

Accurate and time-efficient negative binomial linear model for electric load forecasting in IoE

Yousef-Awwad Daraghmi¹  | Eman Yaser Daraghmi² | Samer Alsaadi³  | Derar Eleyan⁴

¹Computer Systems Engineering Department, Palestine Technical University Kadoorie, Tulkarem, Palestine

²Applied Computing Department, Palestine Technical University Kadoorie, Tulkarem, Palestine

³Electrical Engineering Department, Palestine Technical University Kadoorie, Tulkarem, Palestine

⁴Applied Computing Department, Palestine Technical University Kadoorie/Birzeit University, Tulkarem, Palestine

Correspondence

Yousef-Awwad Daraghmi, Computer Systems Engineering, Palestine Technical University Kadoorie, Tulkarem, Palestine. Email: y.awwad@ptuk.edu.ps

Abstract

Accurate and efficient model predictive control (MPC) is essential for Internet of energy (IoE) to enable active real-time control, decentralized demand-supply balance, and dynamic energy management. The MPC consists of short-term electric load forecasting, whose accuracy is affected by the load characteristics, such as overdispersion, autocorrelation, and seasonal patterns. The forecasting efficiency depends on the computational time that is required to produce accurate results and is affected by the IoE data volume. Although several fundamental short-term forecasting models have been proposed, more accurate and efficient models are needed for IoE. Therefore, we propose a novel forecasting temporal negative binomial linear model (NBLM) that handles overdispersion and captures nonlinearity of electric load. We also classify the load into low, moderate, and high intraday seasons to increase the forecast accuracy by modeling the autocorrelation in each season, separately. The temporal NBLM was evaluated using real-world data from Jericho city, and its results were compared to other forecasting models. The temporal NBLM is found more accurate than the other models as the mean absolute percentage error (MAPE) is reduced by 29% compared to the ARMA model. In addition, the proposed model is more efficient as its running time is reduced by 63% in the training phase and by 87% in the forecast phase compared to the Holt-Winter model. This increase in accuracy and efficiency makes the proposed model applicable for load forecasting in IoE contexts where data volume is large and load is highly fluctuated, is overdispersed, is autocorrelated, and follows seasonal patterns.

1 | INTRODUCTION

The Internet of energy (IoE) aims at optimizing energy usage, improving the distribution of energy by enabling bidirectional transmission of green energy and providing efficient control methods.¹⁻³ An important application of IoE is balancing the supply and demand by, for example, allowing consumers to determine the peak period and switch their use of electric energy to other less demanded periods. On the other hand, prosumers are also capable of identifying the high demand so that they can push more power to the system. For balancing the supply and demand, accurate and efficient control strategies are required. A control approach that showed better accuracy and efficiency than other conventional methods is the model predictive control (MPC).^{1,3} In the MPC, the actual and future state is forecasted using historical data obtained by the large volume of measurements. For daily energy management and real-time control, short-term

forecasting can be used to evaluate control strategies before putting them in use.¹ More importantly, forecasting has to be accurate and efficient to enable an effective decision-making process within the suitable time.^{4,5}

To address accuracy, forecasting models have to deal with electric load characteristics that affect accuracy, such as nonlinearity,^{6,7} fluctuation,⁷⁻⁹ and temporal seasonal patterns.¹⁰⁻¹² The nonlinearity and the high fluctuations cause overdispersion meaning that the variance of the examined variable is greater than the mean. Clearly, electric load is overdispersed as it fluctuates depending on the consumption rate, place, day time, and life style. Overdispersion is critical in forecasting, and if it is not handled properly, it reduces the forecast accuracy and causes false judgments about the significance of the correlated variables.¹³ Another characteristic is the temporal autocorrelation and the seasonal pattern because the consumption of energy follows specific patterns that are repeated periodically and can be captured using historical data. Although several fundamental electric load forecasting models were proposed in the power systems domain as in the works of Laouafi et al.⁵ and Guan et al.,¹⁴ these models did not assiduously consider the overdispersion characteristic.^{15,16} In addition, if these models consider the overdispersion, they will consume much time and their efficiency will be decreased¹⁷; meanwhile, the IoE domain incorporates big data that require efficient and real time processing.^{4,5}

To address efficiency, any forecasting model should have low time complexity, which quantifies the time demand for running an algorithm and produce accurate results. Determining the time complexity of forecasting models assists in deciding whether these models can be applied online to real time load balancing solutions particularly when data volume is large. Although efficiency is crucial to the success of electric load control, few studies have considered the computation time of the forecasting models in the power systems domain.^{5,14,18} To reduce the computational time, these models focus on very short-term forecasting or utilize light forecasting models, but these models are sensitive to variations. According to literature, most electric load forecasting approaches adopt neural network models that require much computational time and their efficiency can still be improved.^{17,19}

In this paper, we address the forecasting accuracy and computational efficiency. To increase the accuracy, we adopt the negative binomial linear model (NBLM), which is derived from the negative binomial distribution for precisely handling overdispersion. We accommodate the NBLM to electric load data by handling the temporal correlation and classifying the seasonal patterns into low, moderate, and high load patterns. In addition, the NBLM can capture nonlinearity and fluctuation of data in low time complexity, which improves the forecasting efficiency. This can be achieved by using an efficient optimization method to estimate the correlation parameters and by applying statistical measures to determine the significant correlated values. The negative binomial-based models were successfully tested in our previous work on traffic data, which are autocorrelated, are overdispersed, and have seasonal patterns.²⁰⁻²² The negative binomial was also used to mine big data in social science, and it was found an accurate and efficient method.²³

The proposed model is evaluated using a real-world data set of mixed load collected from Jericho city. The NBLM is compared with other classical load forecasting exponential smoothing methods: autoregressive integrated moving average (ARIMA), autoregressive moving average (ARMA), and Holt-Winters (HW). The results show that the proposed NBLM model outperforms these models in terms of accuracy and efficiency, which makes the model applicable within the IoE architecture. In summary, this paper firstly contributes accurate and efficient temporal NBLM for forecasting electric load in different seasons mainly low, moderate, and high load conditions. This paper secondly emphasizes that overdispersion is an important characteristic of electric load and requires careful treatment by any forecasting model.

This paper is organized in six sections. Section 2 discusses the related work and shows the gap in the literature. Section 3 presents the proposed temporal NBLM. The application of the proposed model including the data set, the training, and testing of the model are shown in Section 4. Section 5 discusses the results of model, and finally, Section 6 concludes this paper and identifies the future work.

2 | RELATED WORK

Recently, several fundamental short-term electric load forecasting models have been developed to cope with the problems of the electric load forecasting and with the emerging technologies in the power system domain, such as smart grids and IoE. We do not provide extensive review of the short-term electric load forecasting as there are several papers in the literature discussing that.^{6,24} We instead aim to contribute an accurate and efficient forecasting model based on the gap in the literature. Because most papers in the field of electric load forecasting target the power systems mainly the smart grid, the main focus of this paper is the IoE domain, which is more advanced and has different requirements than other power systems domains.¹⁻³ The core requirement is the need to deal with the big data accurately and efficiently.

The electric load short-term forecasting models can be classified into statistical models and artificial intelligence-based models. The statistical-based models can be categorized as univariate and multivariate models. In the univariate models, the time series of load is modelled according to historical data. Examples of these models are HW, ARMA, and ARIMA.^{15,25,26} The multivariate models allows to include multiple explanatory variables that determine the behavior of the response variable, such as the regression models.^{6,27,28} In the statistical-based model, the electric load time series is ensured to be stationary and to have homogeneous variance by using differentiating transformations, such as Box-Cox. These kinds of transformations increase the computational time. In addition, the aforementioned statistical-based models assume equidispersion or depend on distributions that can not capture the high variation of electric load.

The artificial intelligence-based models include neural networks models,^{16,29} artificial networks models,^{30,31} support vector machine models,^{28,32} fuzzy algorithms,³³ and hybrid models.^{34,35} These models were proposed to deal with the nonlinearity of the historical data so that the accuracy can be increased. However, these methods require more historical data than the statistical-based model, they are very sensitive to irrelevant variables, they need many trials to tune the coefficients, and finally, they require much computational time.^{17,19}

The main challenge of the aforementioned models is dealing with the nonlinearity and high variance.¹⁰⁻¹² The nonlinearity and high variance cause overdispersion, which needs specific handling during the modeling so that the forecast accuracy can be increased.¹³ It is clear from the review that the statistical-based methods do not sufficiently address the nonlinearity of the electric load. In addition, the artificial intelligence-based models require much computational time for treating the nonlinearity while none has specifically accounted for the overdispersion problem. The high computational time reduces the efficiency, which was addressed in few studies.^{5,14,18} The low efficiency makes these forecasting models unsuitable for electric load forecasting in the IoE domain. Therefore, we focus in this paper on proposing a forecasting model for handling the overdispersion problem while maintaining high efficiency so that the model can be used in the IoE domain.

3 | THE PROPOSED MODEL

In this section, we firstly explain the theoretical background behind the proposed model. Then, we secondly show how the proposed model is derived based on the theoretical background to fit with the electric load data.

3.1 | Theoretical background

3.1.1 | Generalized linear models

The generalized linear models (GLMs) are extensions of multiple linear regression models. In multiple linear regressions, a dependent variable (response variable) is correlated with multiple variables for quantifying the relationship between that variable and other several independent variables (predictors). However, the linear regression models only work when the response variable and the predictors have a linear relationship. The GLMs handle nonlinear relationships by assuming a link linear relationship. To explain this, \mathbb{Y} is assumed a random variable that is correlated with other random variables $\mathbb{X}_1, \mathbb{X}_2, \dots, \mathbb{X}_r$. In multiple linear regression, \mathbb{Y} depends on other variables as

$$\mu = \mathbf{X}\boldsymbol{\beta}, \quad (1)$$

where μ is the mean of \mathbb{Y} , \mathbf{X} is the matrix of all predictors, and $\boldsymbol{\beta}$ is the matrix of regression coefficients. The GLM extends the linear relationship by

$$g(\mu) = \mathbf{X}\boldsymbol{\beta}, \quad (2)$$

where $g()$ is a monotonic function for transforming the μ_i to a unconstrained scale, such as $g(\mu) = \log(\mu)$ if $\mu_i > 0$ or $g(\mu) = \log[\mu/(1 - \mu)]$ if $0 < \mu_i < 1$. In this situation, the variables may follow distributions other than the normal distribution. That is because the normal approximation of response is accurate only when the variance is very small. However, for count data and nonnegative data, the variance usually is large and data have more variability. The GLMs allow modeling of variables that belong to nonlinear families like the exponential family.

3.1.2 | Negative binomial generalized linear model (NBLM)

The NB regression is one of the most famous modeling methods that can handle overdispersion. Other modeling methods, eg, the Poisson models have been utilized to model count data and assumes equidispersion, which means the variance and the mean are equal. equidispersion makes the Poisson-based methods inapplicable in the case of overdispersion, ie, the variance is greater than the mean.¹³ In addition, the NB reveals the cause-effect relationship among the variables and how they are correlated.¹³

In the NB regression that belongs to the generalized linear model (GLM) family, the mean μ of the response variable \mathbb{Y} is described by an exponential function of a predictor \mathbb{X} by $\mu = e^{\mathbb{X}\beta}$, where β is the regression coefficient.¹³ The link function $g()$ is selected as natural logarithm (\ln) for determining the best fit model. Therefore, when multiple predictors exist, the response variable \mathbb{Y} is correlated with a set of other random variables $\mathbb{X}_1, \mathbb{X}_2, \dots, \mathbb{X}_r$, and the equation of the regression can be written as

$$\ln y_i = \alpha + \sum_{j=1}^r x_{ij} \beta_{ij} + \varepsilon_i, \quad (3)$$

where α is the intercept and ε is the error, which is independent of all random variables and has a distribution in which the error mean = 1 and the error variance = $1/\varphi$. This model assumes that the predictors only affect the response variable. For handling the overdispersion, the NB employs a parameter φ , which makes the variance σ^2 of the dependent variable \mathbb{Y} equal to $\mu + \varphi\mu^2$.¹³

The best fit model is determined by selecting the predictors that are significantly correlated with the response variable, ie, have significant coefficients. In NB regression, a predictor is significant when its coefficient's P value is less than a significance level, eg, 5%.¹³ The significance level is used in step-wise backward or forward elimination algorithms to estimate the regression coefficients and eliminate insignificant predictors. The coefficients estimation process requires an optimization method such as Fisher scoring method, which performs the estimation based on maximizing a likelihood function that is given by

$$f_{NB}(\beta, \varphi) = \sum_{i=1}^n \left(y_i \ln \left(\frac{\varphi e^{x_i' \beta}}{1 + \varphi e^{x_i' \beta}} \right) - \frac{1}{\varphi} \ln \left(1 + \varphi e^{x_i' \beta} \right) + \ln \Gamma \left(y_i + \frac{1}{\varphi} \right) - \ln \Gamma(y_i + 1) - \ln \Gamma \left(\frac{1}{\varphi} \right) \right), \quad (4)$$

where Γ is the gamma function and n is the size of data (number of observations).¹³

The Fisher scoring is a numerical analysis method derived from Newton method and used to solve maximum likelihood function.³⁶ The Fisher scoring method converges more quickly than other numerical analysis methods on the same data set. This is because the Fisher scoring method replaces the Hessian matrix by the Fisher information matrix in which elements consist of the expectation value of the second derivatives of likelihood function toward each parameter.³⁶

However, optimization techniques such as the Fisher scoring method may demand high computational time that can be reach $O(n^3)$.³⁷ To minimize the computational time of coefficients estimation, we use the L-BFGS-B technique that will be shown in the proposed model section. For multivariate form, the log-likelihood function, which is derived in Solis-Trapala and Farewell,³⁸ can be used to simplify the estimation process. This log-likelihood function is given by

$$f_{NB}(\beta, \varphi) = \sum_{i=1}^r \left(\sum_{j=1}^n \beta_i x_{ij} \ln x_{ij} + x_{ij} \ln \varphi + \ln \Gamma \left(x_{ij} + \frac{1}{\varphi} \right) - \left(x_{ij} + \frac{1}{\varphi} \right) \ln \Gamma(\varphi \beta_i x_{ij} + 1) - \ln \Gamma \left(\frac{1}{\varphi} \right) \right), \quad (5)$$

where r is the number of predictors, n is the data size, and x_{ij} is the the value of the predictor i at time j .³⁸

Modeling large data set such as electrical load requires much computational time because data are big and have large variability. To minimize the computational demand, only the significant predictors will be considered in the final prediction model, which is the optimized NBLM. The final model is the best fit model because it contains only the significant predictors and fits the data with the minimum error. The Akaike information criterion (AIC) score and the P value are measures that can be used to determine the significance of a predictor.^{13,15} A predictor will be significant if P value is less than the level of significance, and the predictor significance will increase if its P value decreases.¹³ The AIC of the entire model including all significant predictors must be minimized. The AIC is given by

$$AIC = 2(par) - 2 \ln(\varphi(M_i)), \quad (6)$$

where M_i is the net model when a new significant predictor is added, $\phi(M_i)$ is maximum likelihood value corresponding to the model M_i , and par is the number of estimated parameters.

The forecasting model requires a model matrix \mathbf{X} to contain the data and a model matrix for the coefficient β . The data are the input observations of predictors that are used to forecast the future values of the response. The data set is used to compute the forecasting model matrix \mathbf{X}^p . The forecasting matrix maps the model coefficients β to the forecast of the response (\check{Y}) as

$$\check{Y} = \mathbf{X}^p \beta, \quad (7)$$

where the coefficient matrix β is obtained during the training phase.¹³ The computational time of the NBLM is expressed by a time complexity function $O(n)$, which shows that the time mainly depends on the data size of the predictors.¹³

3.2 | The proposed temporal NBLM

Our goal is to determine a forecast model that can address the electric load correlation. Therefore, we formulate the model that can deal with the important characteristics of the electric load that are overdispersion and autocorrelation.

We firstly focus on the autocorrelation. Let i be the station (transformer) number and let $y_{i,t}$ be the response variable corresponding to electric current transformer i at time t . The autocorrelation is captured by allowing the electrical current at each transformer to correlate with itself at previous time segments. In this approach, we can focus on the entire temporal autocorrelation, $t - 1, t - 2, \dots, t - n$. We let T the set of all previous time segments at which data were observed $(1, 2, \dots, n)$. The temporal NBLM regression can be written as

$$\ln y_{i,t} = \alpha + \sum_{k \in T} \beta_{i,t-k} y_{i,t-k} + \varepsilon_t. \quad (8)$$

We use the term “temporal” in this paper instead of the autocorrelation term because the model can refer to specific time segment in the historical data. These time segments can be determined at the training stage by identifying the significant time segments. In other words, temporal means that the model at time t can correlate with any leading data at time $t - j$. We do not use the term autocorrelation to describe our model because autocorrelation means that the model always correlates with the previous time segment.

The proposed model handles overdispersion by allowing the variance, σ^2 , to equal $\mu + \phi\mu^2$. It also captures the temporal autocorrelation by incorporating previous time segments in the correlation. The log-likelihood function used by Hilbe¹³ has good results when the number of predictors is small. To enable incorporating multiple predictors, we use the log-likelihood function shown in Equation 5.

The complexity of the proposed model depends on the L-BFGS-B optimization technique, which is an extension of the limited-memory BFGS (Broyden-Fletcher-Goldfarb-Shanno). The L-BFGS-B algorithm requires a the negative of the log-likelihood function to be minimized. We used the negative because the log-likelihood function should be maximized not minimized. The L-BFGS-B also requires a lower bound and an upper bound for the coefficients, which are -1 and 1 , respectively. The computational time and space of the L-BFGS-B algorithm can be maintained in the linear order, ie, $O(n)$ where n is the data size.³⁹ These little computational time and memory demands make the proposed temporal NBLM an efficient model for forecasting future electric load values.

Algorithm 1 of the proposed model works first by defining the input values, which are the response variable y at time t and the predictors $x_{t-1}, x_{t-2}, \dots, x_{t-n}$. We then define the model as an empty model at the beginning and α as the mean value of the response variable. We also set j as an indicator to the predictors at previous time segments. The core of the algorithm is finding the significant predictors and their coefficients, which will be used in the forecasting. The for loop examines the correlation of y with other predictors by finding β of that predictor, its P value, and the AIC . If the P value of the predictor is less than the significance level (0.05), we check the AIC of the model. First, at $j = 1$ and when the model is empty, the AIC of the model AIC_m is set to the computed AIC . Then, the model is updated and the predictor is added to the model if the new AIC computed for the new predictor at the new j is less than the AIC_m . The algorithm returns the final model at the end.

Algorithm 1 ModelTraining (response variable and predictors)**Require:** y_i // electric load at time t $x_{t-1}, x_{t-2}, \dots, x_{t-n}$ // electric load at temporal-previous time segments**Ensure:** $M = \emptyset$ // model initialization $\alpha = \text{Mean}(y_i)$ set j to predictor index**for** $j = 1$ to n **do** Compute β_j // by maximizing log-likelihood function Compute P-value of β_j , AIC **if** P-value of $\beta_j < 0.05$ **then** //a predictor is significant when its coefficient's P-value is less than a significance level 0.05 **if** $|M| = 0$ **then** $AICm = AIC$ $M = \{j\}$ **else** **if** $AICm > AIC$ **then** $M = M \cup \{j\}$ $AICm = AIC$ **end if** update model $\ln y_i = \alpha + \sum_{j \in M} \beta_j x_j$ **end if** **end if****end for** **return** model $\ln y_i = \alpha + \sum_{j \in M} \beta_j x_j$

4 | APPLICATION OF THE MODELS

In this section, we present how the proposed model is applied to real data. We also compare the proposed model with other models using the same data set.

4.1 | Electrical load in Jericho city

The data consist of electrical current recorded hourly from 15 transformers in Jericho city between January to December 2015. A total of 4380 records were collected for each transformer. Figure 1 shows the electric current pattern for 17 consequent days. Initially, the data are treated by identifying invalid records resulted from the malfunction of the measurement devices. The percentage of the invalid records is not large and does not affect the accuracy of the model because the NB-based models are not affected by small number of missing data.¹³ For example, each transformer has less than 100 invalid records (2%) per day. The invalid data records are treated by replacing each invalid record by an average value calculated using an interpolation function. We identify the time of the invalid record; then, we calculate the average of loads measured at the same time from the previous weekdays.

4.2 | Electric current characteristics

The electric load of selected stations during 17 days is represented by a time series. Figure 1 shows the time series for 17 days and Figure 2 shows the time series for a single day. Figure 1 and Figure 2 illustrate that the data have a daily seasonal pattern, which is repeated periodically. This daily pattern contains different seasons, which can be categorized as (1) a low-load season exists in the early morning from 02:00 to 10:00 (8 hours) when the current is less than the daily mean; (2) a high-load season exists from 18:00 to 01:00 when the current is greater than the daily mean (8 hours); (3) a moderate load season exists in the durations 01:00-02:00 and 10:00-18:00 (9 hours) when the current is around the mean.

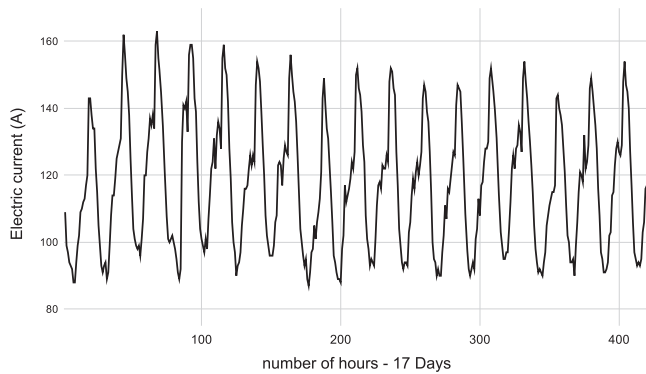


FIGURE 1 The electric current over 17 days from November 13, 2015, to November 29, 2015, in the Jericho Aqabat Jaber transformer

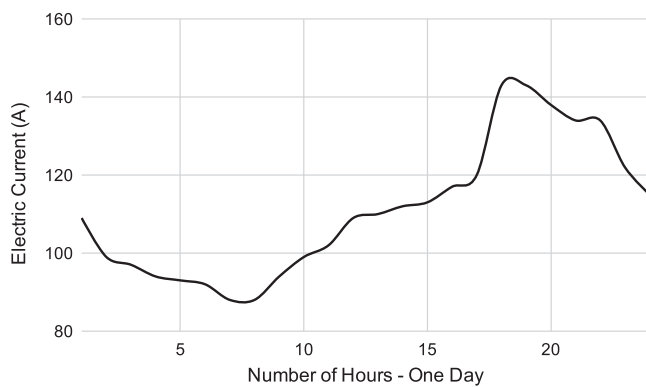


FIGURE 2 The electric current in November 13, 2015, in the Jericho Aqabat Jaber transformer

| Road segment | Aqabat Jaber | Sea level 1 | Sea level 2 | Sea level 3 |
|----------------------------|--------------|-------------|-------------|-------------|
| Daily mean μ | 117.9 | 108.7 | 64.0 | 182.3 |
| Daily Var σ^2 | 381.8 | 194.8 | 163.5 | 352.8 |
| Daily dispersion | 2.3 | 2.6 | 1.8 | 2.4 |
| Daily max | 163 | 140 | 100 | 224 |
| Daily min | 87 | 74 | 44 | 137 |
| High season μ | 153.0 | 149.2 | 87.3 | 205.9 |
| High season σ^2 | 168.1 | 181.4 | 127.7 | 327.5 |
| High dispersion | 2.4 | 2.3 | 2.5 | 2.8 |
| Moderate season μ | 125.7 | 127.6 | 68.3 | 189.3 |
| Moderate season σ^2 | 143.1 | 155.3 | 98.0 | 242.6 |
| Moderate dispersion | 2.7 | 2.8 | 1.9 | 1.9 |
| Low season μ | 101.8 | 97.8 | 46.1 | 151.4 |
| Low season σ^2 | 117.2 | 152.6 | 77.3 | 197.0 |
| Low dispersion | 1.9 | 2.5 | 1.7 | 2.0 |

TABLE 1 The statistical parameters for the four selected transformers in the three load seasons

Electric load in the data shown in Figure 1 and Figure 2 is temporally autocorrelated and is not stationary because each season has different length, and within each season the mean and variance have different values and depend on the time of the day. These characteristics of electric load are common and can be found in electric network worldwide.

Additionally, the electric current has overdispersion because, as shown in Table 1, the variance of the current at each station is greater than the mean. Overdispersion is determined when the value of dispersion, which is the Pearson statistic (χ^2) divided by the degrees of freedom is larger than one.¹³ In the data set, all load are overdispersed because the dispersion values are greater than one.

The overdispersion in the electrical current is related to the consumption variations within seasonal daily patterns. The variations are caused by changing weather conditions, time of the day, place within the city (industry, urban, market), and life style.

4.2.1 | Training the temporal NBLM

To train the proposed models for the selected transformers, we use data of 10 months recorded from January 2015 to October 2015. The model for each load season is trained separately using ten-month long three training data sets that are a set for low load season from 02:00 to 10:00 (8 hours), a set for moderate load season from 10:00-15:00 and 00:00 to 02:00 (7 hours), and a set for high load season from 15:00 to 00:00 (9 hours). Therefore, each season has a different forecast model. The advantages of this approach is eliminating the outliers and reducing the amount of variability and irregularity as the electric current does not fluctuate between peak and off-peak. If a single NB-based model is used for the entire day, the load condition in the larger period will be given more weight, which causes the forecasting model to be accurate only in the large period.

In all analyses, the statistical confidence interval is set to 95% so the significance level is 5%. The temporal correlation is modeled using the stepwise forward elimination algorithm. This means the coefficients are estimated at time segments $t-1$, then $t-2$, to $t-n$, and the estimation stops when the temporal correlations becomes insignificant. In our estimation, we stop at $t-6$ as shown in the Table 2, which presents the detailed temporal correlation coefficients for the Aqabat Jaber transformer.

Table 2 shows the coefficients corresponding to each time segment. The segments with P values that are larger than the significance level (0.05) are insignificant, and the significance increases when the P value decreases. We notice that the current load data is negative-Binomially and temporally correlated with the previous five segments, which are the historical data one hour ago to five hours ago.

4.2.2 | Training the models of comparison

To evaluate the proposed model, we compare it with classical forecasting methods used in in this domain. We mainly focus on the ARMA model as in the work of Boroojeni et al.,¹⁵ the ARIMA as in the work of Williams and Hoel,⁴⁰ and the HW as in the work of Taylor.⁴¹

The HW model is trained to estimate the coefficients of the level, trend, and seasonality using a daily pattern, which is one day cycle of length 24. The coefficients are shown in Table 3 for three transformers.

The coefficients can be interpreted to mean that the level of data is positive and has a coefficient equal to 0.57. The overall trend of the load is constant because the trend coefficient is zero. The high seasonal coefficient expresses that there are seasonal patterns expressed by one day season with a coefficient =0.73. We also examined one week cycle (weekly pattern), but this pattern has lower accuracy than the daily pattern.

The ARIMA model coefficients are estimated using one day cycle of length 24. Seasonal differencing is used to remove nonstationarity. The ARIMA coefficients can form different variations and each variation has a different AIC value. The ARIMA model usually follows the form $ARIMA(p, d, q)$, where p is the number of autoregressive terms, d is the number of nonseasonal differences needed for stationarity, and q is the number of lagged forecast errors in the prediction equation. The best fit model is formed by the variation, which is $ARIMA(2, 0, 2)_{24}$ because it has the lowest AIC and the lowest mean square error. The coefficients of the ARIMA are shown in Table 4 for three transformers.

The variation $ARIMA(2, 0, 2)_{24}$ means the model incorporates two nonseasonal autoregression (AR) components, two nonseasonal moving average (MA), and one seasonal difference when the model is trained for 24 hours season. For

TABLE 2 The regression coefficients of the temporal correlation for the Aqabat Jaber transformer

| Intercepts | Low load | | Moderate load | | High load | |
|-----------------------|----------|------------------------|---------------|------------------------|-----------|------------------------|
| | β | P value | β | P value | β | P value |
| Autocorrelation $t-1$ | 3.37 | 4.3×10^{-12} | 3.45 | 7.31×10^{-12} | 6.35 | 9.43×10^{-12} |
| Autocorrelation $t-2$ | 3.10 | 3.32×10^{-12} | 5.34 | 4.55×10^{-12} | 0.82 | 5.7×10^{-12} |
| Autocorrelation $t-3$ | 8.20 | 5.2×10^{-9} | 3.44 | 5.8×10^{-9} | 9.32 | 2.31×10^{-9} |
| Autocorrelation $t-4$ | 0.883 | 5.4×10^{-5} | 0.649 | 7.4×10^{-5} | 0.56 | 8.1×10^{-5} |
| Autocorrelation $t-5$ | 0.044 | 6.3×10^{-4} | 0.071 | 7.2×10^{-4} | 0.084 | 6.2×10^{-4} |
| Autocorrelation $t-6$ | 0.011 | 0.156 | 0.371 | 0.097 | 1.270 | 0.125 |
| No. of observations | 2432 | | 2128 | | 2736 | |
| Overdispersion ϕ | 3.63 | | 2.33 | | 3.18 | |

TABLE 3 The coefficients of the Holt-Winters model

| Coefficient | Aqabat Jaber | Sea level 1 | Sea level 2 |
|-------------|--------------|-------------|-------------|
| Level | 0.57 | 0.63 | 0.67 |
| Trend | 0 | 0 | 0 |
| Season | 0.73 | 0.79 | 0.86 |

| Coefficient | AR1 | AR2 | MA1 | MA2 | AIC |
|--------------|-------|-------|--------|-------|--------|
| Aqabat Jaber | 0.506 | 0.045 | -1.964 | 1.0 | 432.64 |
| Sea level 1 | 0.593 | 0.181 | -1.981 | 1.0 | 380.65 |
| Sea level 2 | 1.136 | 0.144 | -0.058 | -0.94 | 355.85 |

TABLE 4 The autoregressive integrated moving average coefficients

Abbreviation: AIC, Akaike information criterion.

example, for the station Aqabat Jaber, the two autoregressive components can be written as AR(2): $y_t = 0.506y_{t-1} + 0.045y_{t-2}$ and the error term for the moving average can be written as MA(2): $\epsilon_t = -1.964\epsilon_{t-1} + 1.1\epsilon_{t-2}$.

The ARMA is used after homogenizing the variance and eliminating the nonstationary as in the work of Boroojeni et al¹⁵ but without annual seasonality. The smallest AIC is achieved when the number of daily autoregressive terms and weekly moving average terms are two and seven, respectively.

4.2.3 | Forecasting results

To build the forecasting model, we use the coefficients produced in the training sessions. For example, the coefficients in Table 2 is used to produce the Aqabat Jaber load forecasting for multiple steps ahead: from one step to 24 steps, and each step is one hour. Because the coefficients are different, each load season has a different forecasting model. Forecasting results for other transformers are also shown in Table 5 and Table 6. We compare our model with the HW, ARMA, and ARIMA for a 10-hour prediction horizon. We use a single forecasting model for HW, ARMA, and ARIMA methods during all load seasons because these methods are cyclic and require a repeated pattern. We use the root mean square error (RMSE) and the mean absolute percentage error (MAPE) to compare the accuracy of the forecasting.

| Traffic condition | Method | RMSE Aqabat Jaber | RMSE Sea level 1 | RMSE Sea level 2 |
|---------------------|---------------|-------------------|------------------|------------------|
| Low load season | Temporal NBLM | 5.47 | 3.22 | 2.12 |
| | HW | 7.33 | 5.78 | 4.67 |
| | ARMA | 6.73 | 4.37 | 3.39 |
| | ARIMA | 8.24 | 6.37 | 6.02 |
| Average load season | Temporal NBLM | 6.22 | 4.44 | 3.07 |
| | HW | 7.51 | 7.96 | 5.98 |
| | ARMA | 7.03 | 6.11 | 4.92 |
| | ARIMA | 10.51 | 9.66 | 7.74 |
| High load season | Temporal NBLM | 7.89 | 6.08 | 4.12 |
| | HW | 11.4 | 9.77 | 7.22 |
| | ARMA | 8.45 | 7.50 | 5.93 |
| | ARIMA | 12.45 | 11.21 | 9.50 |

TABLE 5 The root mean square error (RMSE) values for different models during the three load seasons

Abbreviations: ARMA, autoregressive moving average; HW, Holt-Winters; NBLM, negative binomial linear model.

| Traffic condition | Method | MAPE, % Aqabat Jaber | MAPE, % Sea level 1 | MAPE, % Sea level 2 |
|---------------------|---------------|----------------------|---------------------|---------------------|
| Low load season | Temporal NBLM | 1.34 | 1.68 | 1.41 |
| | HW | 3.25 | 3.51 | 3.52 |
| | ARMA | 2.11 | 2.97 | 2.85 |
| | ARIMA | 6.20 | 5.31 | 4.65 |
| Average load season | Temporal NBLM | 3.68 | 2.05 | 2.61 |
| | HW | 5.17 | 3.36 | 4.12 |
| | ARMA | 4.84 | 2.91 | 3.70 |
| | ARIMA | 6.91 | 4.58 | 5.94 |
| High load season | Temporal NBLM | 4.64 | 2.41 | 2.96 |
| | HW | 5.77 | 3.84 | 4.08 |
| | ARMA | 5.24 | 3.25 | 3.48 |
| | ARIMA | 8.15 | 5.66 | 5.66 |

TABLE 6 The mean absolute percentage error (MAPE) values for different models during the three load seasons

Abbreviations: ARMA, autoregressive moving average; HW, Holt-Winters; NBLM, negative binomial linear model.

FIGURE 3 The prediction of electric load for the 10-hour horizon during the low-load season in the Jericho Aqabat Jaber transformer. ARMA, autoregressive moving average; HW, Holt-Winters; NBLM, negative binomial linear model

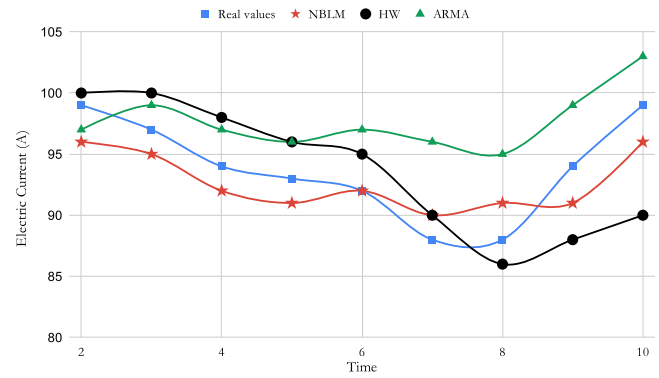


FIGURE 4 The prediction of electric load for the 10-hour horizon during the moderate-load season in the Jericho Aqabat Jaber transformer. ARMA, autoregressive moving average; HW, Holt-Winters; NBLM, negative binomial linear model

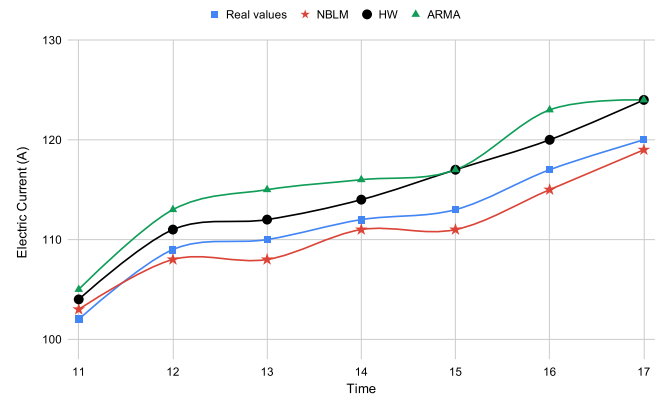


FIGURE 5 The prediction of electric load for the 10-hour horizon during the high-load season in the Jericho Aqabat Jaber transformer. ARMA, autoregressive moving average; HW, Holt-Winters; NBLM, negative binomial linear model

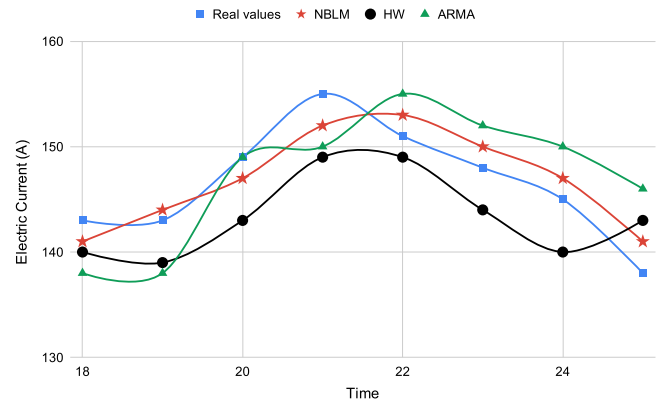


Table 5 and Table 6 present RMSE and the MAPE of the models for examined transformers during the three seasons for a 10-hour horizon. The results of three models during the three load seasons are shown in Figure 3, Figure 4, and Figure 5.

4.2.4 | Time complexity results

We also measure the computational demand of the proposed model. We selected the high load season because it has the biggest data volume, which is 2736. The model is tested on Intel CPU of 2.8 GHz, 64-bit operating system, and 16 GB of RAM. We notice that the NBLM is faster than the others and requires little time during the training stage for optimizing the NBLM regression coefficients and during the forecasting stage, as shown in Table 7.

As our data size is not so big and collected from small scale electrical network, we multiplied the size of data by a factor $F=(10, 20, 30, \dots, 1000)$ to simulate the big load data and test the behavior of the model at these data size. We found that the NBLM model still outperforms others.

| Model | Training (s) | Forecasting (s) |
|-------|--------------|-----------------|
| NBLM | 2.4 | 0.1 |
| ARMA | 6.4 | 3.2 |
| ARIMA | 7.2 | 4.8 |
| HW | 6.4 | 0.8 |

TABLE 7 The measured computation time for different models during the training and forecasting stages

Abbreviations: ARIMA, autoregressive integrated moving average; ARMA, autoregressive moving average; HW, Holt-Winters; NBLM, negative binomial linear model.

5 | DISCUSSION

In Table 5 and Table 6, the RMSE and MAPE values of the proposed model are less than HW, ARMA, and ARIMA, which implies that the proposed model is more accurate than the others. The plots in Figure 3, Figure 4, and Figure 5 also emphasize that the temporal NBLM regression has better performance than the other models. We also notice that the number of the missing values and the way we treated them do not affect the accuracy of the temporal NBLM.

As stated in other works,⁷⁻⁹ handling the nonlinearity and the high variance of the electric load by the forecasting model gives more accurate results. This is clear in our results, as the proposed temporal NBLM handles the overdispersion, ie, the high variance, and therefore, its forecasting accuracy is high. On the other hand, the HW, ARMA, and ARIMA do not handle the overdispersion, and consequently, they have lower accuracy than the proposed model. Although the ARIMA and ARMA model were modified by removing the nonstationary, the proposed model still outperforms these models.

The temporal characteristic of the proposed model allows to incorporate significant historical data. In the used data set, the previous five segments (five hours) are used in the forecasting because only these hours were identified as significant. The temporal correlation with the previous five hours is enough to produce high accuracy because the daily load season was divided into three intraday seasons that are low, moderate, and high electric load. By this, the variance can be guaranteed to be lower than the variance of a complete daily season, as in Table 1, and therefore, the forecast accuracy is increased. The other models still do not perform well even during each load season.

Regarding the computational time, the proposed temporal NBLM also outperforms the ARMA, ARIMA, and HW during the training stage and the forecasting stage. The important point is that the training stage, which costs more time than the forecasting, depends on the data during the load season and every load season has different coefficients. It is found that the NBLM learns the pattern in each season quickly particularly because the variance within one season is lower than one day variance as in Table 1. In addition, the forecasting stages are not largely affected by the data size because the coefficients catch the pattern by referring to the previous five time segments. This makes the forecast run quickly with low computation demand, in contrast to the other models, which require complete seasons (cycles) of data for the forecast. In a real-time scenario, the model maintains adjusting its coefficients to include the previous five hours and uses the data in these hours for the forecasting.

6 | CONCLUSION

Internet of energy (IoE) requires MPC to allow dynamic and distributed balance of electric supply and demand. The short-term load forecasting is a major component of the MPC and it has to be accurate and efficient to achieve effective energy management. Accuracy and efficiency are affected by the characteristics of electrical load, such as temporal autocorrelation and seasonal patterns. Additionally, electrical load exhibits overdispersion that is caused by the high and rapid fluctuation in the electrical network and consumption rates. Overdispersion affects the forecast accuracy and requires special treatments during building the forecast model. Therefore, we have proposed a short-term forecasting model that is based on NB, which handles overdispersion. The model can also capture the trend of the load by including the autocorrelation component in the explanatory variables. To address the seasonal pattern, we divide one day load season into low, moderate, and high load. A different temporal NBLM is used for each load season. The proposed model is accurate and efficient during all electrical load seasons as it produces accurate results in small computation time, and it outperforms other models used in this field.

In the future work, the temporal NBLM will be used as a load forecasting component in the microgrid, which will be built in Jericho. The microgrid will be based on IoE and IoT technologies. For achieving the best use of IoT, we will follow the approach in other works.⁴²⁻⁴⁴ A cloud-based system will be built to ensure the success of the big data processing and hosting. The cloud system also allows other stakeholders mainly electric companies to participate in the system and utilize the smart services as in other works.⁴⁵⁻⁴⁷ By the IoE system, utilities will be able to manage and operate all smart device-related processes and data acquisition systems, communicate with all grid devices, receive problem notifications, and manage new metering points. The security of the system can not be neglected so we will employ different approaches to achieve secure cloud and IoT services as in the work of Ghafir et al.⁴⁸ These developments will also motivate us to investigate other forecasting-related issues, such as the spatial correlation of one station with other stations, the micro-scale forecasting for small urban areas, and the effect of other conditions, eg, weather, on the forecast. We will utilize machine learning methods derived for abnormal conditions as in the work of Khalaf et al.⁴⁹ and Farhan et al.⁵⁰

ORCID

Yousef-Awwad Daraghmi  <https://orcid.org/0000-0001-9126-376X>

Samer Alsaadi  <https://orcid.org/0000-0002-7310-7269>

REFERENCES

1. Siozios K, Anagnostos D, Soudris D, Kosmatopoulos E. *IoT for Smart Grids: Design Challenges and Paradigms*. Cham, Switzerland: Springer; 2019.
2. Kalyanaraman S. Back to the future: lessons for Internet of Energy networks. *IEEE Internet Comput*. 2016;20(1):60-65.
3. Hannan MA, Faisal M, Ker PJ, et al. A review of Internet of Energy based building energy management systems: issues and recommendations. *IEEE Access*. 2018;6:38997-39014.
4. Zjavka L, Snášel V. Short-term power load forecasting with ordinary differential equation substitutions of polynomial networks. *Electr Power Syst Res*. 2016;137:113-123. <https://doi.org/10.1016/j.epsr.2016.04.003>
5. Laouafi A, Mordjaoui M, Haddad S, Boukelia TE, Ganouche A. Online electricity demand forecasting based on an effective forecast combination methodology. *Electr Power Syst Res*. 2017;148:35-47. <https://doi.org/10.1016/j.epsr.2017.03.016>
6. Fan S, Hyndman RJ. Short-term load forecasting based on a semi-parametric additive model. *IEEE Trans Power Syst*. 2012;27(1):134-141.
7. Quilumba FL, Lee W-J, Huang H, Wang DY, Szabados RL. Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities. *IEEE Trans Smart Grid*. 2015;6(2):911-918.
8. Wang Y, Chen Q, Sun M, Kang C, Xia Q. An ensemble forecasting method for the aggregated load with subprofiles. *IEEE Trans Smart Grid*. 2018;9(4):3906-3908.
9. Li B, Zhang J, He Y, Wang Y. Short-term load-forecasting method based on wavelet decomposition with second-order gray neural network model combined with ADF test. *IEEE Access*. 2017;5:16324-16331.
10. Da Silva PG, Ilić D, Karnouskos S. The impact of smart grid prosumer grouping on forecasting accuracy and its benefits for local electricity market trading. *IEEE Trans Smart Grid*. 2014;5(1):402-410. <https://doi.org/10.1109/TSG.2013.2278868>
11. Kwac J, Flora J, Rajagopal R. Household energy consumption segmentation using hourly data. *IEEE Trans Smart Grid*. 2014;5(1):420-430. <https://doi.org/10.1109/TSG.2013.2278477>
12. Chaouch M. Clustering-based improvement of nonparametric functional time series forecasting: application to intra-day household-level load curves. *IEEE Trans Smart Grid*. 2014;5(1):411-419. <https://doi.org/10.1109/TSG.2013.2277171>
13. Hilbe JM. *Negative Binomial Regression*. 2nd ed. Cambridge, UK: Cambridge University Press; 2011.
14. Guan C, Luh PB, Michel LD, Chi Z. Hybrid Kalman filters for very short-term load forecasting and prediction interval estimation. *IEEE Trans Power Syst*. 2013;28(4):3806-3817. <https://doi.org/10.1109/TPWRS.2013.2264488>
15. Boroojeni KG, Amini MH, Bahrami S, Iyengar SS, Sarwat AI, Karabasoglu O. A novel multi-time-scale modeling for electric power demand forecasting: from short-term to medium-term horizon. *Electr Power Syst Res*. 2017;142:58-73. <https://doi.org/10.1016/j.epsr.2016.08.031>
16. Khwaja AS, Zhang X, Anpalagan A, Venkatesh B. Boosted neural networks for improved short-term electric load forecasting. *Electr Power Syst Res*. 2017;143:431-437. <https://doi.org/10.1016/j.epsr.2016.10.067>
17. Bertoni A, Palano B. Structural complexity and neural networks. In: Marinaro M, Tagliaferri R, eds. *Neural Nets*. Berlin, Germany: Springer; 2002:190-216. *Lecture Notes in Computer Science*; vol. 2486.
18. Khan ZA, Jayaweera D, Alvarez-Alvarado MS. A novel approach for load profiling in smart power grids using smart meter data. *Electr Power Syst Res*. 2018;165:191-198. <https://doi.org/10.1016/j.epsr.2018.09.013>
19. Massana J, Pous C, Burgas L, Melendez J, Colomer J. Short-term load forecasting in a non-residential building contrasting models and attributes. *Energy Build*. 2015;92:322-330. <https://doi.org/10.1016/j.enbuild.2015.02.007>
20. Daraghmi Y-A, Yi C-W, Chiang T-C. Space-time multivariate negative binomial regression for urban short-term traffic volume prediction. In: Proceedings of the 12th International Conference on ITS Telecommunications; 2012; Taipei, Taiwan.

21. Daraghmi Y-A, Yi C-W, Chiang T-C. Negative binomial additive models for short-term traffic flow forecasting in urban areas. *IEEE Trans Intell Transp Syst.* 2014;15(2):784-793.
22. Daraghmi Y-A, Yi C-W, Chiang T-C. Mining overdispersed and autocorrelated vehicular traffic volume. In: Proceedings of the 5th International Conference on Computer Science and Information Technology; 2013; Amman, Jordan.
23. Wang H, Kifer D, Graif C, Li Z. Crime rate inference with big data. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16); 2016; San Francisco, CA.
24. Raza MQ, Baharudin Z. A review on short term load forecasting using hybrid neural network techniques. In: Proceedings of the 2012 IEEE International Conference on Power and Energy (PECon); 2012; Kota Kinabalu, Malaysia.
25. Taylor JW, McSharry PE. Short-term load forecasting methods: an evaluation based on European data. *IEEE Trans Power Syst.* 2007;22(4):2213-2219. <https://doi.org/10.1109/TPWRS.2007.907583>
26. Taylor JW. Short-term load forecasting with exponentially weighted methods. *IEEE Trans Power Syst.* 2012;27(1):458-464. <https://doi.org/10.1109/TPWRS.2011.2161780>
27. Goude Y, Nedellec R, Kong N. Local short and middle term electricity load forecasting with semi-parametric additive models. *IEEE Trans Smart Grid.* 2014;5(1):440-446. <https://doi.org/10.1109/TSG.2013.2278425>
28. Ghelardoni L, Ghio A, Anguita D. Energy load forecasting using empirical mode decomposition and support vector regression. *IEEE Trans Smart Grid.* 2013;4(1):549-556. <https://doi.org/10.1109/TSG.2012.2235089>
29. Quan H, Srinivasan D, Khosravi A. Short-term load and wind power forecasting using neural network-based prediction intervals. *IEEE Trans Neural Netw Learn Syst.* 2014;25(2):303-315. <https://doi.org/10.1109/TNNLS.2013.2276053>
30. Cecati C, Kolbusz J, Różycki P, Siano P, Wilamowski BM. A novel RBF training algorithm for short-term electric load forecasting and comparative studies. *IEEE Trans Ind Electron.* 2015;62(10):6519-6529. <https://doi.org/10.1109/TIE.2015.2424399>
31. Saviozzi M, Massucco S, Silvestro F. Implementation of advanced functionalities for distribution management systems: load forecasting and modeling through artificial neural networks ensembles. *Electr Power Syst Res.* 2019;167:230-239. <https://doi.org/10.1016/j.epsr.2018.10.036>
32. Zhang X, Wang J, Zhang K. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by cuckoo search algorithm. *Electr Power Syst Res.* 2017;146:270-285. <https://doi.org/10.1016/j.epsr.2017.01.035>
33. Coelho VN, Guimarães FG, Reis AJR, Coelho IM, Coelho BN, Souza MJF. A heuristic fuzzy algorithm bio-inspired by evolution strategies for energy forecasting problems. In: Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE); 2014; Beijing, China.
34. Nazar MS, Fard AE, Heidari A, Shafie-khah M, Catalão J. Hybrid model using three-stage algorithm for simultaneous load and price forecasting. *Electr Power Syst Res.* 2018;165:214-228. <https://doi.org/10.1016/j.epsr.2018.09.004>
35. Ghayekhloo M, Menhaj MB, Ghofrani M. A hybrid short-term load forecasting with a new data preprocessing framework. *Electr Power Syst Res.* 2015;119:138-148. <https://doi.org/10.1016/j.epsr.2014.09.002>
36. Longford NT. A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika.* 1987;74(4):817-827.
37. Ksantini R, Ziou D, Colin B, Dubeau F. Weighted pseudometric discriminatory power improvement using a Bayesian logistic regression model based on a variational method. *IEEE Trans Pattern Anal Mach Intell.* 2008;30(2):253-266.
38. Solis-Tripala IL, Farewell VT. Regression analysis of overdispersed correlated count data with subject specific covariates. *Statist Med.* 2005;24(16):2557-2575.
39. Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. *SIAM J Sci Comput.* 1995;16(5):1190-1208.
40. Williams BM, Hoel LA. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. *J Transp Eng.* 2003;129(6):664-672.
41. Taylor JW. Triple seasonal methods for short-term electricity demand forecasting. *Eur J Oper Res.* 2010;204(1):139-152.
42. Qamar A, Asim M, Maamar Z, Saeed S, Baker T. A Quality-of-Things model for assessing the Internet-of-Things' nonfunctional properties. *Trans Emerg Telecommun Technol.* 2019.
43. Hammoudeh M, Al-Fayez F, Lloyd H, et al. A wireless sensor network border monitoring system: deployment issues and routing protocols. *IEEE Sens J.* 2017;17(8):2572-2582. <https://doi.org/10.1109/JSEN.2017.2672501>
44. Maamar Z, Baker T, Sellami M, Asim M, Ugljanin E, Faci N. Cloud vs edge: who serves the Internet-of-Things better? *Internet Technol Lett.* 2018;1(5):e66. <https://doi.org/10.1002/itl2.66>
45. Baker T, Aldawsari B, Asim M, Tawfik H, Maamar Z, Buyya R. Cloud-SEnergy: a bin-packing based multi-cloud service broker for energy efficient composition and execution of data-intensive applications. *Sustain Comput Inform Syst.* 2018;19:242-252. <https://doi.org/10.1016/j.suscom.2018.05.011>
46. Baker T, Ngoko Y, Tolosana-Calasanz R, Rana OF, Randles M. Energy efficient cloud computing environment via autonomic meta-director framework. In: Proceedings of the 2013 6th International Conference on Developments in eSystems Engineering; 2013; Abu Dhabi, United Arab Emirates.
47. Baker T, Asim M, Tawfik H, Aldawsari B, Buyya R. An energy-aware service composition algorithm for multiple cloud-based IoT applications. *J Netw Comput Appl.* 2017;89:96-108. <https://doi.org/10.1016/j.jnca.2017.03.008>
48. Ghafir I, Prenosil V, Svoboda J, Hammoudeh M. A survey on network security monitoring systems. In: Proceedings of the 2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW); 2016; Vienna, Austria.

49. Khalaf M, Hussain AJ, Al-Jumeily D, et al. A data science methodology based on machine learning algorithms for flood severity prediction. In: Proceedings of the 2018 IEEE Congress on Evolutionary Computation (CEC); 2018; Rio de Janeiro, Brazil.
50. Farhan L, Kharel R, Kaiwartya O, Hammoudeh M, Adebisi B. Towards green computing for internet of things: energy oriented path and message scheduling approach. *Sustain Cities Soc.* 2018;38:195-204. <https://doi.org/10.1016/j.scs.2017.12.018>

How to cite this article: Daraghmi Y-A, Daraghmi EY, Alsaadi S, Eleyan D. Accurate and time-efficient negative binomial linear model for electric load forecasting in IoE. *Trans Emerging Tel Tech.* 2019;e3732. <https://doi.org/10.1002/ett.3732>