

Electrical and Computer Engineering

Convolutional Neural Networks in Load Forecasting

A proposal in partial fulfilment of the MScE

|  |  |  |
| --- | --- | --- |
|  |  | Version <3.0>  Created: 2020 Dec-14  Updated: 2021-Apr-29 |

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

Table of Contents

[1 Load Forecasting Overview 1](#_Toc69486050)

[1.1 Load Forecasting Techniques 2](#_Toc69486051)

[1.1-a Statistical Techniques 2](#_Toc69486052)

[1.1-b Artificial Intelligence Techniques 3](#_Toc69486053)

[1.1-c Finding the One Size Fits All Technique 3](#_Toc69486054)

[1.2 Research Focus 3](#_Toc69486060)

[2 Investigation 4](#_Toc69486061)

[2.1 The Benchmark Algorithms 4](#_Toc69486062)

[2.1-a Seasonal Naïve Approach 5](#_Toc69486064)

[2.1-b Auto-Regressive Integrated Moving Average with Exogenous Variables (ARIMAX) 6](#_Toc69486065)

[2.1-c Multiple Linear Regression 7](#_Toc69486066)

[2.1-d Artificial Neural Network Short Term Load Forecaster (ANNSTLF) – Generation Three 7](#_Toc69486067)

[2.2 Convolutional Neural Network (CNN) 9](#_Toc69486068)

[2.3 Data Sets and Metrics for Evaluation 9](#_Toc69486070)

[3 Contributions 10](#_Toc69486071)

[4 References 11](#_Toc69486072)

Table of Figures

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

<Title>

Updated: 2021-Apr-16 by <Author>

# Load Forecasting Overview

Load forecasting is an integral part of the process of planning and operation of electric utilities; it has played a vital role in the power industry for over a century. But load forecasting can also be useful 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. This 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 behaviour, that need to be taken into consideration to predict the future electricity demand. Finding an appropriate forecasting model for a specific electricity network is not a trivial task [9][4][5]. Different factors can affect load forecasting such as; the location of the area, the type of customers in the region, weather factors (temperature, etc.), a trend in the time series dataset, the time of the day, day of the week, and other unpredictable factors (coronavirus outbreak, etc.). Electricity demand is assessed by tracking it periodically - hourly, daily, weekly, monthly, and/or yearly. 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 research, concentrating on horizons of less than 2 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]. Our aim to create a forecaster that could easily adapt to sudden changes in the demand profile, as well as it being able to easily 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) modelling [16], [17]. AI algorithms are considered to be smarter and better, as they provide the capacity to learn and adapt to the non-linear and complex relationships between the load and other influencing factors (i.e., weather, time of day) automatically [18]. Examples of these algorithms are Artificial Neural Networks [19][20], Fuzzy Regression Models [21], [22], Support Vector Machines [23], Gradient Boosting Machines [24].

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 is useful in modelling sequential data. However, RNNs have a weakness and it often fails because of vanishing/exploding gradient [25][6]. This led to the creation of Long Short-Term Memory (LSTM). LSTM is an RNN that overcomes the issue of vanishing gradient, and it provides a model that is capable of storing information for a long period. Munem[26] et al. argue that LSTM stores information better than other deep neural networks because of its memory cell configuration. Convolutional Neural Networks (CNNs) have also gained the attention of partitioners in this field. CNN is a feedforward network that was created after the human neurons in the way it’s structured. It is normally used for processing data with grid topology; its major 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. From their results and observations, they concluded that CNN is a viable technique that produces accurate load forecasts.

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

In this research work, our goal is to determine how much value some deep learning approaches (e.g. CNN and LSTM) adds in terms of forecasting accuracy; when they are compared to some of the current ones that have been available for many decades. The four benchmark algorithms are Seasonal Naïve, Multiple Linear Regression (MLR), Auto-Regressive Integrated Moving Average (ARIMAX), and Artificial Neural Network (ANN). These benchmark algorithms have been available for many years and have been used by researchers and utilities [29][4][30].

# Investigation

Many papers lack detailed information about their experimental set-ups, and this makes it challenging to conduct direct comparisons with the results that they report [1]. To make our work reproducible, some of the datasets we are using are from an independent system operator and the selected benchmark algorithms have a lot of documentation about how they could be implemented. 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. It’s therefore important to create algorithms that could adapt easily to these changes as they occur [31].

Recently, deep learning has gained the eyes of partitioners in this field. Therefore, we aim to compare the performance of some deep techniques with the current ones starting with the CNN and LSTM [32]. This will help us to figure out how much extra value is added to the forecasts when compared to the ones currently used by utilities. First, we begin by implementing all the algorithms, then we compare their performances on two 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 have been chosen 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 considered to be the most cost-effective forecasting model; it has often been used as a benchmark for developing much more sophisticated models [34][35]. 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 larger 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 being added to the current forecasting process. The formula for SNF is shown below;



Where;  is the time series and is the seasonal period. In summary, 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 can then be used to forecast future values [37]. The formula of the ARIMA is shown below;



Where  is the lag1 of the time series,  is the coefficient of lag1 estimated by the model,  is the intercept that has been estimated by the model,  are the error terms from respective lags. ARIMA in its basic form is; the 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 one where the time series was differenced at least once to make it stationary and you 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 from one another. The Auto-Regressive order p refers to the number of lags of the data that are used as predictors. While 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 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 nonlinear relationships between the electrical load and weather variables, which turns out to be false [1] [42]. For example, polynomial regression models can describe nonlinear 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  is a representation of a set of random variables that are independent and identically distributed and having a mean of zero. MLR models are fitted such that the sum-of-squares of differences of actual and forecasted values are minimized. 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 model is built as a multi-layer feed-forward Artificial Neural Network (ANN) as identified by the creators in this paper [45]. The ANN models are still being used today due to their ability to learn complex and non-linear relationships in the data on their own. The specification of independent variables explicitly in ANNs is not mandatory, like in the case of MLR models. This ANNSTLF model has been identified as the best-known ANN implementation for STLF [1], [42][30]. The ANNSTLF and its improvements of it have been used by several utilities in Canada and the US. The figure below shows the block diagram of the system:

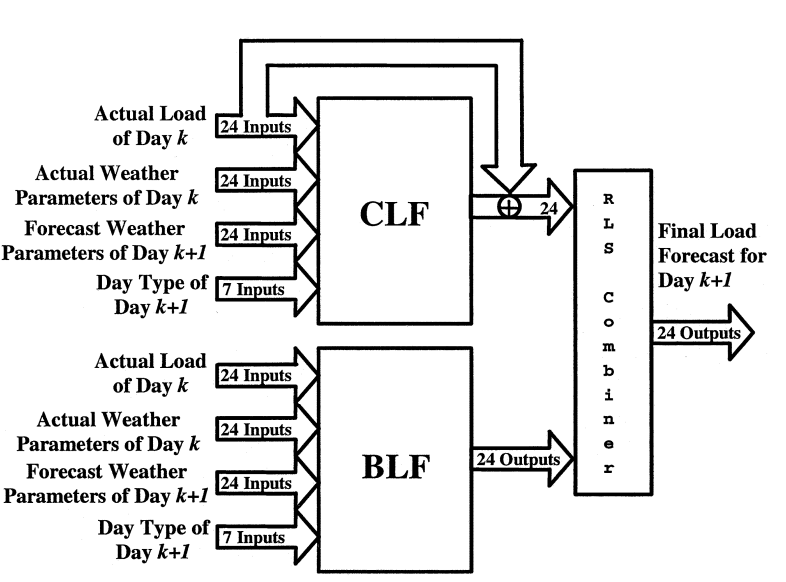


Figure 1:- The Block Diagram of the third generation ANNSTLF [45]

ANNSTLF is a multilayer perceptron that is trained with the error backpropagation 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 can be 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 also 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 feedforward neural network which mimics the human neurons in its design. CNN has been applied broadly in image and audio processing, natural language processing, and video recognition [48] [6]. CNNs are normally 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 2-dimensional grids and time series data are 1 dimensional, which makes the time series conversion necessary. The CNN architecture we are using in this research study consists of six layers namely; the image input layer, the 2D convolution layer, the rectified linear unit activation layer (relu), max-pooling 2D layer, fully connected layer, and a regression layer.

Because the ANNSTLF structure has been identified to be the best forecaster for short-term load forecasting [1], [42]. Our approach is to mimic 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

%% To be completed

## Data Sets and Metrics for Evaluation

Three time-series datasets are used, 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 optimally when trained with at least 3 years of data. Two datasets were gotten from the Ontario independent system operator [50]; one was from the city of Toronto, and the other was from Ottawa. Both datasets run from January 2010 until December 2019. The third dataset was gotten from Saint John Energy [32]. This dataset runs from January 2018 until December 2020. The weather variables for both datasets were gotten from the Government of Canada’s website [51]. These datasets were selected because the load and weather data are easily accessible compared to other ones we tried using.

The global metrics being used 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; it works best when there are no near zeros or extreme values. ME is the mean error of the forecasts across the entire horizon. MAE is a measure of the average magnitude of the forecast errors without the consideration of their direction. RMSE shows the absolute fit of the model; the closeness of the actual values to the predicted ones. Standard deviation tells how spread the errors are; it’s a measure of how far each error is from the mean error.

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

# Contributions

This study would help researchers get a better idea of how much value deep learning algorithms like the CNN and LTSM, adds when compared with some of the currently used ones. We aim to create an algorithm that could easily adapt to the rising trend in power demand that most utilities face every year, sudden changes in temperature, as well as every other random variable that affects the load demand. We also aim to have an algorithm that would understand and interpret the complex relationships in the data better, without the need for explicit specification from the user. This research work would be a reproducible experiment for other researchers to use as a point of reference in the future. The major reasons for this are because the datasets are from an independent system operator and there is also a lot of documentation about the algorithms used in this study.

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

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

[52] “Historical Climate Data - Climate - Environment and Climate Change Canada.” https://climate.weather.gc.ca/ (accessed Jan. 05, 2021).