

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-26 |

|  |  |  |
| --- | --- | --- |
|  | 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. For the stable supply of electricity, reserve power must be available to serve consumers (e.g., in case of high demand or failure in the current grid supply). 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 [ref].

Electric load forecasting is well studied [ref] and most current research is focused on developing more accurate forecasts. This is particularly relavent in today’s context, with the advent of new smart grid technologies. The demand patterns used to drive these technologies is complex due to deregulation of energy markets, and the number of different random variables, often governed by human behavior, that need to be taken into consideration to predict demand. Finding an appropriate forecasting model for a specific electricity network is not a trivial task [3][4][5].

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) [6]. Shorter-term forecasting has been the focus in most research, concentrating on horizons of less than 2 weeks [7][6][1].

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.).

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]. []Statistical approaches can forecast the current value of a variable through the use of a mathematical combination of past historical values of the variable, and previous or current values of other variables [8]. Examples of statistical techniques applied to electrical load forecasting include multiple linear regression analysis [ref], exponential smoothing [ref], and auto-regressive integrated moving average (ARIMA) modeling [ref]. AI algorithms are considered to be smarter and better, as they provide 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 [9]. Examples of these algorithms are Artificial Neural Networks, Fuzzy Regression Models, Support Vector Machines, Gradient Boosting Machines. In recent years, deep learning approaches have also become enticing to researchers in this field. The Recurrent Nerual Network (or RNN) introduced the concept of memory into neural networks, which is useful in modeling sequential data []. it’s known to boost the power of the ANN as it has deeper layers and can interpret load data better [10], [11].

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, and different periods.

In this research work, one of the datasets used is from an independent system operator (Ontario IESO), as this would make it easier for our work and experiments to be reproduced by researchers and practitioners in this field.

In this research work, we aim to compare the performance of the convolutional neural network (CNN) technique with some classical 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 [12][4][13]. It is important to see how much more value CNN adds when it is compared with the classical and currently used techniques.

# Investigation

Many papers lack detailed information about their experiment setup, this makes a hard for their experiments to be reproduced [1]. For this reason, we are including a dataset from an independent system operator and also, the selected benchmarks algorithms are classical approaches with a lot of documentation about how they can be created. Also, there is a rising trend in the power demand at most utilities each year due to new systems and more sophisticated equipment been added. Therefore, it’s important to have algorithms that could adapt easily to these changes as they occur [14].

Our solution is to make a comparison between a deep learning algorithm (CNN) and some benchmark algorithms. First, we begin by implementing all the algorithms, then we compare their performances on two datasets, and see when one performs better or worse, then we will make an improvement based on the new information we find.

## The Benchmark Algorithms

[]Four algorithms have been chosen for benchmark comparison: seasonal naïve approach, 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 [3][5][1]. The seasonal naïve is simple but it still performs well enough to be a benchmark for other sophisticated models. ARIMAX is a classical model, it has been used for many years and has been proven to be quite good for load forecasting [15], [16]. The MLR method is still used today for load forecasting; it can model non-linear relationships with the specification of independent variables to the dataset. ANNSTLF has been identified as the best forecaster in short-term load forecasting, therefore it serves as a good benchmark [1], [17].

### Seasonal Naïve Approach

[]The naïve approach is considered to be the most cost-effective forecasting model; it often serves as a benchmark for developing much more sophisticated models [18][19]. In the naïve approach, the forecast is taken as the previously observed value; this type of forecast is only suitable for time series data. This approach works best if the previous observation has a high similarity with the current; it is sometimes called a similar day approach. If there is seasonality in the time series; the seasonal naïve approach is preferable, because forecasts will be equal to the value from the last season. The seasonal naïve approach is most useful when there is a very high level of seasonality in the dataset [20].

The naïve approach, when used as a baseline for other methods; it gives us an understanding of how much value is being added to the current forecasting process. The formula for the naïve approach and the seasonal naïve approach is shown below respectively;



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. 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 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 [21]. 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 [22]. If we would love to take into consideration exogenous variables like temperature, day of the week, etc.; the ARIMAX model would have to be taken into consideration [15].

### Multiple Linear Regression

Multiple linear regression is one of the most used statistical techniques for load forecasting [23] [24]. 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] [17]. 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 [17][7][24].

### 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 [25]. 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], [17][13]. 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:

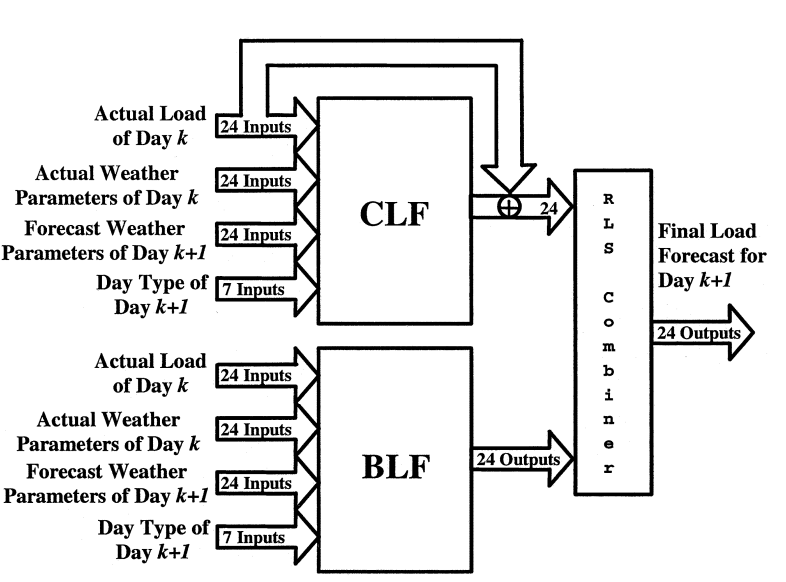


Figure 1:- The Block Diagram of the third generation ANNSTLF [25]

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 [17][26][27].

## 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 [28] [10]. CNNs are normally used with image data; time-series data can be arranged to mimic image data and it can then be fed into a CNN [29][30]. CNNs normally process data with a grid topology; images are 2-dimensional grids and time series data are 1 dimensional, which makes this 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], [17]. Our approach is to mimic the ANNSTLF structure by creating a Base Load Forecaster, Change in the Load Forecaster, and RLS combiner; with the CNN algorithm. The architecture will have the same inputs and structure as the ANNSTLF, but the BLF and CLF algorithms will be created using CNNs.

## Data Sets and Metrics for Evaluation

The datasets to be used are time series sampled on an hourly horizon, and they contain the temperature variable as it plays a huge role in power demand. The more data available for training, the better the algorithms perform. The authors of the ANNSTLF algorithm mentioned that the algorithm works optimally when trained with at least 3 years of data. The first dataset selected was gotten from the Ontario independent system operator [31]. The second dataset was gotten from Saint John Energy [32].

The city of Toronto’s dataset runs from January 2010 until December 2019, and the Saint John’s datasets run from January 2018 until December 2020. The Saint John’s dataset isn’t publicly available but I believe it would be useful for my research work as it would help us to see how well the algorithms perform when using a bigger and a smaller dataset from different cities with large population gaps. The weather variables for both datasets were gotten from the government of Canada website [33].

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 shows us the distribution of error for each hour, month, and day; therefore, it helps us to identify situations where the algorithms perform better or worse.

# Contributions

Researchers would get a better idea of how much value the CNN algorithm adds when compared with some of the best classical approaches. We aim to create an algorithm that could easily adapt to the rising trend in power demand that most utilities face every year. We also aim to have an algorithm that understands and interprets the relationships in the data better, without the need for explicit specification. This will also be a reproducible experiment for other researchers to use as a point of reference in the future, because of the publicly available dataset being used in this research study (Ontario IESO).

# 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] 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.

[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] 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.

[7] 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.

[8] N. Amjady, “Short-term hourly load forecasting using time-series modeling with peak load estimation capability,” *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 798–805, 2001, doi: 10.1109/59.962429.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] R. Bonetto and M. Rossi, “Parallel multi-step ahead power demand forecasting through NAR neural networks,” in *2016 IEEE International Conference on Smart Grid Communications, SmartGridComm 2016*, Dec. 2016, pp. 314–319, doi: 10.1109/SmartGridComm.2016.7778780.

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

[18] 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.

[19] 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.

[20] 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.

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

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

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

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

[29] 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.

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

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

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

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