Long Term Electricity Load Forecast Based on Machine Learning for Cameroon’s Power System

Terence K. Lukong1,2, Derick N. Tanyu3, Thomas T. Tatietsé4 & Detlef Schulz2

1 Doctoral School, National Advanced School of Engineering University of Yaoundé I, Yaoundé, Cameroon
2 Department of Electrical Power Systems, Helmut Schmidt University, Hamburg, Germany
3 Centre for Industrial Mathematics (ZeTeM), University of Bremen, Bremen, Germany
4 IUT FV of Bandjoun, Bandjoun, Cameroon

Correspondence: Terence K. Lukong, Doctoral School, National Advanced School of Engineering Yaoundé, University of Yaoundé I, Cameroon. Tel: 237-67-477-8879. E-mail: terencelukong@yahoo.com

Received: April 30, 2022            Accepted: May 25, 2022            Online Published: May 30, 2022
doi:10.5539/eer.v12n1p45            URL: https://doi.org/10.5539/eer.v12n1p45

Abstract
A reliable power supply has long been identified as an important economic growth parameter. Electricity load forecasts predict the future behavior of the electricity load. Carrying out a forecast is important for real-time dispatching of power, grid maintenance scheduling, grid expansion planning, and generation planning depending on the forecasting horizon. Most of the methods used in long-term load forecasting are regressions and are limited to predicting peak loads of a yearly or monthly resolution with low accuracy. In this paper, we propose a method based on long short-term memory-recurrent neural networks (LSTM-RNN) cells with relations between identified influential econometric load-driving parameters which includes: the Gross Domestic Product (GDP), Population (H), and past Electric Load Data. To the best of our knowledge, the use of the GDP and H as two additional independent variables in load forecast modelling using machine learning techniques is a novelty in Cameroon. A comparison was performed between a linear regression (LR)-based long-term load forecast model (a model currently used by the Transmission System Operator of Cameroon) and LSTM-RNNs model constructed. The results generated were evaluated using a Mean Absolute Percentage Error (MAPE) within the same period of evaluation, and the overall value of the MAPE obtained for LSTM-RNNs model was 5.4962 whereas that for the LR model was 7.5422. Based on these results, the LSTM-RNN model is considered highly accurate and competent. The model was used to generate a forecast for the period of 2022–2026 with an hourly resolution. A MAPE of 5.4962 was obtained with a computational time of approximately ten minutes, making the model vital for offline use by utilities due to its capacity to quantitatively and accurately predict long-term load with an hourly resolution.

Keywords: electricity load forecast, LSTM-RNN model, machine learning, load parameters

1. Introduction
Electric Load forecasting is the act of predicting the future behavior of electric load. In emerging economies like Cameroon, installed power system facilities are approaching their nominal useful lives in the face of a steady increase in electricity demand. This situation is similar to studies carried out by (Melodi, Adeniyi, & Oluwaniyi, 2017) for Nigeria’s power system. Planning for increased power capacity and reinforcement or expansion of the electrical network should be based on an accurate prediction of long-term demand to ensure electric energy service continuity and reliability. Developing a reliable load forecast model is very crucial and essential for the strategic planning of the national network. There are three different time horizons of forecast (Suganthi & Samuel, 2012): short-term forecast, usually from a few hours ahead to a week ahead which is essentially useful for real-time electric energy economic dispatching and unit commitment; medium-term forecast, varies from a month to one year ahead and is mainly useful for maintenance scheduling of the grid; long-term forecast which varies from one year to ten years or several years and is mainly considered for grid expansion planning, power flow studies and generation investments decisions (Agrawal, Muchahary, & Tripathi, 2018). By principle, a load forecasting model aims at a mathematical representation of the relationship between load and influential parameters. Such a model is identified with coefficients that are used to forecast the future values by...
extrapolating the relationship to the desired lead time as argued by (Khuntia, Rueda, & van der Meijden, 2016). To plan the efficient operation and economical capital expansion of a power system, including generation, transmission, and distribution, the system owner must be able to anticipate the need for power delivery, the quantity of power to be delivered, and where and when it will be needed. In the 1980s, (Vlahović & Vujošević, 1987) proposed a mathematical model based on regression to forecast long-term load demands. In addition, different time series statistical models have been proposed by other researchers like the autoregressive, moving average, and the autoregressive moving average forecast models which are popular and widely accepted by power utilities (Po, Min, & Di-chen, 2006). A model based on advance grey theory developed by (Zhao & Niu, 2010) has been proven to produce a more accurate mid-term forecast than the traditional grey model. They argued that the proposed model improves the fault prediction precision of power load. However, most of these traditional methods have unexpected variations in environmental parameters and most studies using these methods used an annual or monthly granularity of peak load to perform forecasts with relatively small datasets. The results obtained though arguably reasonable is limited in use on a practical basis for utility companies and main stakeholders to make good decisions. This is due to the low level of accuracy, low granularity of predicted load and the nonlinear characteristic of electrical load not captured precisely (Tamer & Ehab, 2021). In addition, studies realized by (Soliman & Mohammad, 2010) demonstrated that electric loads depend on a number of complex factors. They equally have nonlinear characteristics and satisfactory results which may not be obtained using statistical methods (Hagan & Behr, 1987). They suggested that a better method of long-term load forecasting would be one that could determine nonlinear relationships between load and various economic and other factors that can adapt to changes.

The recent successes of machine learning for the task of prediction as demonstrated by (LeCun, Bengio, & Hinton, 2015) provides a promising direction for the field of long-term load forecasting. Specifically, Recurrent Neural Networks (RNNs) have been known to produce good predictions when dealing with time series data. Long Short-Term Memory-Recurent Neural Networks (LSTM-RNN) which is an extension of RNNs maintains both long-term and short term states, establishes a non-linear relation between load and the economic factors, converge quicker and are also able to capture longer term dependencies, thus making them ideal for longer term forecasting of electricity demands (Agrawal, Muchahary, & Tripathi, 2018).

This paper presents a long-term electric load forecast model for Cameroon’s power system using Long-Short-Term-Memory based Recurrent Neural Network (LSTM-RNN) model for forecasting electricity demand for a period of five years. The model is realized based on past electric load data provided by the Société Nationale de Transport d’Électricité (SONATREL), the Transmission System Operator of Cameroon. The forecasts generated have an hourly time-step, thus it is operationally usable by the utility companies. The remainder of this paper is divided as follows. In section 2, we describe the application domain, the LSTM-RNN model and the dataset. In section 3, we explain the methodology. The results obtained and discussions are presented in section 4, and in section 5, the conclusion.

2. Application Domain, LSTM-RNN Model, and Dataset

2.1 Application Domain

Cameroon is divided into 10 administrative regions. The electrical transmission grid in Cameroon is composed of three separate grids having a spatial coverage of Cameroon as follows:

- South Interconnected Grid (SIG): serves demand for six regions with the largest towns which include the Center, Littoral, West, Northwest, Southwest, and South regions of Cameroon. It is the largest grid serving almost 75% of the entire population of Cameroon and most industries.
- North Interconnected Grid (NIG): serves three regions which includes the Adamawa, the North, and the Far North regions.
- East grid which serves the East region.

The global urban population percentage covered by the aforementioned grids are 78%, 18%, and 3.3% respectively (Ministry of Water Resources and Energy, 2014)

This paper focuses on long term load forecast for the SIG. The SIG is composed of 225 kV lines of 483 km and 90 kV lines of 1011 km with 30 transmission and sub transmission substations. It evacuates all the energy generated by the hydropower plants in SIG (Songloulou: 384 MW, Edea: 264 MW, and Memve’ele: 200 MW), gas power plant (Kribi Power Plant capacity 200 MW), and other thermal plants to different load centers.
Table 1. Peak load evolution in the SIG

| Year | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Peak Load (MW) | 490 | 514 | 555 | 585 | 619 | 667 | 670 | 684 | 744 | 764 | 790 | 836 | 844 | 892 | 890 |

As shown in Table 1, the evolution of peak load for the SIG from 2006 to 2020 has an average demand growth rate of 5.5%. This observed average growth rate could vary based on the factors influencing demand.

2.2 The LSTM-RNN Model

Artificial Neural Networks (ANN) seek to recognize relationships between input data and output data. Inspired by the biological neural networks in the human brain, they imitate the functioning of the human brain in recognition/learning. ANNs are thus made up of a collection of artificial neurons, connected to each other. Each neuron receives information (a real number), processes it, and transmits it to the next neuron(s). This process continues till the output neuron, which finalizes the learning task. Each neuron typically has a weight and a non-linear activation function which are applied to the input information (real number). The result (output signal) is another real number which is a function of the applied weight. Training an ANN entails learning these weights.

Recurrent Neural Networks (RNN) are a special type of ANN that achieve their task of learning with feedback connections linking the neurons in one layer to those in the previous layers (Kasabov, 1996). In this way, memory is preserved as information propagates through the RNN from one neuron to the other. However, RNN suffers from short-term memory and vanishing gradient decent problem during back-propagation especially when dealing with long time series data. This problem is overcome in the long short-term memory (LSTM), introduced by (Hochreiter & Schmidhuber, 1997) and refined and popularized by many researchers. LSTMs are more sophisticated RNN and they permit the learning of more complex relationships between sequential inputs and outputs with little feature engineering (Graves, 2012). They have recorded tremendous success when working with large variety of time series data, as they remember information for longer period. They achieve this with a cell state, which works like a conveyor belt that carries the relevant information from the earlier steps to later steps. LSTMs introduce the notion of gates; forget, input and output gates. Figure 1 shows the schematic of an LSTM memory block with one cell.

The three gates are nonlinear summation units that collect activations from inside and outside the block and control the activation of the cell via multiplications (small black circles). The input and output gates multiply the input and output of the cell, whereas the forget gate multiplies the previous state of the cell. No activation function is applied within the cell. The gate activation function ‘f’ is usually the logistic sigmoid, such that the gate activations are between 0 (gate closed) and 1 (gate open). The input and output activation functions of the cell (‘g’ and ‘h’) are usually a tanh or a logistic sigmoid, though sometimes ‘h’ is the identity function. The weighted ‘peephole’ connections from the cell to the gates are shown with dashed lines. All other connections within the block are unweighted (or equivalently, have a fixed weight of 1.0). The only outputs from the block to the rest of the network emanate from the multiplication of the output gate. The multiplicative gates allow LSTM memory cells to store and access information over long periods, thereby mitigating the vanishing gradient...
The memories $h$ and $c$ are updated and then the output $y_t$ is computed as follows

$$y_t = \tanh(W_y h_t + b_y)$$  \hspace{1cm} (8)
entire Cameroon with a yearly time set. H data was also obtained from (The World Bank data, Population of Cameroon) for the entire Cameroon with a yearly time step. The 2 variables obtained are representative of SIG and the LSTM-RNN learns this data. Hourly Past Load data for the SIG was equally obtained as indicated earlier for the same period.

3. Methodology
The model input includes the past system (previous year) load data for all the months in a year, the monthly population and GDP, while the corresponding load data are the LSTM-RNN targets. The inputs were fed into the model and after sufficient training, the model was used to predict the load output. Subsequently, forecast GDP and H data was used as inputs to the trained LSTM model to predict the system load for the forecast period (in this case, the next five years) because input data of GDP and H were needed for the forecast years (2006–2014). A linear time series regression equations based on historical data was applied to grow the GDP and H data for desired time steps Y (months).

For the training and testing of the proposed model, the input dataset (past system load data) from the national electricity transmission grid of Cameroon (SONATREL) was used. This data consisted of hourly load demands from 2006 to 2020 as shown in Figure 2. The population of Cameroon from 2006 to 2020 and the GDP growth of the country over this same period were also used. These were obtained from the World Bank database and are shown in Figure 3 and Figure 4. During training, data from 2006 to 2014 were used as training data, whereas those from 2015 were used as validation data. This enabled us to obtain suitable hyper parameters that permit good optimization and avoid over fitting. The trained model was then used to predict the load forecast for 2015.

To ensure proper learning by the model, some feature engineering was performed on the data. The engineered features were then used as the input to the model.
Given the various factors that affect electricity consumption, feature engineering is used to identify influencing factors. While the data for the electricity demand is readily available hourly, this is not the case for the population and the GDP growth. As a workaround, the GDP (and population) for a particular year was used for every hour of that year. The date and time are equally used as input to the neural network, but not in the original forms. First, they are converted to hour, day of the week, month, and week of the year because these are important factors in electricity demand. For example, electricity demand on Fridays (the fifth day of the week) can be observed to have a particular trend throughout the year and over many years. A similar remark can be made for other dates and time data. In addition, it is important to ensure that the cyclical nature of these date and time data is reflected in the features, meaning the closeness of Day 7 and Day 1, Hour 24 and Hour 1 should be evident in the feature. This is made possible using cyclical feature encoding which leverages ideas from trigonometry. Finally, a data pair used to train, validate, and test the model consist of an hourly value as label (output) and sequential data (size 168) of the past week as feature (input). Figure 5 shows an illustration of the feature engineering step. A good practice is to ensure that the input to the model for both training and test data originates from the same distribution. This makes training easier as faster convergence is ensured. Consequently, a scaling was used to ensure data values are within a certain interval [minimum, maximum].

Note that with the demand value of the last year included as one of the input features, the training data effectively begins from 2007 to 2014 (61320 observations). The validation data is that of 2015 (8760 observations). Once the model is trained on the training data, and hyper parameters tuned based on the validation data, this model is then used to predict the load demand for 2016. The predicted load for 2016 will then serve as input to predict the load of 2017. This process continues until the load forecast for 2016 to 2020 (5 years) is obtained. Thus, our model does not only predict short-term load but long-term load as well.

### 3.2 Model

The dataset is sequential, as a result we employed a sequence model. The sequence model is an LSTM model owing to its good performance with time series data. The best-performing model had 3 LSTM layers with an input size of 12 features, 168 (24 hours for 7 days) time steps, and 64 features in the hidden state. A dropout of 0.2 was applied to the outputs of each LSTM layer. At the exit of these three LSTM layers was a fully connected linear layer leading to a single dimension output. The model was trained for 100 epochs; each iteration in the epoch made use of a batch size of 64. The Adam optimizer was used during training. These set of hyperparameters, which resulted from an extensive hyperparameter tuning, when put together, provided the best
convergence of our model during training as shown in Figure 7. **Variation of loss during training**. The Adam optimizer was used during training as it tends to provide faster convergence even if the gradient becomes sparse during training, as opposed to other optimizers like Stochastic Gradient Descent, RMSprop, Adagrad. (Sun, Cao, Zhu & Zhao, 2019) provides a good review of these optimization algorithms.

Conversely, a model based on simple LR was equally constructed for benchmarking.

4. Results and Discussions

As earlier stated, the obtained model was used to forecast the load demand for 2016 to 2020. The model performance was evaluated using the MAPE metric.

\[
\text{MAPE} = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|
\]

(9)

Where \( A_t \) is the actual value and \( F_t \) is the forecast value, \( n \) is the number of observations in the test set.

Table 2 shows the result of the yearly MAPE for both the LSTM-RNN model and the LR benchmark models. Figure 7 to Figure 11 also show the observations. From the table we observe how the prediction capability of the LSTM model decreased over the years. This is expected, because we accumulate prediction errors as we progress in years: in order to predict year 2017, load forecast from 2016 is used together with the load of the previous years. Similarly, in order to predict year 2018, load forecast from 2016 and 2017 is used. However, the ability of LSTM to keep memories still enables the predictions to be better than that of the linear regression. In Figure 7, which shows a visual comparison of the LSTM load forecast and the Linear Regression load forecast, we observe how the LSTM predictions overlaps more often with the Ground Truth, while the Linear Regression predictions always stay within a certain interval. As expected, this overlap is even more evident in the first year (2016) as we observe in Figure 8. Figures 9, 10 and 11 further stress this overlap as we zoom even closer to few the first month, first week and first day respectively.

Table 2. Yearly Performance metric for the LSTM model and Linear Regression model.

| Year | 2016 | 2017 | 2018 | 2019 | 2020 | Overall |
|------|------|------|------|------|------|---------|
| LSTM-RNN MAPE | 4.2053 | 4.8271 | 4.6938 | 7.0613 | 6.3617 | 5.4297 |
| LR MAPE   | 7.2953 | 7.8883 | 7.0927 | 7.9616 | 7.1199 | 7.5422 |
Figure 8. Five years (2016 to 2020) prediction

Figure 9. One-year (2016) load forecast
Figure 10. One-month (January 2016) load forecast

Figure 11. One-week (January 01 to January 07, 2016) prediction
5. Conclusion

This paper proposes an LSTM-RNN model with a satisfactory MAPE of 5.4297%, which outperforms the current method (Linear regression model) used for load forecasting at SONATREL for Cameroon’s power system. The model is found to be highly accurate and can reliably be used to predict the future years of load. The years (2021 to 2025) was predicted on a rolling basis. Population and GDP projections for the said years was equally needed as inputs. Figure 11 shows such a prediction. Consequently, based on the good performance of this model, it will assist utilities in long term grid expansion strategies and in investment decisions. Considering another variant of LSTM; gated recurrent unit (GRU) which focuses only on gate functions rather than the long-term memory. GRU can reduce the number of parameters in comparison to LSTM since there are two gates in GRU: reset gate and update gate. The objective of the authors to keep the ability of LSTM to reserve long-term information while the whole network is made as simple as possible. We would like to know further whether GRU can outperform our LSTM model, this would be our future research orientation.

References

Agrawal, R.K., Muchahary, F., & Tripathi, M.M. (2018). Long-term load forecasting with hourly predictions based on long-short-term-memory networks, *IEEE Texas Power and Energy Conference (TPEC)*, College Station, TX, 1-6. https://doi.org/10.1109/TPEC.2018.8312088

Farrag, T. A., & Elattar, E. E., (2021). Optimized Deep Stacked Long Short-Term Memory Network for Long-Term Load Forecasting, in *IEEE Access*, 9, 68511-68522. https://doi.org/10.1109/ACCESS.2021.3077275
Graves, A. (2012). Long short-term memory. In: Supervised Sequence Labelling with Recurrent Neural Networks”. Studies in Computational Intelligence, 385. Springer, Berlin, Heidelberg. 38-39. https://doi.org/10.1007/978-3-642-24797-2_4

Hagan, M. T., & Behr, S. M., (1987). The Time Series Approach to Short Term Load Forecasting. IEEE Trans. Power Syst., 2(3), 785–791. https://doi.org/10.1109/TPWRS.1987.4335210

Hochreiter, S., & Schmidhuber, J., (1997). Long short-term memory. Neural computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kasabov, N.K., (1996). Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering (2nd ed.). A Bradford Book The MIT Press Cambridge, Massachusetts London, England, pp 292.

Khuntia, S. R., Rueda, J. L., & van der Meijden, M. A. M. M. (2016). Volatility in electrical load forecasting for long-term horizon — An ARIMA-GARCH approach, International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2016, pp. 1-6. https://doi.org/10.1109/PMAPS.2016.7764184

LeCun, Y., Bengio, Y. & Hinton,G. (2015). Deep learning. Nature, 521, 436–444 https://doi.org/10.1038/nature14539

Melodi, A. O., Adeniyi, & Oluwaniyi, R. (2017). Long term load forecasting for Nigeria's electric power grid using ann and fuzzy logic models, IEEE 3rd International Conference on Electro-Technology for National Development (NIGERCON), 2017, pp. 962-968. https://doi.org/10.1109/NIGERCON.2017.8281961

Ministry of water resources and energy. (2014). “Plan de Developpement du Secteur De L'Electricité A l’Horizon 2030 (PDSE 2030).

Po, L., Min, L., & Di-chen, L. (2006). Power Load Forecasting Based on Improved Regression. Power System Technology, 30(1), 99-104.

Soliman, S.A., & Mohammad, A. (2010). Electrical Load Forecasting: Modeling and model construction, ISBN 978-0-12-381543-9.

Souza, R.C., Esteves, G.R.T., Bastos, B.O., Cyrino, F.L., & Calili, R.F. (2015). Long term electricity forecast: a systematic review. Procedia Computer Science, 55, 549-558. https://doi.org/10.1016/j.procs.2015.07.041

Suganthi, A., & Samuel. (2012). Energy models for demand forecasting—A review. Renew. Sust. Energy. Rev., 16, 1223–1240. https://doi.org/10.1016/j.rser.2011.08.014

Sun, S., Cao, Z., Zhu, H., & Zhao, J. (2019). A survey of optimization methods from a machine learning perspective. IEEE transactions on cybernetics, 50(8), 3668-3681. https://doi.org/10.1109/TCYB.2019.2950779

Ushiku Y. (2021) Long Short-Term Memory. In: Ikeuchi K. (eds) Computer Vision. Springer, Cham. https://doi.org/10.1007/978-3-030-63416-2_856

Vlahović, V.M., & Vujošević, I.M. (1987). Long-term forecasting: a critical review of direct-trend extrapolation methods. International Journal of Electrical Power & Energy Systems, 9(1), 2-8.ISSN 0142-0615. https://doi.org/10.1016/0142-0615(87)90019-6

Zhao, W., & Niu, D. (2010). A mid-long term load forecasting model based on improved grey theory. 2010 2nd IEEE International Conference on Information Management and Engineering, 633-635. https://doi.org/10.1109/ICIME.2010.5478173

Web Pages
The World Bank Database, GDP of Cameroon. Retrieved from https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=CM

The world Bank Database, Population of Cameroon. Retrieved from https://data.worldbank.org/indicator/SP.POP.TOTL?locations=CM

Copyrights
Copyright for this article is retained by the author(s), with first publication rights granted to the journal.
This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).