UC Riverside
2019 Publications

Title

Statistical and Deep Learning Methods for Electric Load Forecasting in Multiple Water Utility Sites

Permalink

https://escholarship.org/uc/item/2fx5x9fm

Authors

Yusuf, Jubair
Faruque, Rumana
Ula, Sadrul
et al.

Publication Date

2019-11-01

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Statistical and Deep Learning Methods for Electric Load Forecasting in Multiple Water Utility Sites

Jubair Yusuf, *Student Member, IEEE*, Rumana Binte Faruque, A S M Jahid Hasan, Sadrul Ula, *Senior Member, IEEE*,
Department of Electrical and Computer Engineering
University of California Riverside
Riverside, CA, USA
jyusu001@ucr.edu, rfaru001@ucr.edu, ahasa006@ucr.edu, sula@cert.ucr.edu

***Abstract*—Most of the water utilities in the U.S. consume a lot of electrical energy for water treatment and delivery. Despite being large energy consumers, priority is not given to electric load forecasting in water utilities. An accurate forecast of electric load can pave the way to shaving peak demand and reducing high electricity bills. This paper applies a popular statistical approach named Auto Regressive Integrated Moving Average (ARIMA) and Deep Learning techniques to forecast daily electric load over a period of a month and 15-minute moving average electric load of a day for two sites in a southern California water utility. A comparative performance of these techniques with relevant error metrics has been introduced. The electric load of a water treatment plant and a pumping station have been forecasted with these two methods. Deep Learning techniques result in better load prediction for both accounts and in both time resolutions. This allows operators to take possible appropriate actions resulting in reduced electrical demand for any given billing period.**

Keywords— water utility, water energy nexus, load forecasting, Auto Regressive Integrated Moving Average, Long Short Term Memory, deep learning, statistical approach

I. INTRODUCTION

Water districts are one of the largest energy consumers. The water supply and wastewater facilities use 69.4 billion kWh energy per year in the U.S. [1].Despite heavy energy consumption, water utility companies do not pay much attention to their electricity usage for treating and delivering water. The main priorities of the water utilities include the maintenance of the water quality and assurance of the uninterrupted service to its customers. The lack of attention in adopting necessary steps for using energy efficiently results in higher demand during critical time periods, unregulated demand strategies and increases in electricity bills.

Water Treatment Plants (WTP) and Pumping Stations (PS) are two main types of electricity use in water utilities. Water treatment plant collects water from either surface reservoirs or underground aquifers, and then treat water to disinfect it and make it drinkable. Purification and maintaining the quality of water are the main responsibilities for the operators of the water treatment plants. Electric load demand for WTPs vary according to the treatment process and time of the day. Pumping stations are of two major types: 1. water pumping from wells and 2. booster stations and water storage tanks for water distribution. Their operation is somewhat unpredictable in nature and they do not strictly follow any pattern or seasonality due to the limitations of the availability and quality of water.

High electricity usage in a typical water district puts them into Time of Use (TOU) based rate schedules, which have charges such as on-peak, mid-peak and off-peak for the same day. Demand charge at on-peak periods is very high compared to other periods of the day. Demand charges are based on the highest 15 minute rolling averages over a billing period occurring at the appropriate time periods. The consumer is also responsible for facilities related demand charge, which is related to the maximum demand of the month irrespective of the time period. In most cases, the demand charge constitutes more than half of the total electric charge for any billing cycle. Inappropriate water treatment work scheduling and ignoring TOU based electricity rates result in large electricity bill for water utilities. Prior knowledge of electric load profile can alleviate this situation and promote efficient operation of a water utility.

Electric load forecasting has been used for utilities, residential and commercial consumers for a long time. As water demand is needed to be known for water utility operators, water demand forecasting has been a common practice for a long time. Econometric models, Artificial Neural Network (ANN) approaches and probabilistic prediction methods are done for long and short term water demand forecasting [2-3]. To the best of our knowledge, not many works have been done to forecast the water utilities electric load. The electric load and water demand of the water utilities are not linearly correlated. Statistical approaches of load forecasting are being used as standard operations of electric load forecasting. Auto Regressive Moving Average (ARIMA) is the most popular time series approach of load forecasting for stationary data. Recently, different Deep Learning techniques are being developed and implemented for time series forecasting. Recurrent Neural Network (RNN) is being widely used due to it's robustness and capability of handling data for a longer time interval [4-6].

It may be possible to reduce electric demand for any 15 minute period without interrupting water service. Optimal use of multiple pump motors can help reduce the high peak demand especially during the high cost peak demand period and handle the TOU electric bill issues. With prior knowledge based plan, it is possible to implement demand management strategies, which include putting a limit on the maximum number of pumps running at the same time, and longer operation of pump motors instead of shorter periods of on-off operations [7]. These steps can eventually help to reduce the peak demand and result in lower electricity bill.

The main contributions of this work are threefold.

1. Load forecasting techniques have been developed to forecast the electric load of a water utility.

2. Evaluation of both statistical and deep learning approaches for different sizeable water utility accounts.

3. Selection of the appropriate method for load forecasting to introduce the opportunity of electricity bill saving for the water utilities.

978-1-7281-5650-7/19/$31.00 ©2019 IEEE

This paper has the following subsections. Section II describes the system, Section III states the methodology, and Section IV discusses the experiment, while simulation results and performance evaluation are shown in Section V and VI. Finally, the conclusions are drawn in Section VII

II. System Description

The water district is a municipal water district in southern California which supplies adequate water for landowners and residents. It consists of 17 water storage reservoirs with a capacity of nearly 80 million gallons of water using over 400 miles of potable water pipelines. The water treatment plant treats up to 34 million gallons of water each day. In addition, the water reclamation facility produces up to 2 million gallons of recycled water per day.

After a brief review of the 15-minute demand profile of different sites, two sites have been selected. One is water treatment plant and the other one is a general water pump station. The water treatment plant has a daily average demand of 331.06 kW and the pumping station has a daily average demand of 51.28 kW. There are 10 large pumps in water treatment plant whereas there are 5 pumps in the pumping station. All pumps are from 50 hp to 250 hp range.

III. Methodology

Auto Regressive Integrated Moving Average (ARIMA) is a popular method for time series analysis and forecasting the future values. Deep learning techniques like Long Short Term Memory (LSTM) are also capable of time series data prediction. The water utility electric load has been forecasted using both these approaches.

A. ARIMA

ARIMA has been proposed by Box and Jenkins [8]. This approach is being used to predict the stationary data. In case the data is non stationary, the data can be made stationary by differencing and ARIMA can be applied after that. Two types of ARIMA are usually used: a. seasonal, b. non-seasonal. As the data used here have not shown any trend or seasonality after the seasonal decomposing, non-seasonal ARIMA has been used here. Non-seasonal ARIMA can be modeled by the following equation.

$$(1-\Phi_1 B-\Phi_2 B^2-\ldots-\Phi_p B^p)\,(1-B)^d\, y_t = \theta_0+ (1+\theta_1 B+\theta_2 B^2+\ldots\theta_q B^q)\,\varepsilon_t \quad (1)$$

where B is backshift operator and can be expressed as $By_t = y_{t-1}$, ε_t is white noise and θ_0 is a constant. p is the auto regressive order and q defines the moving average order. d is the degree of differencing to make the data stationary. Φ_1 Φ_2 Φ_3 …. Φ_p and θ_1 θ_2 θ_3….. θq are autoregressive and moving average coefficients respectively.

B. LSTM Network

Long Short Time Memory (LSTM) network is a special type of Recurrent Neural Network (RNN). Conventional RNN approaches can not learn and apply the long term sequential values to predict the data for the future values. LSTM eliminates the vanishing gradient problem of vanilla RNN which fails to hold the memory of long term lagged values [9].

LSTM has a hidden unit embedded with four layers inside it. The first one is forget gate (f_t) that decides which data should be kept or forgot. The other two gates are input (i_t) and output (o_t) gates. All these three gates are sigmoid functions and can be expressed in the form of equations (2-4).The new cell state (c_t) and the output of the hidden layer (h_t) are being updated based on the values of previous three gates. They can be modeled by equations (5-6). σ_g and σ_h are sigmoid and hyperbolic tangent functions respectively. W_f, W_i, W_o, W_c and b_f, b_i, b_o, b_c are the weights and biases of the forget gate, input gate, output gate and cell state respectively. The input vector at each time step t is denoted by x_t.

$$f_t = \sigma_g(W_f\,.\,[h_{t-1}, x_t] + b_f) \quad (2)$$

$$i_t = \sigma_g(W_i\,.\,[h_{t-1}, x_t] + b_i) \quad (3)$$

$$o_t = \sigma_g(W_o\,.\,[h_{t-1}, x_t] + b_o) \quad (4)$$

$$c_t = f_t * c_{t-1} + i_t * [\sigma_h(W_c\,.\,[h_{t-1}, x_t] + b_c)] \quad (5)$$

$$h_t = o_t * \sigma_h c_t \quad (6)$$

IV. Experiment

A. Data Collection and Preprocessing

The load forecasting is done for both 15-minute and 24-hour resolution. April 2018 data are used for 15-minute load prediction. The electricity demand data from May 2017-April 2018 have been used as baseline data for daily load prediction. The daily average was calculated from available 15-minute rolling average data. The water treatment plant has a daily high demand whereas the electricity demand for the pumping station is low.

The daily load data of first eleven months have been used for model validation in ARIMA modeling and test set in LSTM network for both accounts. The data of month of April has been used for validating models. The last day of April is used for validating the 15 minute resolution model. Figure 1 shows the daily average electric load of first eleven months for treatment plant and pumping station.

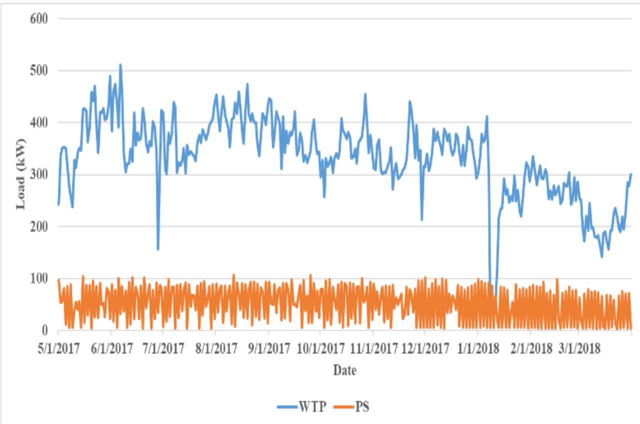

Fig 1: Average Daily Load Profile used as Test Set for a Day Ahead Model Formulation: Water Treatment Plant and Pumping Station

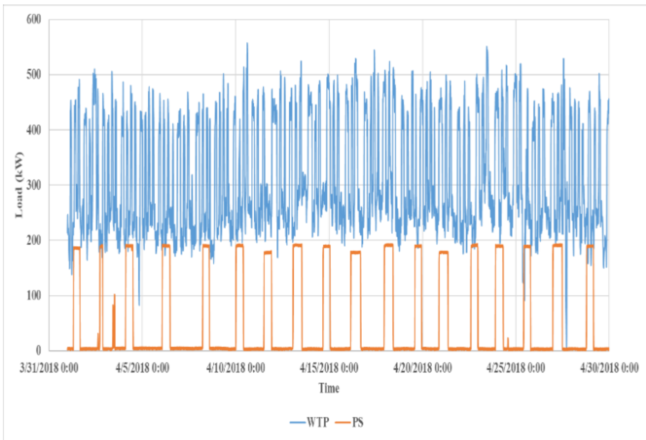

Fig 2: April (First 29 Days) Load Profile used as Test Set for a 15 Minute Ahead Model Formulation: Water Treatment Plant and Pumping Station

B. *Model Selection for ARIMA*

The first step of applying ARIMA model is to check the stationarity of the data. Stationarity of any time series data set means that the corresponding value does not change with time. The mean of the data set remains constant. Augmented Dicky-Fuller (ADF) test is widely used to check the stationarity of the data. The WTP data have showed non stationarity property whereas the pumping station data have done the opposite of ADF test for both daily and 15-minute data sets. By assigning value to the degree of differencing (d), the non-stationary data can be made stationary. Seasonal decomposition is done to find any trend or seasonality. No data sets have showed any particular trend or seasonality. Hence, non-seasonal ARIMA is used to perform load forecasting.

Next, the autoregressive and moving average orders are needed to be selected for modeling. To find out the best suited autoregressive order and moving average order for modeling ARIMA, the Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) have been plotted for four datasets. ACF plots of all data sets are non-decaying functions and converge into 95% confidence interval after a lot of lags. This plot intends that a high autoregressive order is needed for modeling which is computationally expensive too. The PACF plots also converge after a high number of lags. This explains similarly that higher moving average orders are needed to perform ARIMA modeling. Higher orders of p,d,q to perform modeling are computationally expensive. Therefore, in order to find the best fit, ARIMA has been implemented for $0 \leq p,q \leq 3$ and the best fit has been chosen for this interval.

To find the best fit of the ARIMA model, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) determines the goodness of fit of the developed model [10-11].

$$\text{AIC}= -2\log(\text{maximum likelihood}) + 2k \tag{7}$$

where k is independently adjusted number of parameters within the model.

$$\text{BIC}= -2\log(\text{maximum likelihood}) + \frac{k\log(n)}{n} \tag{8}$$

where k is as stated above, and n is the number of samples used for modeling.

AIC is used to select the best ARIMA model here. ARIMA is an inappropriate model for very short term load forecasting such as 15 minute load forecasting. Hence, the ARIMA model of 15 minute data sets for given interval of p,d,q results in a very high value of AIC and large residuals. This leads to a large prediction error in turn. So, ARIMA is modeled here for daily load prediction only. AIC values for ARIMA model of daily load forecasting are showed in table I.

TABLE I. ARIMA MODEL SELECTION FOR DAILY LOAD FORECASTING

Account Name	**(p,d,q) order**	**AIC**
Treatment Plant	(1,1,1)	3431.9993
Pumping Station	(2,1,2)	2918.1322

C. *Model Selection for LSTM*

The LSTM model has been applied to the datasets using Keras. This modeling has been applied for one step ahead forecasting of utility load. Firstly, the data sets are made stationary by applying differencing operations. Then this sequential series is transformed to supervised problem for the LSTM method to be applied. To implement hyperbolic tangent function and sigmoid function, the data have been rescaled. Adam optimizer is used with default parameters of learning rate (0.001), exponential decay rates (0.9 and 0.999) and epsilon (10E-8) for Keras [12]. Adam optimizer is mostly used for non convex problems and the hyperparameters require less tuning. The batch size, no of layers and no of epochs have been tuned to find the best fit for the fitted model. Finally, inverse scaling is applied to convert the data into actual scale. Figure 3 shows the flow of modeling LSTM.

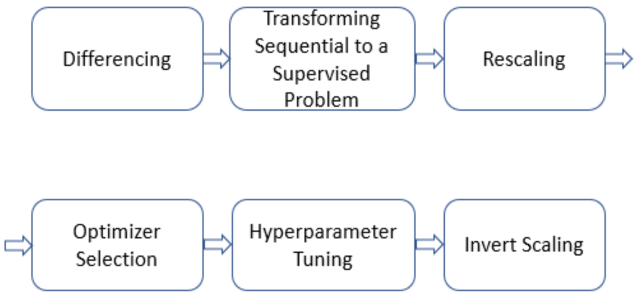

Fig 3: LSTM Modeling Flow Diagram

V. SIMULATION RESULTS

A. *ARIMA Modeling Results*

Figure 4 and 5 demonstrate the forecasted load values for the month of April along with the actual load values for the respective accounts. The daily load data for WTP is steady but does not follow any pattern. Our optimal ARIMA model for the given bound can not predict the data properly beforehand. Rather, it results in a steady load value after a few predicted values and shows that ARIMA modeling is inappropriate for the WTP data prediction.

On the other hand, the load of the pumping station fluctuates rapidly. The motors turn on and off rapidly. ARIMA modeling can't keep track of the sudden fluctuation of this load behavior and hence results in inappropriate forecasting. Absence of seasonality or pattern is another reason for this failure in forecasting.

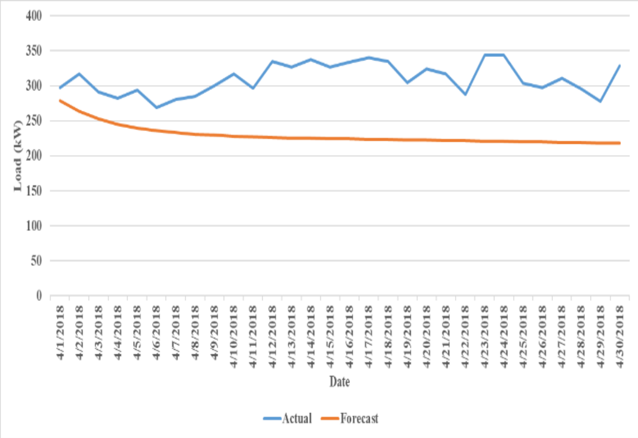

Fig 4: Forecasted Daily Load with ARIMA Modeling: WTP

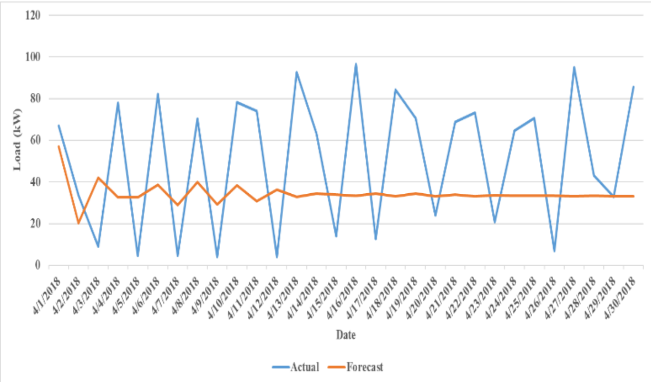

Fig 5: Forecasted Daily Load with ARIMA Modeling: PS

B. *LSTM Modeling Results*

Figure 6 and 7 are showing the LSTM forecasted daily load values for the month of April along with the actual load values for the respective accounts. Our optimal LSTM model can predict the WTP data properly beforehand. It captures the rise and fall of WTP data and results in a good prediction. For pumping station load, LSTM predicts better than ARIMA too. The ARIMA model can not follow the rapid fluctuation whereas LSTM follows the fluctuation in most of the cases. It gives poor results on only two occasions. Hence, it results in a good prediction.

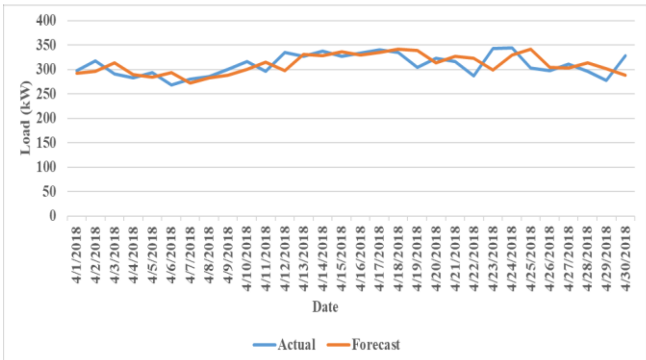

Fig 6: Forecasted Daily Load with LSTM modeling: WTP

Fig 7: Forecasted Daily Load with LSTM modeling: PS

Figure 8 and 9 are showing the 15 minute LSTM forecasted load values for the last day of April. LSTM works better for higher resolution and short term load prediction too whereas ARIMA can not produce better predictions for shorter periods.

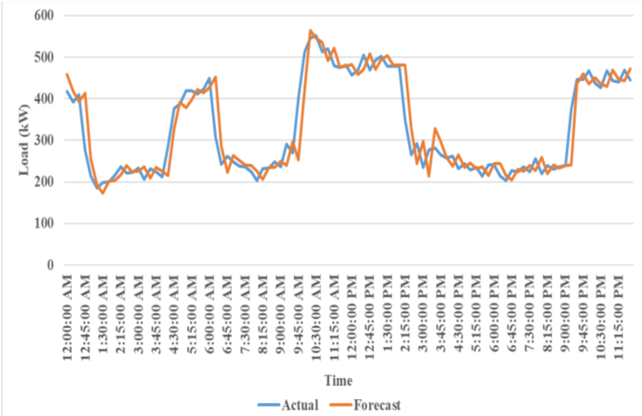

Fig 8: Forecasted 15 Minute Load with LSTM Modeling: WTP

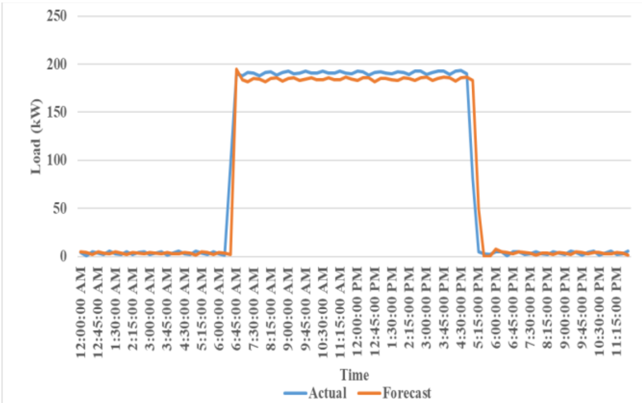

Fig 9: Forecasted 15 Minute Load with LSTM Modeling: PS

VI. Performance Evaluation

To evaluate the performance of both models for different time resolutions, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) metrics have been used.

$$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}(Predicted-Actual)^2}{n}} \tag{9}$$

$$\text{MAPE}=\left(\frac{1}{n}\sum_{i=1}^{n}\frac{|Actual-Predicted|}{|Actual|}\right)\times 100\% \tag{10}$$

RMSE (eq. 9) is a common metric to evaluate the performance as it gives larger weightage to the error calculation and penalizes the large errors. Lower RMSE values are preferable for good prediction. LSTM results in lower RMSE for daily load prediction of both data sets.

MAPE (eq. 10) is a metric to show the variances in data. RMSE values don't change for high variances in data sets. The WTP load does not have high variance unlike the PS data sets. Hence, both the modeling result in lower MAPE values for WTP daily load prediction. But the PS load results in higher MAPE values for large fluctuations in demand.

Table II shows the RMSE and MAPE values for the daily predicted load. LSTM results in lower RMSE and MAPE values for all cases. In case of RMSE, LSTM results in 75.5% error reduction for WTP and 17.3% reduction for PS. In case of MAPE, LSTM results in 78.4% error reduction for WTP and 10% reduction for pumping station.

For 15 minute resolution, ARIMA results in higher error metric values. Hence, it is incompatible with LSTM modeling for the same time resolution. For WTP data, LSTM results in 42.13 RMSE and 9.27% MAPE. It produces 15.64 RMSE and 46.7% MAPE for 15 minute resolution data of pumping station. 15-minute load forecasting works better for PS data in comparison to day ahead forecasting whereas daily forecasting is more preferable to WTP demand prediction. As, pumping station activities are fixed for a particular time period in a day, its load consumption does not usually depend on very earlier values. On the other hand, the nature of WTP load is dependent on older values.

TABLE II. PERFORMANCE EVALUATION FOR DAILY LOAD PREDICTION

Modeling	Accounts	RMSE	MAPE (%)
ARIMA	WTP	85.68	25.58
	PS	35.94	158.71
LSTM	WTP	20.96	5.51
	PS	29.71	142.9

VII. CONCLUSION

The electric load forecasting is important for large water utilities as unregulated pumping activities result in large electricity bills caused by high kW demand. Knowing the predicted values ahead of time, initiatives may be taken to reduce this high demand. This load forecast was done for both 15 minute and daily resolution of electric load. Electric utilities use highest 15-minute peak demand data for calculating monthly peak demand. Auto Regressive Integrated Moving Average (ARIMA) results in poor values and high residuals for higher data resolution and is not as accurate as Long Short Term Memory (LSTM) forecasting of the same resolution. LSTM can be handy for very short term load forecasting for water utilities. The electric load variation of water utilities does not have any resemblance to any other large electricity users. LSTM provides a 75.54% reduction in RMSE and 78.46% reduction in MAPE in comparison to ARIMA for water treatment plants. For pumping station, LSTM results in 17.33% lower RMSE and 9.96% lower MAPE values compared to ARIMA. As, the data sets do not show any specific trend or seasonality and highly non linear in nature, deep learning techniques and neural networks are likely to be more useful for accurate prediction.

REFERENCES

[1] C. Copeland, N. T. Carter, “Energy-Water Nexus: The Water Sector's Energy Use”, January 24, 2017,CRS Report

[2] Emmanuel A. Donkor, Thomas A. Mazzuchi; Refik Soyer; and J. Alan Roberson, “Urban Water Demand Forecasting: Review of Methods and Models”, Journal of Water Resources Planning and Management, Volume 140, Issue 2, February 2014

[3] Jean-Daniel Rinaudo “Long-Term Water Demand Forecasting”, 2015

[4] T. Panapongpakorn and D. Banjerdpongchai, "Short-Term Load Forecast for Energy Management Systems Using Time Series Analysis and Neural Network Method with Average True Range," *2019 First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP)*, Bangkok, Thailand, 2019, pp. 86-89.

[5] K. Goswami, A. Ganguly and D. A. Kumar Sil, "Comparing Univariate and Multivariate Methods for Short Term Load Forecasting," *2018 International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, Uttar Pradesh, India, 2018, pp. 972-976.

[6] H. Matsila and P. Bokoro, "Load Forecasting Using Statistical Time Series Model in a Medium Voltage Distribution Network," *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, Washington, DC, 2018, pp. 4974-4979.

[7] J. Yusuf, A. J. Hasan, S. Ula, “Successful Demand Management Strategies Development and Implementation for Large Water Districts”, Unpublished

[8] E. P. Box, G. M. Jenkins, G. C. Reinsel, “Time Series Analysis : Forecasting and Control”

[9] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory”, Neural Computation, vol. 9, no. 8, pp. 1735-1780

[10] H. Akaike, "A new look at the statistical model identification," in *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716-723, December 1974.

[11] G. Schwarz, “Estimating the dimension of a model,” Ann. Statist., vol. 6, no. 2, pp. 461–464, 03 1978.

[12] Available Online: github.com/keras-team/keras