

Electrical and Computer Engineering

Deep Learning Techniques in Load Forecasting

A proposal in partial fulfillment of the MScE

|  |  |  |
| --- | --- | --- |
|  |  | Version <3.1>  Created: 2020 Dec-14  Updated: 2021-May-08 |

|  |  |  |
| --- | --- | --- |
|  | Supervised By: | Tolulope Olugbenga  Dr Dawn MacIsaac, PhD  Dr Julian Cardenas, PhD |

Table of Contents

[1 Load Forecasting Overview 1](#_Toc71471468)

[2 Investigation 4](#_Toc71471469)

[2.1 The Benchmark Algorithms 4](#_Toc71471470)

[2.1-a Seasonal Naïve Forecaster (SNF) 5](#_Toc71471471)

[2.1-b Auto-Regressive Integrated Moving Average with Exogenous Variables (ARIMAX) 6](#_Toc71471472)

[2.1-c Multiple Linear Regression 7](#_Toc71471473)

[2.1-d Artificial Neural Network Short Term Load Forecaster (ANNSTLF) – Generation Three 7](#_Toc71471474)

[2.2 Convolutional Neural Networks (CNN) 9](#_Toc71471475)

[2.3 Long Short-Term Memory (LSTM) 9](#_Toc71471476)

[2.4 Data Sets and Metrics for Evaluation 10](#_Toc71471477)

[3 Contributions 11](#_Toc71471478)

[4 References 13](#_Toc71471479)

Table of Figures

[Figure 1:- The Block Diagram of the third generation ANNSTLF [25] 8](#_Toc70354493)

<Title>

Updated: 2021-May-08 by <Author>

# Load Forecasting Overview

Load forecasting is an integral part of the planning and operation of electric utilities; it has played a vital role in the power industry for over a century. However, load forecasting can also be helpful to organizations other than electric utilities, such as load aggregators, power marketers, independent system operators, regulatory commissions, and even industrial/commercial companies, banks, trading firms, and insurance companies [1][2]. These organizations use load forecasting in power systems planning/operations, revenue projection, rate design, energy trading, and other activities [3]–[5]. For example, to have a stable supply of electricity, reserve power must be prepared beforehand to serve consumers in the future (e.g., in case of high demand or failure in the current grid supply).

Electric load forecasting is well studied [1], [6]–[8], and most current research is focused on developing more accurate forecasts. Load forecasting is particularly relevant in today’s context, with the advent of new smart grid technologies. The demand patterns used to drive these technologies are complex due to the deregulation of energy markets, and the number of different random variables, often governed by human behavior, which need to be considered to predict future electricity demand. Finding an appropriate forecasting model for a specific electricity network is not a trivial task [4][5][9]. Different factors can affect load forecasts, such as the location of the area, the type of customers in the region, weather factors (e.g., temperature), a trend in the time series dataset, the time of the day, day of the week, and other unpredictable factors (i.e., coronavirus outbreak). Also, electricity demand can be assessed by tracking it periodically - hourly, daily, weekly, monthly, or yearly and forecasting processes can be applied to various horizons: very short-term load forecasting (VSTLF, <1-day), short-term load forecasting (STLF, <2-weeks), medium-term load forecasting (MTLF <3-years), and long-term load forecasting (LTLF >3years) [10]. Shorter-term forecasting has been the focus in most current research, concentrating on horizons of less than two weeks [11][10][1].

Both statistical techniques and artificial intelligence (AI) have been applied to provide load forecasts, and with the advent of ubiquitous application of data science, the boundary between these two approaches is becoming more equivocal [1]. We aim to create an algorithm(s) that could quickly adapt to sudden changes in the demand profile and quickly learn and interpret the complex relationships in the dataset without explicit specification from the user. Examples of statistical techniques applied to electrical load forecasting include multiple linear regression analysis [12][13], exponential smoothing [14], [15], and auto-regressive integrated moving average (ARIMA) modeling [16], [17]. AI algorithms are more intelligent and can be better, as they provide the capacity to learn and adapt to the non-linear and complex relationships between load and other influencing factors (e.g. weather, time of day) automatically [18]. Artificial Neural Networks [19][20], Fuzzy Regression Models [21], [22], Support Vector Machines [23], Gradient Boosting Machines [24] have all been applied to electrical load forecasting.

In recent years, deep learning approaches have also become enticing to researchers in this field. The Recurrent Neural Network (or RNN) introduced the concept of memory into neural networks, which helps model sequential data. However, RNNs have a weakness in that they are susceptible to the effects of a vanishing or exploding gradient [25][6]. This weakness led to the development of the Long Short-Term Memory (LSTM) network. The LSTM is an RNN that overcomes a vanishing gradient by providing a model capable of storing information for an extended period. Munem[26] et al. argue that LSTM is better than other deep neural networks because of its memory cell configuration. Convolutional Neural Networks (CNNs) have also gained the attention of researchers studying load forecasting [ref]. The CNN is a feed-forward network that is created after the human neurons in the way it is structured. CNNs has typically been used for processing data with grid topology; its primary application has been for image classification [27][3]. []CNNs are known to boost the power of the ANN as it has deeper layers and can interpret load data better [6], [28]. Amaradinghe[3] et al. compared the CNN with the LSTM, SVM, ANN, and other algorithms for individual building level load forecasting. Their results and observations concluded that CNN is a viable technique that produces accurate load forecasts.

It is not likely that one approach will be helpful in all load forecasting scenarios. Tao Hong spoke about the myth of finding the best technique [1]. He concluded that it is essential that researchers and users know that a universally best technique does not exist. The approach applied to load forecast should be based on forecasting needs and the dataset being analyzed. Different algorithms perform better or worse with different datasets. Forecast errors differ significantly for different utilities, utility zones, different horizons, etc.

The goal of this work is to determine whether or not deep learning approaches (e.g., LSTM and CNN) can improve forecasting accuracy for particular data sets by comparing the accuracy of deep learning forecasters to some of the current forecasters used by utilities. [state the characteristics of the data sets of interest, and the horizon of interest]. Four benchmark forecasters will be uses for comparison: a Seasonal Naïve forecaster, a Multiple Linear Regression (MLR) forecaster, an Auto-Regressive Integrated Moving Average (ARIMAX) forecaster, and a forecaster based on a shallow Artificial Neural Network (ANN). These benchmark algorithms have been available for many years and have been implemented and used by researchers and utilities [29][4][30].

# Investigation

Many papers lack detailed information about their experimental set-ups, making it challenging to conduct direct comparisons with the results they report [1]. To make our work reproducible, two of the datasets we are using are from an independent system operator, and the selected benchmark algorithms have much documentation about how their implementation could be done. Over the previous years, we have noticed a rising trend in the power demand at most utilities due to new systems and more sophisticated equipment been added. Therefore, it is essential to create algorithms that could adapt quickly to these changes as they occur [31].

We will start with the CNN and LSTM [32] algorithms, as these are two of the most popular ones. This comparison will help us figure out how much extra value deep learning adds to the forecasts compared to the benchmarks. First, we begin by implementing all the algorithms, then we compare their performances on three datasets and make observations of when one performs better or worse. We can then make improvements based on the new information we find.

## The Benchmark Algorithms

Four algorithms were selected for benchmark comparison: seasonal naïve forecaster (SNF), the autoregressive integrated moving average with exogenous variables (ARIMAX), multiple linear regression (MLR), and the artificial neural network short term load forecaster technique (ANNSTLF). All of these benchmark algorithms have been used for many years by researchers and utilities [9][5][1][33].

### Seasonal Naïve Forecaster (SNF)

The naïve forecaster is the most cost-effective forecasting model; it has often been implemented as a benchmark for developing much more sophisticated models [34][35]. Let us think about it; “The simplest way to predict the next value in a time series is to assume it is going to have the same values as the current value.” This assumption holds reasonably well for load forecasting, and this forms the basis of the naïve forecaster. Of course, the naïve forecaster is susceptible to significant errors when there are trends in the data. When there is seasonality in the time series data, the seasonal naïve forecaster is preferable because forecasts will be equal to the value from the last season (e.g., a week ago). SNF is most useful when there is a very high level of seasonality in the dataset [36].

The naïve forecaster is usually used as a baseline for other methods because it gives us an understanding of how much value is added to the current forecasting process. When the naïve forecaster performs better or similar to a more sophisticated technique, this tells us that the technique might not be a viable option. The formula for SNF can be seen below;



Where;  is the time series and is the seasonal period (m=1 for naïve forecaster without seasonality?). The naive formula takes the last observed value as the future value, while the seasonal naive formula takes the value from the previous season.

### Auto-Regressive Integrated Moving Average with Exogenous Variables (ARIMAX)

ARIMA is a statistical technique that describes a given time series distribution based on its past values (its lags and the lagged forecast error); the final equation is then used to forecast future values [37]. The formula of the ARIMA can be seen below;



Where  is the lag1 of the time series,  is the coefficient of lag1 estimated by the model,  is the intercept that the model has estimated,  are the error terms from respective lags. In its basic form, ARIMA’s forecast  is the sum of a constant, the linear combination lags of (up to p lags), and the linear combination of lagged forecast errors (up to q lags). An ARIMA model is characterized by p, d, q, where p is the order of the AR term, q is the order of the MA term, and d is the number of differences required to make the time series stationary.

An ARIMA model is where the time series was differenced at least once to make it stationary and combine the AR and the MA terms [38]. Building an ARIMA model requires the time series to be stationary because the term “Auto-Regressive” in ARIMA means we are dealing with a linear regression model that uses its lags as predictors. Also, linear regression models work better in situations where the predictors are not correlated and independent. The Auto-Regressive order p refers to the number of lags of the data that are selected as predictors. At the same time, the Moving Average order q refers to the number of lagged forecast errors that go into the creation of the ARIMA Model [39]. When we want to consider exogenous variables (temperature, day of the week, etc.), the ARIMAX model would have to be used [40].

### Multiple Linear Regression

Multiple linear regression is one of the most commonly used statistical techniques for load forecasting [41] [13]. The idea of MLR is to model the relationships between a continuous dependent variable (electricity demand) and one or more independent variables (i.e., temperature, the hour of the day, etc.) A common misunderstanding is that MLR models cannot model the non-linear relationships between the electrical load and weather variables, which turns out to be false [1] [42]. For example, polynomial regression models can describe non-linear relationships between dependent and independent variables using polynomials. The equation below shows an MLR with two independent variables:



Where  is the dependent variable, and  are the independent variables, ’s are parameters to be estimated and is the error. The error term  represents a set of random variables that are independent and identically distributed and have a mean of zero. MLR models are fitted such that the sum-of-squares of differences of actual and forecasted values are reduced. Although a large number of alternatives are currently available, linear regression models are still quite popular [43], [44][42][11][13].

### Artificial Neural Network Short Term Load Forecaster (ANNSTLF) – Generation Three

The ANNSTLF forecaster, is the best-known ANN implementation for STLF [1], [42][30]. The ANNSTLF model is built as a shallow multi-layer feed-forward Artificial Neural Network (ANN) identified by the creators in this paper [45]. ANN models are popular in use today due to their ability to learn complex and non-linear relationships in the data on their own – unlike MMLR models, the specification of independent variables explicitly in ANNs is not required. The ANNSTLF and its improvements of it have been implemented by several utilities in Canada and the US. The figure below shows the block diagram of the system:

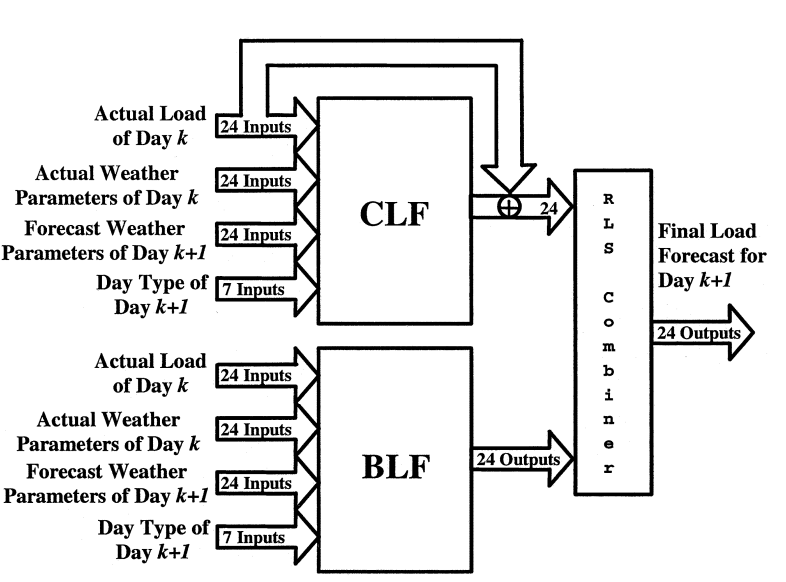


Figure :- The Block Diagram of the third generation ANNSTLF [45]

ANNSTLF is a multi-layer perceptron, and it is trained with the error back-propagation algorithm. The third generation of ANNSTLF has three models; a baseload forecaster (BLF), a change in load forecaster (CLF), and a recursive least squares (RLS) combiner. The two forecasters are created the same and given the same inputs; the difference is found in their output. The BLF is trained to forecast the regular load of the next day, while the CLF is trained to forecast the change in the hourly load between yesterday and today. The final CLF forecast is the addition of the change in load forecast and the actual load of yesterday. The RLS combiner takes the outputs from these forecasts and combines them adaptively using the recursive least squares algorithm. It is argued that the CLF forecaster allows the model to rapidly adapt to abrupt changes in temperature [42][46][47].

## Convolutional Neural Networks (CNN)

CNN shares some similarity with the ANN; it is a feed-forward neural network which mimics the human neurons in its design. CNN has been applied in image and audio processing, natural language processing, and video recognition [48] [6]. CNNs are usually used with image data, but time-series data can be arranged to mimic image data, and it can then be fed into a CNN [49][3]. CNNs usually process data with a grid topology; images are two-dimensional grids and time series data are one dimensional, making the time series conversion necessary. In this research study, the CNN architecture consists of six layers: the image input layer, the 2D convolution layer, the rectified linear unit activation layer (relu), the max-pooling 2D layer, and fully connected, and a regression layer.

Because the ANNSTLF structure was recognized as the best forecaster for short-term load forecasting [1], [42]. Our approach mimics the ANNSTLF structure by creating a Base Load Forecaster, Change in the Load Forecaster, and RLS combiner; while using the CNN algorithm in place of the ANN. The architecture will have the same inputs and structure as the ANNSTLF, but the BLF and CLF algorithms will be trained using CNNs.

## Long Short-Term Memory (LSTM)

The Recurrent Neural Networks (RNNs) is typically trained using either a Back-propagation or Real-Time Recurrent Learning algorithm. The issue is, training with these methods usually fails due to the vanishing gradient [32], [50][26]. This LSTM is an RNN created to fix vanishing gradient problems; it can store information for long periods. Its memory cell configuration helps retain information more than any other deep neural network currently available [26]. We also plan on trying out the LSTM algorithm in a similar fashion as the CNN using the ANNSTLF structure. The computational graph of the LSTM consists of five critical elements: 1) input gate, 2) forget gate, 3) output gate, 4) cell and 5) state output. The cell memory state is responsible for operations such as writing, reading, and erasing. The equations below give a mathematical representation of the model [32].













Where  represents the input of the input gate,  represents the input of the forget gate,  represents the output gate’s input,  represents the update signal,  represents the state value at a time  , and  represents the output of the LSTM cell. The input gate’s decision to use a sigmoid feature with an on/off state will change the memory state. There will be no improvement in the state cell memory  if the input gate value is minimal and close to zero. In the network model, stacked LSTM can be implemented by using multiple LSTM layers [32].

## Data Sets and Metrics for Evaluation

Three time-series datasets were selected , they are sampled on an hourly horizon, and they also contain the temperature variable as changes in temperature plays a huge role in load forecasting. The authors of the ANNSTLF algorithm mentioned that the algorithm works best when trained with at least three years of data. Two datasets were obtained from the Ontario independent system operator [51]; one was from Toronto, and the other was from Ottawa. Both datasets run from January 2010 until December 2019. The third dataset was obtained from Saint John Energy [52]; this dataset runs from January 2018 until December 2020. The weather variables were obtained from the Government of Canada’s website [53].

The selected global metrics are: Mean Absolute Percent Error (MAPE), Mean Error (ME), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Standard Deviation. The MAPE tells us the size of the error of the forecasted values in terms of percentage. MAPE is the most common measure used for load forecasting [1]; it works best when there are no near zeros or extreme values. ME stands for the average error of all forecasts over the entire horizon. The MAE is a calculation of the average magnitude of forecast errors without taking their direction into account. The RMSE indicates the model’s absolute fit or how similar the actual values are to the expected values. Standard deviation tells how spread the errors are, measuring how far each error is from the mean error.

Since the global metrics only show one value that was gotten from the entire dataset. There is the need to classify the errors on hourly, daily, and monthly horizons. This approach gives us a better picture of the distribution of errors for each hour, month, and day. It also helps us to identify situations where the algorithms perform better or worse.

# Contributions

Researchers will be able to compare the value added by deep learning algorithms (such as CNN and LTSM) to more traditional algorithms with the help of this research. We want to develop an algorithm (or a series of algorithms) that can easily adjust to annual increases in power demand, as well as sudden shifts in temperature and any other random variable that affects load demand. We also want to create an algorithm or algorithms capable of comprehending and interpreting complex data relationships without the need for explicit user feedback. Furthermore, this project will be a reproducible experiment that other researchers can use in the future. The main reasons for this are that two of our datasets come from an independent system operator, and the benchmark algorithms we will be working with are well-documented.

# References

[1] T. Hong and S. Fan, “Probabilistic electric load forecasting: A tutorial review,” *Int. J. Forecast.*, vol. 32, no. 3, pp. 914–938, 2016, doi: 10.1016/j.ijforecast.2015.11.011.

[2] S. Saurabh, H. Shoeb, and A. B. Mohammad, “Short Term Load Forecasting Using Artificial Neural Network,” pp. 159–163, 2017.

[3] K. Amarasinghe, D. L. Marino, and M. Manic, “Deep neural networks for energy load forecasting,” 2017, doi: 10.1109/ISIE.2017.8001465.

[4] J. Zhang, Y. M. Wei, D. Li, Z. Tan, and J. Zhou, “Short term electricity load forecasting using a hybrid model,” *Energy*, 2018, doi: 10.1016/j.energy.2018.06.012.

[5] C. Kuster, Y. Rezgui, and M. Mourshed, “Electrical load forecasting models: A critical systematic review,” *Sustainable Cities and Society*. 2017, doi: 10.1016/j.scs.2017.08.009.

[6] A. Almalaq and G. Edwards, “A review of deep learning methods applied on load forecasting,” *Proc. - 16th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2017*, vol. 2017-Decem, pp. 511–516, 2017, doi: 10.1109/ICMLA.2017.0-110.

[7] B. Yildiz, J. I. Bilbao, and A. B. Sproul, “A review and analysis of regression and machine learning models on commercial building electricity load forecasting,” *Renewable and Sustainable Energy Reviews*. 2017, doi: 10.1016/j.rser.2017.02.023.

[8] A. Baliyan, K. Gaurav, and S. Kumar Mishra, “A review of short term load forecasting using artificial neural network models,” 2015, doi: 10.1016/j.procs.2015.04.160.

[9] I. K. Nti, M. Teimeh, O. Nyarko-Boateng, and A. F. Adekoya, “Electricity load forecasting: a systematic review,” *J. Electr. Syst. Inf. Technol.*, 2020, doi: 10.1186/s43067-020-00021-8.

[10] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, and Z. Zhu, “Multi-scale convolutional neural network with time-cognition for multi-step short-Term load forecasting,” *IEEE Access*, vol. 7, pp. 88058–88071, 2019, doi: 10.1109/ACCESS.2019.2926137.

[11] T. Hong, J. Wilson, and J. Xie, “Long term probabilistic load forecasting and normalization with hourly information,” *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 456–462, 2014, doi: 10.1109/TSG.2013.2274373.

[12] S. Kumar, S. Mishra, and S. Gupta, “Short term load forecasting using ANN and multiple linear regression,” 2016, doi: 10.1109/CICT.2016.44.

[13] A. Y. Saber and A. K. M. R. Alam, “Short term load forecasting using multiple linear regression for big data,” *2017 IEEE Symp. Ser. Comput. Intell. SSCI 2017 - Proc.*, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/SSCI.2017.8285261.

[14] P. Ji, D. Xiong, P. Wang, and J. Chen, “A study on exponential smoothing model for load forecasting,” 2012, doi: 10.1109/APPEEC.2012.6307555.

[15] J. F. Rendon-Sanchez and L. M. de Menezes, “Structural combination of seasonal exponential smoothing forecasts applied to load forecasting,” *Eur. J. Oper. Res.*, 2019, doi: 10.1016/j.ejor.2018.12.013.

[16] L. Tang, Y. Yi, and Y. Peng, “An ensemble deep learning model for short-term load forecasting based on ARIMA and LSTM,” 2019, doi: 10.1109/SmartGridComm.2019.8909756.

[17] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji, “Electricity load forecasting using clustering and ARIMA model for energy management in buildings,” *Japan Archit. Rev.*, 2020, doi: 10.1002/2475-8876.12135.

[18] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, and Z. Zhu, “Multi-scale convolutional neural network with time-cognition for multi-step short-Term load forecasting,” *IEEE Access*, vol. 7, pp. 88058–88071, 2019, doi: 10.1109/ACCESS.2019.2926137.

[19] A. Badri, Z. Ameli, and A. Motie Birjandi, “Application of artificial neural networks and fuzzy logic methods for short term load forecasting,” 2012, doi: 10.1016/j.egypro.2011.12.965.

[20] P. H. Kuo and C. J. Huang, “A high precision artificial neural networks model for short-Term energy load forecasting,” *Energies*, 2018, doi: 10.3390/en11010213.

[21] T. Hong and P. Wang, “Fuzzy interaction regression for short term load forecasting,” *Fuzzy Optim. Decis. Mak.*, 2014, doi: 10.1007/s10700-013-9166-9.

[22] M. Hanmandlu and B. K. Chauhan, “Load forecasting using hybrid models,” *IEEE Trans. Power Syst.*, 2011, doi: 10.1109/TPWRS.2010.2048585.

[23] A. Yang, W. Li, and X. Yang, “Short-term electricity load forecasting based on feature selection and Least Squares Support Vector Machines,” *Knowledge-Based Syst.*, 2019, doi: 10.1016/j.knosys.2018.08.027.

[24] V. Mayrink and H. S. Hippert, “A hybrid method using Exponential Smoothing and Gradient Boosting for electrical short-term load forecasting,” 2017, doi: 10.1109/LA-CCI.2016.7885697.

[25] D. L. Marino, K. Amarasinghe, and M. Manic, “Building energy load forecasting using Deep Neural Networks,” *IECON Proc. (Industrial Electron. Conf.*, pp. 7046–7051, 2016, doi: 10.1109/IECON.2016.7793413.

[26] M. Munem, T. M. Rubaith Bashar, M. H. Roni, M. Shahriar, T. B. Shawkat, and H. Rahaman, “Electric power load forecasting based on multivariate LSTM neural network using bayesian optimization,” *2020 IEEE Electr. Power Energy Conf. EPEC 2020*, vol. 3, 2020, doi: 10.1109/EPEC48502.2020.9320123.

[27] G. H. Yann LeCun, Yoshua Bengio, “Deep learning (2015), Y. LeCun, Y. Bengio and G. Hinton,” *Nature*, 2015.

[28] A. Gasparin, S. Lukovic, and C. Alippi, “Deep Learning for Time Series Forecasting: The Electric Load Case,” 2019, [Online]. Available: http://arxiv.org/abs/1907.09207.

[29] T. Hong, P. Wang, and H. L. Willis, “A naïve multiple linear regression benchmark for short term load forecasting,” 2011, doi: 10.1109/PES.2011.6038881.

[30] K. Methaprayoon, W. J. Lee, S. Rasmiddatta, J. R. Liao, and R. J. Ross, “Multistage artificial neural network short-term load forecasting engine with front-end weather forecast,” *IEEE Trans. Ind. Appl.*, 2007, doi: 10.1109/TIA.2007.908190.

[31] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, and Z. Zhu, “Multi-Scale Convolutional Neural Network With Time-Cognition for Multi-Step Short-Term Load Forecasting,” *IEEE Access*, vol. 7, pp. 88058–88071, 2019, doi: 10.1109/ACCESS.2019.2926137.

[32] A. Almalaq and G. Edwards, “A review of deep learning methods applied on load forecasting,” 2017, doi: 10.1109/ICMLA.2017.0-110.

[33] A. K. Singh, Ibraheem, S. Khatoon, M. Muazzam, and D. K. Chaturvedi, “Load forecasting techniques and methodologies: A review,” 2012, doi: 10.1109/ICPCES.2012.6508132.

[34] P. Wang, B. Liu, and T. Hong, “Electric load forecasting with recency effect: A big data approach,” *Int. J. Forecast.*, 2016, doi: 10.1016/j.ijforecast.2015.09.006.

[35] A. Bracale, G. Carpinelli, P. De Falco, and T. Hong, “Short-term industrial load forecasting: A case study in an Italian factory,” 2017, doi: 10.1109/ISGTEurope.2017.8260176.

[36] Da Liu, K. Sun, H. Huang, and P. Tang, “Monthly load forecasting based on economic data by decomposition integration theory,” *Sustain.*, 2018, doi: 10.3390/su10093282.

[37] A. Shadkam, “USING SARIMAX TO FORECAST ELECTRICITY DEMAND AND CONSUMPTION,” 2020.

[38] R. Weron, *Modeling and ForecastingElectricity Loads and Prices*. 2013.

[39] R. Bonetto and M. Rossi, “Parallel multi-step ahead power demand forecasting through NAR neural networks,” *2016 IEEE Int. Conf. Smart Grid Commun. SmartGridComm 2016*, pp. 314–319, 2016, doi: 10.1109/SmartGridComm.2016.7778780.

[40] K. Goswami, A. Ganguly, and A. K. Sil, “Day ahead forecasting and peak load management using multivariate auto regression technique,” *Proc. 2018 IEEE Appl. Signal Process. Conf. ASPCON 2018*, no. 1, pp. 279–282, 2018, doi: 10.1109/ASPCON.2018.8748661.

[41] T. Hong, M. Gui, M. E. Baran, and H. L. Willis, “Modeling and forecasting hourly electric load by multiple linear regression with interactions,” *IEEE PES Gen. Meet. PES 2010*, pp. 1–8, 2010, doi: 10.1109/PES.2010.5589959.

[42] R. Weron, *Modeling and forecasting electricity loads and prices: A statistical approach*. wiley, 2006.

[43] X. Sun, Z. Ouyang, and D. Yue, “Short-term load forecasting based on multivariate linear regression,” 2017, doi: 10.1109/EI2.2017.8245401.

[44] G. Dudek, “Pattern-based local linear regression models for short-term load forecasting,” *Electr. Power Syst. Res.*, 2016, doi: 10.1016/j.epsr.2015.09.001.

[45] A. Khotanzad and R. Afkhami-Rohani, “ANNSTLF - Artificial neural network short-term load forecaster - generation three,” *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1413–1422, 1998, doi: 10.1109/59.736285.

[46] A. Khotanzad, E. Zhou, and H. Elragal, “A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment,” *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1273–1282, Nov. 2002, doi: 10.1109/TPWRS.2002.804999.

[47] P. R. J. Campbell and K. Adamson, “Methodologies for load forecasting,” 2006, doi: 10.1109/IS.2006.348523.

[48] M. Imani and H. Ghassemian, “Sequence to Image Transform Based Convolutional Neural Network for Load Forecasting,” 2019, doi: 10.1109/IranianCEE.2019.8786456.

[49] N. Singh, C. Vyjayanthi, and C. Modi, “Multi-step Short-term Electric Load Forecasting using 2D Convolutional Neural Networks,” 2020, doi: 10.1109/HYDCON48903.2020.9242917.

[50] D. L. Marino, K. Amarasinghe, and M. Manic, “Building energy load forecasting using Deep Neural Networks,” *IECON Proc. (Industrial Electron. Conf.*, pp. 7046–7051, 2016, doi: 10.1109/IECON.2016.7793413.

[51] “Independent Electricity System Operator Demand - Toronto, ON.” http://reports.ieso.ca/public/DemandZonal/ (accessed Jan. 05, 2021).

[52] “Saint John Energy.” https://www.sjenergy.com/ (accessed Jan. 05, 2021).

[53] “Historical Climate Data - Climate - Environment and Climate Change Canada.” https://climate.weather.gc.ca/ (accessed May 04, 2021).