaviour related to building energy consumption in the Town Energy Balance (SURFEX-TEB v. 8.2)". en. In: *Geoscientific Model Development* 10 (July 2017), pp. 2801–2831. ISSN: 1991-959X, 1991-9603. URL: https://doaj.org (visited on 08/22/2018).

[105] Ad J. Seebregts, Gary A. Goldstein, and Koen Smekens. "Energy/Environmental Modeling with the MARKAL Family of Models". en. In: *Operations Research Proceedings 2001*. Ed. by Peter Chamoni et al. Operations Research Proceedings 2001. Berlin, Heidelberg: Springer, 2002, pp. 75–82. ISBN: 978-3-642-50282-8. DOI: 10.1007/978-3-642-50282-8_10.

[106] Praveen Sehrawat and K. Kensek. "URBAN ENERGY MODELING: GIS AS AN ALTERNATIVE TO BIM". en. In: *undefined*. 2014. URL: /paper/URBAN-ENERGY-MODELING%3A-GIS-AS-AN-ALTERNATIVE-TO-BIM-Sehrawat-Kensek/d284745bccc5c0fc498183a34b45ea1c8dc746ce (visited on 12/09/2020).

[107] Senatro Di Leo et al. "Energy systems modelling to support key strategic decisions in energy and climate change at regional scale". en. In: *Renewable and Sustainable Energy Reviews* 42 (Feb. 2015), pp. 394–414. ISSN: 1364-0321. DOI: 10.1016/j.rser.2014.10.031. URL: https://ezproxy.ucentral.edu.co:2154/science/article/pii/S136403211400848X (visited on 01/15/2020).

[108] Xavier Serrano Guerrero et al. "A Time-Series Treatment Method to Obtain Electrical Consumption Patterns for Anomalies Detection Improvement in Electrical Consumption Profiles". In: *Energies* 13 (Feb. 2020), p. 1046. DOI: 10.3390/en13051046.

[109] Xavier Serrano-Guerrero et al. "Forecasting Building Electric Consumption Patterns Through Statistical Methods". en. In: *Advances in Emerging Trends and Technologies*. Ed. by Miguel Botto-Tobar et al. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2020, pp. 164–175. ISBN: 978-3-030-32033-1. DOI: 10.1007/978-3-030-32033-1_16.

[110] A. K. Shackelford, C. H. Davis, and Xiangyun Wang. "Automated 2-D building footprint extraction from high-resolution satellite multispectral imagery". In: *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*. Vol. 3. Sept. 2004, 1996–1999 vol.3. DOI: 10.1109/IGARSS.2004.1370739.

[111] Muhammad Shahbaz et al. "Industrialization, electricity consumption and CO2 emissions in Bangladesh". en. In: *Renewable and Sustainable Energy Reviews* 31 (Mar. 2014), pp. 575–586. ISSN: 1364-0321. DOI: 10.1016/j.rser.2013.12.028. URL: http://www.sciencedirect.com/science/article/pii/S1364032113008423 (visited on 12/03/2020).

[112] Shanhong Liu. *Number of connected devices per person*. en. URL: https://www.statista.com/statistics/678739/forecast-on-connected-devices-per-person/ (visited on 06/02/2020).

[113] Meng Shen et al. "Prediction of household electricity consumption and effectiveness of concerted intervention strategies based on occupant behaviour and personality traits". en. In: *Renewable and Sustainable Energy Reviews* 127 (July 2020), p. 109839. ISSN: 1364-0321. DOI: 10.1016/j.rser.2020.109839. URL: http://www.sciencedirect.com/science/article/pii/S1364032120301337 (visited on 06/02/2020).

[114] Brian L Smith, Billy M Williams, and R Keith Oswald. "Comparison of parametric and nonparametric models for traffic flow forecasting". en. In: *Transportation Research Part C: Emerging Technologies* 10.4 (Aug. 2002), pp. 303–321. ISSN: 0968-090X. DOI: 10.1016/S0968-090X(02)00009-8. URL: http://www.sciencedirect.com/science/article/pii/S0968090X02000098 (visited on 12/09/2020).

[115] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. "Sequence to Sequence Learning with Neural Networks". In: *Advances in Neural Information Processing Systems 27*. Ed. by Z. Ghahramani et al. Curran Associates, Inc., 2014, pp. 3104–3112. URL: http://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks.pdf (visited on 01/30/2019).

[116] Mikhail S. Tarkov and Ivan A. Chernov. "Time Series Prediction by Reservoir Neural Networks". en. In: *Advances in Neural Computation, Machine Learning, and Cognitive Research IV*. Ed. by Boris Kryzhanovsky et al. Studies in Computational Intelligence. Cham: Springer International Publishing, 2021, pp. 303–308. ISBN: 978-3-030-60577-3. DOI: 10.1007/978-3-030-60577-3_36.

[117] Gregorio Fidalgo Valverde et al. "Copper Price Time Series Forecasting by Means of Generalized Regression Neural Networks with Optimized Predictor Variables". en. In: *15th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2020)*. Ed. by Álvaro Herrero et al. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2021, pp. 681–690. ISBN: 978-3-030-57802-2. DOI: 10.1007/978-3-030-57802-2_65.

[118] Khairul Wagiman et al. "Lighting system control techniques in commercial buildings: Current trends and future directions". In: *Journal of Building Engineering* (Mar. 2020), p. 101342. DOI: 10.1016/j.jobe.2020.101342.

[119] Z. Wang and Y. Lou. "Hydrological time series forecast model based on wavelet de-noising and ARIMA-LSTM". In: *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. Mar. 2019, pp. 1697–1701. DOI: 10.1109/ITNEC.2019.8729441.

[120] Zhiqiang Wang et al. "Modeling of regional electrical heating load characteristics considering user response behavior difference". en. In: *International Journal of Electrical Power & Energy Systems* 123 (Dec. 2020), p. 106297. ISSN: 0142-0615. DOI: 10.1016/j.ijepes.2020.106297. URL: http://www.sciencedirect.com/science/article/pii/S0142061519343224 (visited on 07/22/2020).

[121] Parag Wate and Volker Coors. "3D Data Models for Urban Energy Simulation". en. In: *Energy Procedia*. 6th International Building Physics Conference, IBPC 2015 78 (Nov. 2015), pp. 3372–3377. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2015.11.753. URL: http://www.sciencedirect.com/science/article/pii/S1876610215024856 (visited on 12/09/2020).

[122] C. Wei and Y. Li. "Design of energy consumption monitoring and energy-saving management system of intelligent building based on the Internet of things". In: *2011 International Conference on Electronics, Communications*

and Control (ICECC). Sept. 2011, pp. 3650–3652. DOI: 10.1109/ICECC.2011.6066758.

[123] Joakim Widén and Ewa Wäckelgård. "A high-resolution stochastic model of domestic activity patterns and electricity demand". en. In: *Applied Energy* 87.6 (June 2010), pp. 1880–1892. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2009.11.006. URL: http://www.sciencedirect.com/science/article/pii/S0306261909004930 (visited on 06/02/2020).

[124] Tatsuhiro Yamamoto et al. "Fundamental study of coupling methods between energy simulation and CFD". en. In: *Energy and Buildings* 159 (Jan. 2018), pp. 587–599. ISSN: 03787788. DOI: 10.1016/j.enbuild.2017.11.059. URL: https://linkinghub.elsevier.com/retrieve/pii/S0378778817318868 (visited on 01/29/2019).

[125] Jihui Yuan et al. "Predictive artificial neural network models to forecast the seasonal hourly electricity consumption for a University Campus". en. In: *Sustainable Cities and Society* 42 (Oct. 2018), pp. 82–92. ISSN: 2210-6707. DOI: 10.1016/j.scs.2018.06.019. URL: http://www.sciencedirect.com/science/article/pii/S2210670718301203 (visited on 07/22/2020).

[126] Aston Zhang et al. *Dive into Deep Learning*. 2020. URL: https://d2l.ai.

[127] G. Peter Zhang. "Time series forecasting using a hybrid ARIMA and neural network model". en. In: *Neurocomputing* 50 (Jan. 2003), pp. 159–175. ISSN: 0925-2312. DOI: 10.1016/S0925-2312(01)00702-0. URL: http://www.sciencedirect.com/science/article/pii/S0925231201007020 (visited on 12/09/2020).

[128] Yinping Zhang et al. "Application of latent heat thermal energy storage in buildings: State-of-the-art and outlook". en. In: *Building and Environment* 42.6 (June 2007), pp. 2197–2209. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2006.07.023. URL: http://www.sciencedirect.com/science/article/pii/S0360132306002058 (visited on 06/02/2020).

[129] J. Zheng and M. Huang. "Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning". In: *IEEE Access* 8 (2020). Conference Name: IEEE Access, pp. 82562–82570. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.2990738.

[130] Bin Zhou et al. "Multi-energy net load forecasting for integrated local energy systems with heterogeneous prosumers". en. In: *International Journal of Electrical Power & Energy Systems* 126 (Mar. 2021), p. 106542. ISSN: 0142-0615. DOI: 10.1016/j.ijepes.2020.106542. URL: http://www.sciencedirect.com/science/article/pii/S0142061520322626 (visited on 11/09/2020).