International Journal of Software Innovation
Volume 11 • Issue 1

Machine Learning-Based Electricity Load Forecast for the Agriculture Sector

Megha Sharma, Malaviya National Institute of Technology Jaipur, India*
Namita Mittal, Malaviya National Institute of Technology Jaipur, India
Anukram Mishra, Genus Power Infrastructures Limited, India
Arun Gupta, Genus Power Infrastructures Limited, India

ABSTRACT

A large section of the population has a source of income from the agriculture sector, but their share in the Indian GDP is low. Thus, there is a need to forecast energy to improve and increase productivity. The main sources of energy in agriculture are electricity, coal, and diesel. Among them, electricity plays an important role in land irrigation. Power forecasting is also essential for demand response management. Thus, any process that dissolves future consumption is favorable. This article presents a time series-based technique for forecasting medium-term load in agriculture. The aim is to find the peak periods of power consumption by months and seasons using statistical and machine learning-based techniques. The result shows that SARIMA has lower RMSE and exponential smoothing has lower RMSPE error than random forest and LSTM, which makes the statistical approach more efficient than intelligent approach for historical datasets. The season-wise peak demand occurs during the Rabi season. Finally, five-year ahead load in the agriculture sector was determined using the best models.

KEYWORDS

Agriculture Sector, Load Forecasting, Machine Learning, Statistical Approach, Time Series Forecasting

1. INTRODUCTION

Power forecasting is required for consumers, utilities, distributors and generators as it plays a vital role in electricity procurement and planning (Kaytez et al., 2015). Consumers of different sectors demand energy in various purpose and need. Thus, load forecasting is a challenging task to regulate the demand and supply of electricity in different sectors. Energy demand forecasting (Alvarez et al., 2010) is essential in every sector but in case of agriculture sector it becomes challenging to forecast due to various factors such as climate, rainfall, and land type. Existing work, forecasting energy demand in the agriculture sector relies on factors from national averages. However, energy demand varies by season, crops, land area, agricultural machinery and technologies, pump sets, and groundwater (Kumar, 2005). In India, after the sixth five-year plan period (1980-1985), agricultural

DOI: 10.4018/IJSI.315735
*Corresponding Author

output growth increased due to improved seed quality, usage of fertilizers, and irrigation. Modern farming requires energy as an input in all stages of agricultural production such as farming, water management, irrigation, agricultural machinery and harvesting. Out of these stages the power consumption for motor pump irrigation is high so it is necessary to forecast the power consumption for this sector (Moulik et al., 1990).

According to the 2011 census, 61.5% of India's population is rural and is dependent on agriculture. The contribution of the agriculture sector to the Indian GDP is 14.4% (as per the 2018-19 Economy Survey). Electricity consumption is increasing every year and the agriculture sector also has high energy consumption every year for the purpose of irrigation, agricultural machinery and cold storage.

The total power consumption in the year 2018-19 was 1158 TWh and growth rate of 2018-19 as compared to 2017-18 was 3.11% and the Compound Annual Growth Rate (CAGR) was 6.58% from 2009-10 to 2018-19. The agriculture sector shared 8% of total consumption in 1980 and 17.49% of the total consumption in 2018-19. In 2018-19, 207 TWh was consumed by the sector, and the growth rate of 2018-19 over 2017-18 was 4.29% and the CAGR rate 2009-10 to 2018-19 was 7.43% (Energy Statistics, 2020). The electricity supplied is subsidized or free of charge and most areas are without meters (Prayas (Energy Group), 2018). Farmers need incentives to grow suitable crops, but they are wrongly blamed for high electricity consumption and face poor-quality electricity due to lack of meters. Accurate forecasting in the agriculture sector benefits the farmers, utilities, and the government. The low level of metering and unmetered connection challenges in the agriculture sector to find the true level of power consumption and to predict electricity, because when utilities deliver electricity, it becomes difficult to predict how much electricity a consumer consumes and how much electricity is lost due to technical and commercial reasons. Farmers having the issue of poor-quality electricity supply and overestimation of the electricity consumption occurs in the agriculture sector that leading to the underestimation of the poor-quality supply. To accurately estimate the power consumption, there is a need to install a meter in the fields. In the presented work, power consumption (metered and flat) is considered for estimating future consumption.

The agricultural cropping season in India is July to June and the major crops are classified as food grain, cash crops, plantation crops, and horticulture crops. Rice, Wheat, Maize, Pulses are the food grains that are required for the country as food-stuff. Cotton, Jute, Sugarcane, Oilseeds are cash crops that are not used for food-stuff but used for commercial purposes. Tea, Coffee, Rubber, and Coconut are plantation crops that are also used for commercial purposes. Fruits and vegetables come under horticulture crops. Kharif, Rabi, and Zaid are the cropping seasons in India. Indian crops are classified into 3 parts according to these seasons, shown in Table 1. In India, extensive cultivation occurs in the Kharif season. Cultivation during the Rabi is limited to land with irrigation of the crops. Crops are grown only once a year in about 70% of the cultivated area (Arthapedia, 2015; Singh, 1976). Zaid is a short duration summer season that does not require rain. The classification of the crop is shown in Table 1.

Season wise forecast of electricity consumption benefits the farmers as they use the resources available during the season accordingly. Various forecasting approaches have been applied to accurately measure electricity consumption. These forecasting approaches are classified into two groups (Cai et al., 2019; Eskandarnia et al., 2018; Hong & Fan, 2016):

Table 1. Crop classification according to monsoon

Types of crops	Kharif	Rabi	Zaid
Duration	June-October	November-March	April-June/July
Sown	Beginning of the first rain (June)	Winter (Mid November)	Between Rabi and Kharif (March/April)
Harvested	End of Monsoon (Sept/Oct)	Spring (March/April)	June/July
Known as	Monsoon crops (Rainwater required)	Winter crops	Summer crops/no need for rain
Example	Bajra, Jowar, Maize, Millet, Rice	Wheat, barely, Gram, Mustard	Cucumber, Muskmelon, Watermelon

a) Statistical approach: Auto-Regressive Integrated Moving Average (ARIMA), Seasonal ARIMA and Exponential Smoothing
b) Intelligent approach: Machine learning and deep learning models.

The motivation for this work is to make predictions of power consumption considering past data. Three techniques are applied to the dataset: statistical, ensemble learning and deep learning, of which the statistical approach shows less error and better performance. The authors use ensemble learning, deep learning techniques and compare them with statistical models.

In this paper, the authors explore SARIMA, Exponential Smoothing, Random Forest and LSTM model in the agricultural sector for time series forecasting to identifying electricity consumption patterns of the consumers. The dataset is collected from Jaipur Vidyut Vitran Nigam Limited (JVVNL)[1]and work is done with the collaboration of Genus Power Infrastructures Limited[2]. The proposed work aims to find the peak power consumption months and the season so that efficient and good quality electricity can be provided to the farmers using demand response management (Torriti et al., 2010).

The main contributions of this manuscript are highlighted below:

- To explore SARIMA, ES, Random Forest and LSTM model in the agriculture sector for load forecasting.
- A comparative analysis of electricity consumption patterns using statistical and intelligent models.
- Identifying electricity consumption pattern in the agriculture sector.
- Determining month-wise and season-wise peak demand for electricity and forecast five year ahead load.

This article is organized in the following manner; Section 2 discusses the previous work done on estimating power consumption with various indicators while section 3 gives a description of the proposed work and specifies the various time series-based models to be used for forecasting power consumption patterns. Section 4 gives information about the results obtained by various forecasting techniques. Section 5 discusses the conclusion.

2. LITERATURE REVIEW

Forecasting electricity demand is an essential tool for maintaining efficient and secure energy infrastructure (Powell et al., 2014). Energy indicators play an important role in identifying key drivers and trends to optimize power consumption. According to “Energy Indicators for Sustainable Development: Guidelines and Methodology”, there are three energy indicators: social, economic and environmental (Vera & Langlois, 2007). Of these indicators, the economic indicator provides an estimate of power consumption for rural areas (Saravanan & Karunanithi, 2018). Electricity consumption and predictability depend on per capita GDP, population and agricultural land (Günay, 2016; Valasai et al., 2012).

Artificial Neural Networks (ANN) are applied for power prediction where the economic indicators are input and Agriculture Sector- Electricity Consumption (AS-EC) is output. Three different learning algorithms (trainscg, traincgf, trainlm) are considered as activation functions in the hidden layer. The result is evaluated based on MAPE error determined by actual and forecasted AS-EC of different neurons used in the hidden layer and finally predicted for the next 12 years with the ANN model using economic indicators.

Similarly, an economic indicator-based work is presented by (Saravanan et al., 2012) where 11 variables per capita GDP, CO_2 emission, population, national income, etc. are analyzed to forecast the electricity demand, as it is important in planning. ANN technique is used for forecasting where 29-year historical data is used for training and the next 19 years electricity consumption is forecasted based on 29-year data and it is used as test data. Forecasting is the long term it is not season-wise forecasting.

The Internet of Thing based model is given by (Muangprathub et al., 2019) to monitor, diagnose, and control crop yield factors and optimize watering demands. The parameters humidity, temperature, and soil moisture are analysis for optimal watering needed in the field. Sensors have also been installed in the fields to estimate the moisture content of the soil and analyze the water requirement of the crop in the future. A genetic algorithm-based model (Ou, 2012) was proposed for forecasting agricultural production. A forecast model for rice yield upstream under climate change is given by (Zhang et al., 2019), where multiple linear regression and boosted tree are used for modeling rice yield.

An ARIMA model is given by (Kaur & Ahuja, 2017) to estimate power consumption in an institution for monthly time series data. The Health Care Institute dataset is used for the task. To estimate future power consumption a model has been selected which has least SSE and MPE errors. Similarly (Panapongpakorn & Banjerdpongchai, 2019) used weather-based models for Short Term Load Forecasting (STLF) using neural networks, RNN, RNN with Average True Range (ATR) as well as ARIMA, SARIMA models. Of the five models, the RNN with ATR gives better results for 3 seasons, May–June, July–Sept, Oct–Dec, and different seasons with each weather forecast model to predict future values.

A time-series analysis model with forecast consumption of electricity for public transportation is presented by (Tepedino et al., 2014), in which a new seasonality coefficient is introduced and named as the Triple Seasonality TSA Model (TSM) compared with the double seasonality model. In double seasonality, only two parameters daily and weekly are added in the model with the coefficient and the triple seasonality model also added the monthly behavior or monthly coefficient. The model can be applied for different seasons using seasonal coefficients.

(Katara et al., 2014) showed that the ARIMA model estimates 7 years ahead of power consumption in many regions. Each area has a distinct ARIMA(p,d,q) model where p is the degree of autoregression, d is the degree of trend, and q is the degree moving average term. This article focuses on residential, industrial and commercial sectors, but does not consider agricultural sector.

The short-term-based load prediction model is given by (Moon et al., 2018), where Multi-Layer Perceptron (MLP) and random forest predict one week load. SVR and fuzzy logic based FSVR model (Sina & Kaur, 2021) provide accurate load forecasting for a day. EUNITE-1997 and New England data sets are used for STLF. Similarly, a Sequential Hybrid Neural Network (SHNN) architecture given by (Eapen & Simon, 2019) for STELF based on similar day and day ahead approaches. It gives better results than BPNN and particle swarm optimization-based approaches.

(Voyant et al., 2011) defined a forecasting model using ANN for global solar irradiation and compared with different forecasting methods like ARIMA reference predictor and naive forecaster. (Wang et al., 2020), proposed a deep learning and clustering-based ensemble learning approach to forecast urban areas' load. In the ensemble learning method, stacking is used in which each model is trained individually and then integrated. Similarly, (Sadaei et al., 2019) proposed a deep learning CNN model for STLF based on fuzzy time series. The data includes hourly temperature, hourly load data and fuzzified load time series.

Most of the work has been done in different sectors but the agriculture sector has remained untouched, thus the authors explore different time series models in the agriculture sector to estimate the power consumption.

3. MATARIAL AND METHODOLOGY

3.1 Dataset

JVVNL (Jaipur Discom) distributes and supplies the power in 12 districts of Rajasthan which are Alwar, Baran, Bundi, Bharatpur, Dausa, Dholpur, Jaipur, Jhalawar, Karauli, Kota, Sawaimadhopur, and Tonk (JVVNL, 2020). Category Wise details of the consumer's bills are available on an actual or an average basis. According to climate, these 12 districts are divided into 3 zones:

a) Zone1: Jaipur, Alwar, Dausa, Tonk comes under the semi-arid eastern plain,
b) Zone2: Bharatpur, Dholpur, Karauli, comes under flood-prone eastern plain,
c) Zone3: Sawaimadhopur, Jhalawar, Bundi, Kota, Baran comes under the humid south-eastern plain (Components of RACP, 2012).

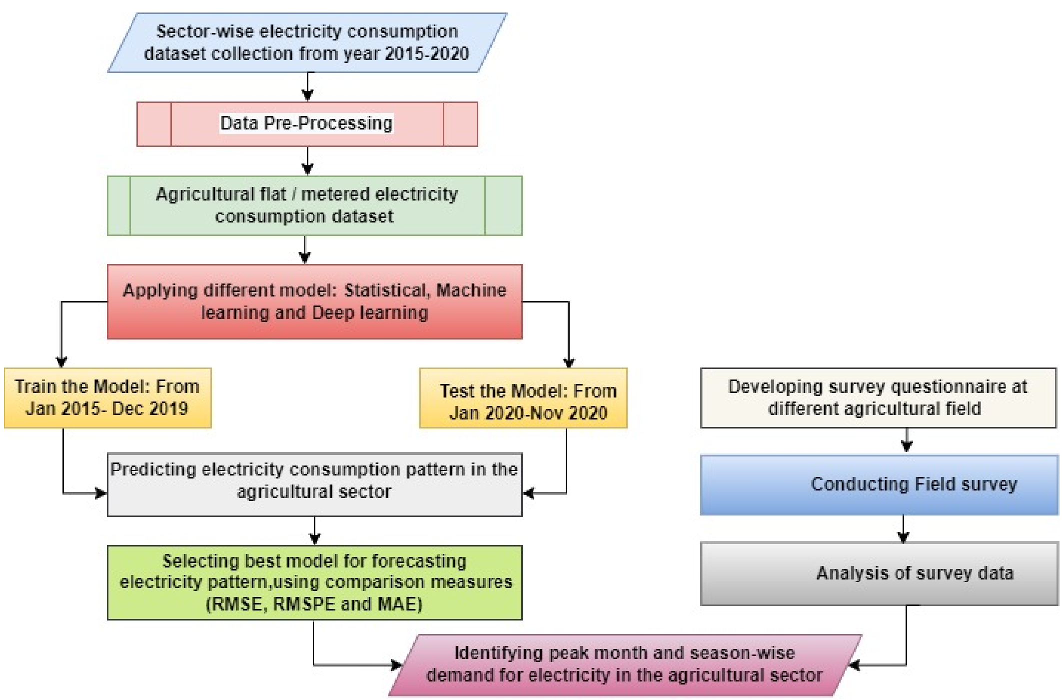
The data set consists of sector-wise power consumption information. In preprocessing, agricultural (metered and flat) monthly electricity consumption data from January,2015 to November,2020 is collected. To train the model past 5 years monthly data is used (from Jan-2015 to Dec-2019). Jan-2020 to Nov-2020 monthly data is used for testing purpose it is also known as actual data and predict electricity consumption for the next year that is known as predicted data. Finally, the authors compare the actual consumption to the predicted consumption to find the best model.

3.2 Methodology

The main objective of time series forecasting of loads in agriculture is to determine seasonally peak demand, so a demand response program can be designed using the forecast. Using historical consumption data, accurate forecasts can be made for future time intervals or multiple time intervals. The proposed work steps are shown in Figure 1.

Firstly, the dataset was collected from JVVNL, which has sector-wise (small industrial power, medium industrial power, commercial, agricultural flat, agricultural metered and residential) consumer details and historical electricity consumption data in Kilo Watt (KW) / Kilo Volt Ampere (KVA) unit. After pre-processing, the dataset contains only agricultural electricity consumption data. The

Figure 1. Proposed Methodology Flowchart



authors explore different statistical, machine learning, and deep learning techniques on the dataset, explained in section 3.3 (Forecasting Model). The authors trained and tested the different (ARIMA, SARIMA, Exponential Smoothing, Random Forest, and LSTM) models to predict future electricity consumption. Then, select the best model for load forecasting and predict future electricity consumption patterns using comparison measures. The Authors visited different agricultural fields for the survey and based on survey analysis, identified the crop-wise electricity need of the farmer. Finally, the authors are determining month-wise and season-wise demand for electricity in the agricultural field.

3.3 Forecasting Model

In time series analysis, the forecasting value of any variable is dependent on its previous pattern and data set collected through time (Mart\'\inez-Álvarez et al., 2015). From Equation (1), X's previous pattern over the period is calculated and according to that next period, data will be forecasted.

$$\mathrm{X}_{\mathrm{t}} = \mathrm{X}_{\mathrm{t}-1} + \mu \tag{1}$$

3.3.1 Statistical model-based forecasting

For analysis, time series data need to be transformed into stationary time series. Stationarity shows no change in the mean, no change in the variance, without periodic fluctuations (Box et al., 2015). Auto-Correlation Function (ACF) is used to find the relation between time series data and it defines a similar pattern at some time lag. Autocorrelation coefficient at time lag k is represented as $\rho\left(k\right) = \frac{\gamma\left(k\right)}{\gamma\left(0\right)}$, where $\gamma\left(k\right)$ is the autocovariance at time lag k and $\gamma\left(0\right)$ is the autocovariance at time lag 0. A correlogram plot is used to evaluate the autocorrelation of the data. ACF computes degree of Moving Average (MA) i.e., MA(q) that reaches to maximum level at q lags. Partial ACF (PACF) computes the degree of Auto-Regressive (AR) i.e. AR(p) that reaches to maximum level at p lags (Kaur & Ahuja, 2017).

Ex: - Auto-Regressive Integrated Moving Average (ARIMA), Exponential Smoothing (ES), and Seasonal ARIMA (SARIMA)

3.3.1.1 Auto-Regressive Integrated Moving Average

It is a combination of Autoregressive and Moving Average with integrated differencing (Conejo et al., 2005; Contreras et al., 2003; Kaur & Ahuja, 2017; Panapongpakorn & Banjerdpongchai, 2019). ARMA(p,q) is represented as:

$$X_{t=}\varnothing_1 X_{t-1} + \varnothing_2 X_{t-2} + \ldots .. \varnothing_p X_{t-p} + Z_t + \beta_1 Z_{t-1} + \ldots \beta_q Z_{t-q} \tag{2}$$

where $X_{t=}\varnothing_1 X_{t-1} + \varnothing_2 X_{t-2} + \ldots .. \varnothing_p X_{t-p}$ is the autoregressive series and p defines the degree of AR term, and $Z_t + \beta_1 Z_{t-1} + \ldots \beta_q Z_{t-q}$ is the Moving average term which is of degree q.

The value of (p and q) is decided by ACF and PACF functions.

Equation (2) can be written as:

$$\varnothing\left(B\right) X_t = \beta\left(B\right) Z_t \tag{3}$$

where AR polynomial represented as:

$$\varnothing\left(B\right) = 1 - (\varnothing_1 B + \varnothing_2 B^2 + \ldots \ldots + \varnothing_p B^p) \tag{4}$$

and MA polynomial represented as:

$$\beta\left(B\right) = \beta_0 + \beta_0 B^1 + \ldots \ldots + \beta_q B^q \tag{5}$$

If roots of the polynomial lie outside of the circle then ARMA(p,q) will be stationary and invertible. The non-stationary data can be molded by removing the trend using the difference operator.

$$\nabla = 1 - B \tag{6}$$

$$\nabla X_t = X_t - X_{t-1} = \left(1 - B\right) X_t \tag{7}$$

ARIMA(p,d,q) process: A process X_t is Autoregressive integrated Moving Average process of order (p,d,q) is represented as:

$$\varnothing\left(B\right)\nabla^d X_t = \beta\left(B\right) Z_t \tag{8}$$

From Eq. 6 and Eq. 8

$$\varnothing\left(B\right)\left(1 - \mathrm{B}\right)^d X_t = \beta\left(B\right) Z_t \tag{9}$$

3.3.1.2 Seasonal Autoregressive Integrated Moving Average Process (SARIMA)

In seasonal ARIMA data contain seasonal periodic components. It also correlates with recent lags. It repeats at some observation. SARIMA(p,d,q,P,D,Q)s has two parts Non-seasonal part(p,d,q) and seasonal part(P,D,Q)s (Tepedino et al., 2014) and represented as (Cai et al., 2019):

$$\varphi_P(B^s)\ \varnothing_p\left(B\right)\left(1 - B^s\right)^D\left(1 - B\right)^d X_t = \vartheta_Q\left(B^s\right)\theta_q\left(B\right) Z_t \tag{10}$$

where seasonal AR term is represented as:

$$\varphi_P(B^s) = 1 - \varphi_1 B^s - \varphi_2 B^{2s} - \ldots \ldots - \varphi_P B^{Ps}, \tag{11}$$

non-seasonal AR term represented as:

$$\varnothing_p\left(B\right) = 1 - \varnothing_1 B - \varnothing_2 B^2 - \ldots \ldots - \varnothing_p B^p, \tag{12}$$

seasonal MA term defined as:

$$\vartheta_Q(B^s) = 1 + \vartheta_1 B^s + \vartheta_2 B^{2s} + \ldots \ldots + \vartheta_Q B^{Qs}, \tag{13}$$

non-seasonal MA term define as:

$$\vartheta_Q(B^s) = 1 + \vartheta_1 B^s + \vartheta_2 B^{2s} + \ldots \ldots + \vartheta_Q B^{Qs} \,, \tag{14}$$

d is the non-seasonal differencing order, and D is the seasonal differencing order.

SARIMA Modeling: This model transform input data into stationary time-series data and afterword ACF and PACF is used to evaluate the degree of AR, MA, Seasonal Auto-Regression (SAR), and Seasonal Moving Average (SMA) (Gonzalez-Romera et al., 2006). SARIMA modeling steps are given below:

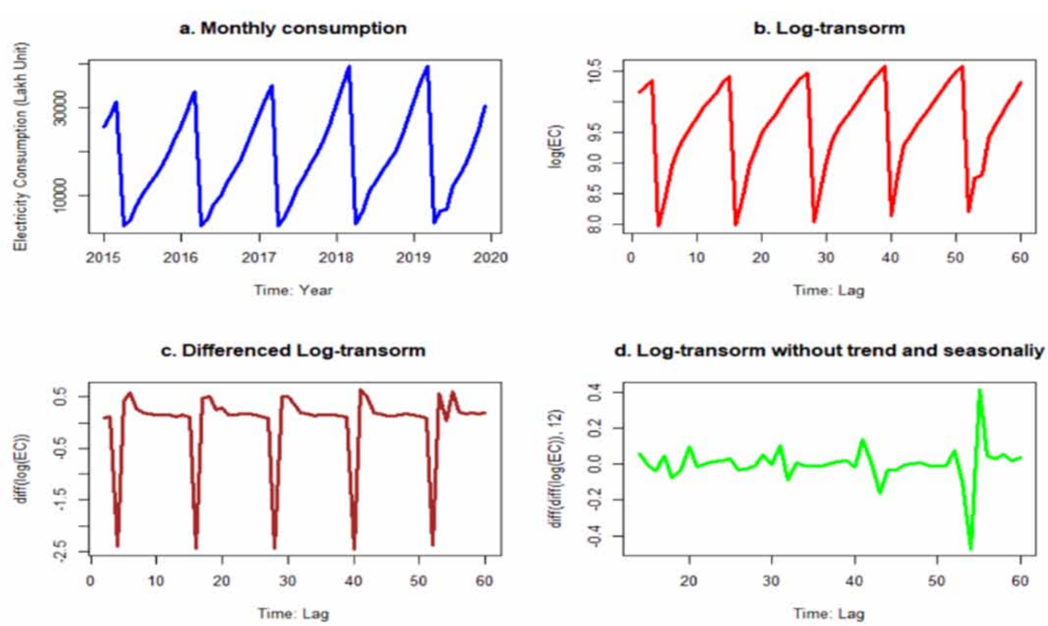
Step1 Time series analysis and removal of trend and variance:

In this step, the time series data is converted into stationary time series data using log and difference transformations to remove seasonality and trend to make them stable. Figure 2 represents the change of month-wise electricity load data into stationary load data, where the monthly consumption pattern of electricity is shown in Figure 2(a), Log transformation of the data given in Figure 2(b), To remove trend the difference of log transformation is shown in Figure 2(c) and the stationary time series pattern without trend and seasonality is shown in Figure 2(d).

Step2 ACF for MA(q), SMA(Q) and PACF for AR(p), SAR(P):

In this step, ACF is used to identify the order of MA and Seasonal MA term at different time lags. Similarly, PACF is used to calculate the order of AR and Seasonal AR term. The degree of MA

Figure 2. a) Month-wise electricity load data b) Log transformation c) Differenced Log-transform data d) Stationary (without trend and seasonality) load data



(q) is 1 and there is a seasonal component after every 12 lag hence SMA(Q) has degree of 5 (shown in Figure 3(a)) while figure 3(b) shows AR (p) degree i.e., either 0 or 1 and SAR (P) degree i.e., 1.

Step3 Model selection with the least AIC value:

Here, identifies the best model among (0,0,0,0,0,0), (1,0,0,0,0,0), ……. till (1,1,1,1,1,5) based on Akaike Information Criteria (AIC) and SSE value. ARIMA (1,1,1,0,1,2) model results minimum AIC and less SSE value. Thus ARIMA (1,1,1,0,1,2) model is selected to fit and forecast future consumption.

Step4 Ljung-Box Q-statistics test:

There is no correlation among the residuals defined using Ljung-Box test for the residuals as shown in Figure 4.

Step5 Estimate and fit the model to time series:

After performing the Ljung-Box test, the best model is selected for predicting the load consumption based on minimum AIC and SSE value. Figure 5 represents the forecasted power consumption pattern.

3.3.1.3 Exponential Smoothing (ES)

The simple exponential smoothing is used when trend and seasonality are not clear while plotting a time-series graph. In ES future values is dependent on the past weighted average value. It includes three component level, trend, and seasonal components with parameter alpha(α), beta(β), gamma(γ) respectively. The seasonality may be additive or multiplicative (Christiaanse, 1971; Gardner Jr, 2006; Taylor, 2003).new level ln represented as:

$$l_n = \propto .(X_n - s_{n-m}) + \left(1 - \propto\right)(l_{n-1} + t_{n-1}), \tag{15}$$

Figure 3. a) Autocorrelation coefficient at different time lag b) Partial autocorrelation coefficient at different time lag

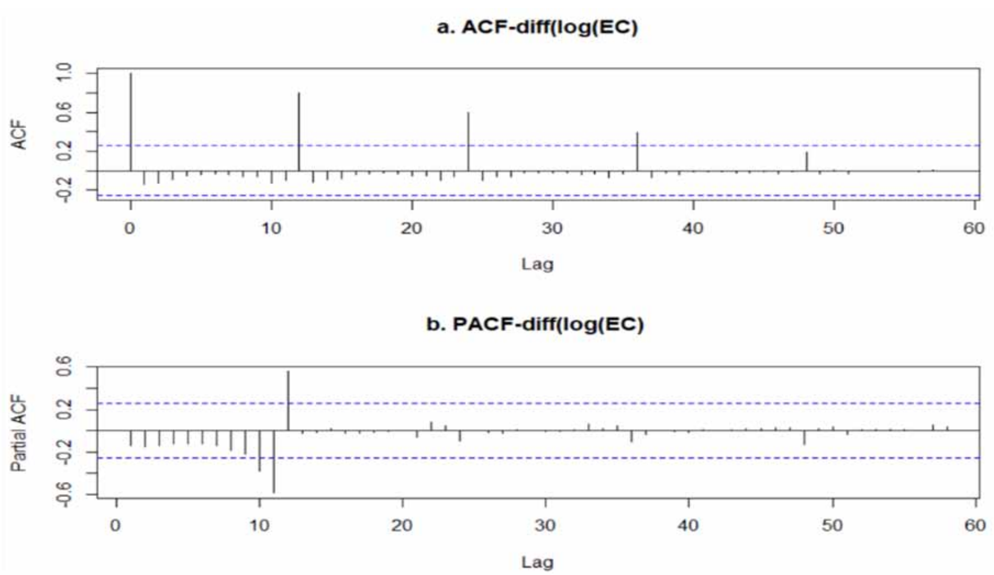


Figure 4. Ljung-Box test for Residuals

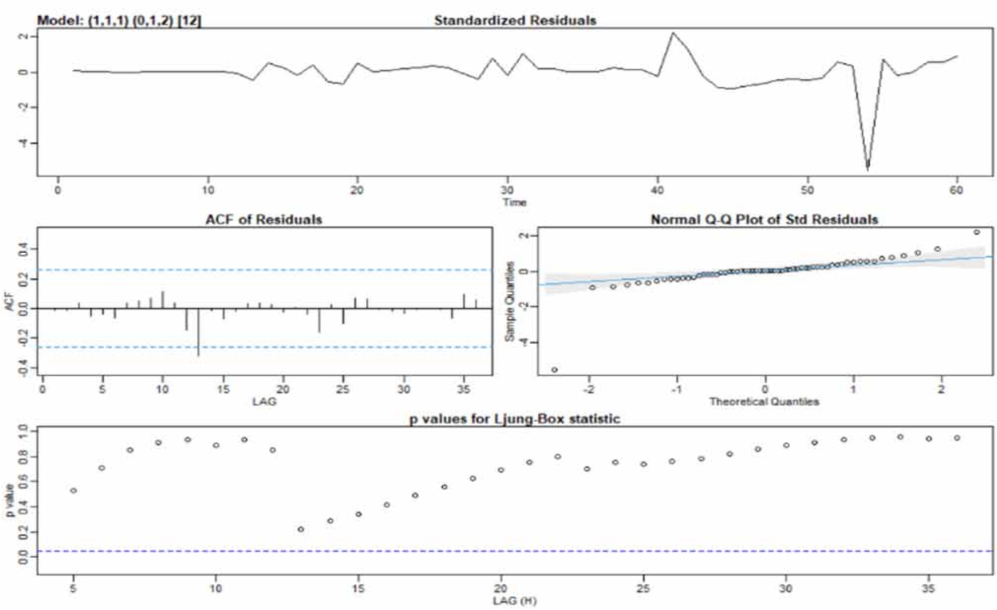
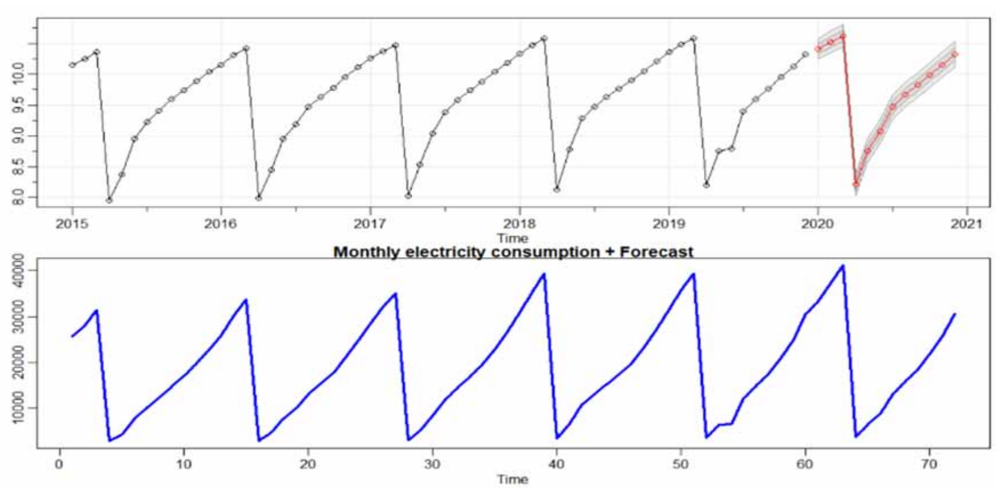


Figure 5. Forecasted Model using SARIMA



new trend tn represented as:

$$t_n = \beta\left(l_n - l_{n-1}\right) + \left(1 - \beta\right)t_{n-1}, \qquad (16)$$

new seasonal sn represented as:

$$s_n = \gamma.\left(X_n - l_n\right) + \left(1 - \gamma\right)s_{n-m} \qquad (17)$$

where h= number of steps into the future for forecastm=length of the season, in our case it is 12.

a) For additive seasonality forecasted value is

$$X_{n+h} = l_n + h.t_n + s_{n+h-m} \tag{18}$$

b) For multiplicative seasonality forecasted value is

$$X_{n+h} = \left(l_n + h.t_n\right).s_{n+h-m} \tag{19}$$

Holt-winters method: Holt-winters function preforms exponential smoothing which is based on level (α), trend (β), and seasonality (γ). A normal distribution curve of electricity consumption shown in Figure 6, thus forecasting for future consumption can be performed.

3.3.2 Ensemble learning-based forecasting

In time series modeling time lags are considered as features and machine learning techniques identify these important features. Random Forest (RF) is an example of an ensemble learning paradigm. RF comprises a large set of Decision Trees (DTs) so it is the ensemble of DTs and these trees predicts variable's value (Galicia et al., 2019). Random forest trains many trees simultaneously. The bootstrap sample creates DTs of the forest from the training set (Cheng et al., 2012). There are m possible features and a random feature that is the best feature among them features is used to split the node of the DT (Karvelis et al., 2017).

Time intervals are essential features in time-series data. RF estimates the current observation based on generation of new time series along with 12-month interval values. Later on, Random Forest identifies the important time lags from 12-month interval to predict current values. Time lag t-2 is the most significant of the two time periods, followed by the value of t at the current time. Figure 7 shows that t, t-2, t-9, t-10 are important time interval that predict the current observation of response variable.

3.3.3 Deep learning-based forecasting

Recurrent neural network and LSTM are mostly used for time series data. Long-Short-Term Memory (LSTM) is one of the most popular deep learning methods as it has self-connected memory cells. It

Figure 6. Histogram and Normality test for data

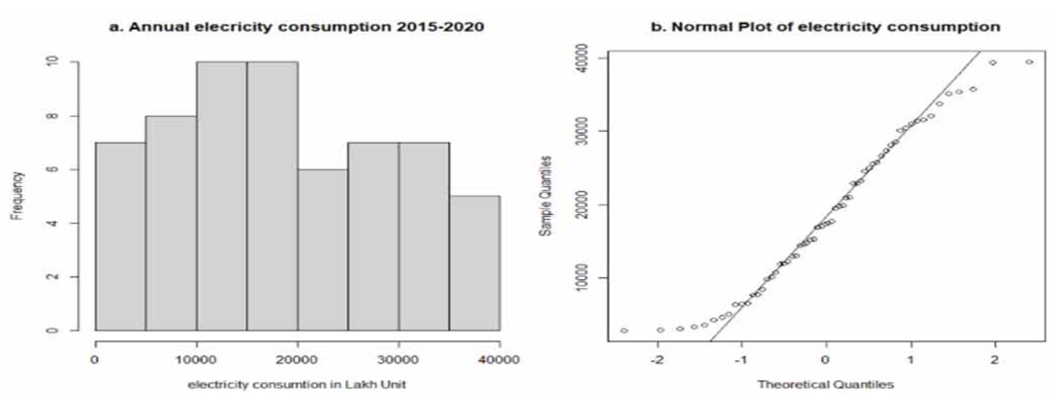
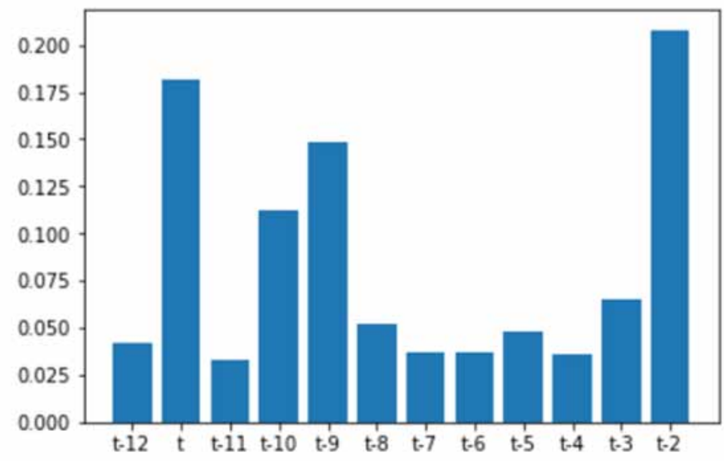


Figure 7. 12 months lag values

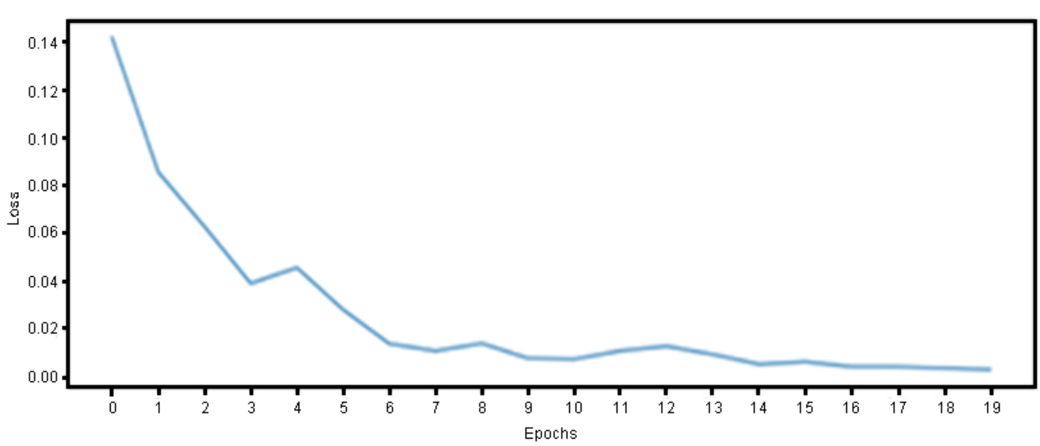


consists of four gates input gate, update gate, forget gate, and output gate (Somu et al., 2020). Time series dataset has particular characteristics and these characteristics are learnt themselves using LSTM. In this work, LSTM architecture contains one input layer, one LSTM layer with activation function "reLU" and one Dense layer. Different losses in 20 epochs are shown in Figure 8.

4. EXPERIMENTAL RESULT AND DISCUSSION

In this section, the authors present the main results obtained by different models to identify peak demand months and season. The work has been done on cumulative data. In this article, the Root Mean Square Error (RMSE), Root Mean Square Percentage Error (RMSPE) and Mean Absolute Error (MAE) has been used for measuring the performance of load forecasting.

Figure 8. Loss versus Epochs in LSTM



4.1 Result using SARIMA Model

Electricity load is affected by the months and seasons. The one-year-ahead predicted power consumption pattern using the SARIMA model is depicted in Figure 9, where predicted power consumption represents almost similar pattern to the actual power consumption. Electricity consumption is measured in lakh units (LU). Table 2 represents the coefficient values along with standard error of SARIMA model.

4.2 Result using Exponential Smoothing Model

The one year ahead estimated power consumption pattern using the ES model is shown in Figure 10, where predicted power consumption has the similar pattern as the actual power consumption.

Table 2. SARIMA ((1,1,1) (0,1,2)) model's Standard Error w.r.t. coefficient values

Coefficient	Value	Standard Error
ar(1)	0.63	0.12
ma(1)	-0.99	0.05
sma(1)	-0.30	0.21
sma(2)	0.99	0.86

Figure 9. Forecasted power consumption pattern using Seasonal ARIMA model a) Month-wise b) Year-wise

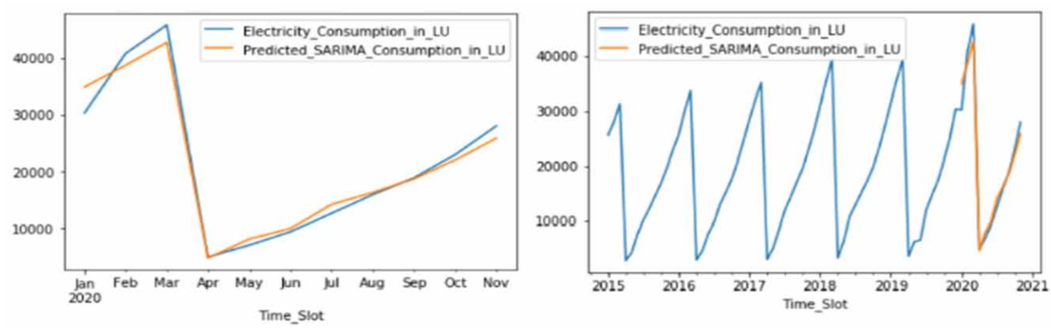
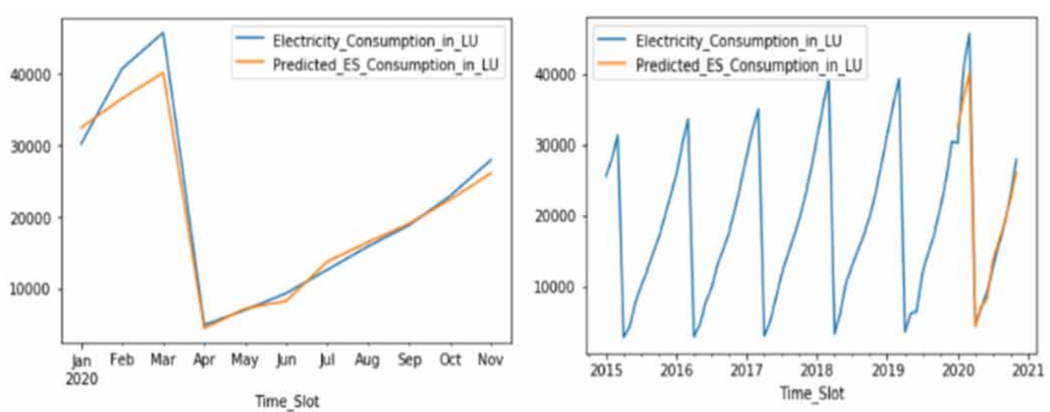


Figure 10. Forecasted power consumption using Exponential Smoothing a) Month-wise b) Year-wise



Apart from seasonal parameters (see Table 3) value the model also computes smoothing parameters whose values are: α=0.18, β =0 and γ=1. From the smoothing parameters, it is observed that the series is not trend dependent but seasonal dependent as value of γ=1 which is most influence compared to other parameters. The smoothing parameter shows that trend has less effect and seasonality have more effect.

4.3 Result using Random Forest Model

The forecasted electricity consumption of year 2020 and compare actual consumption with forecasted consumption shown in Figure 11, where blue line used for actual power consumption and orange line used for the predicted power consumption pattern using Random Forest model.

4.4 Result using LSTM Model

The result of forecasted electricity consumption of year 2020 and compare actual consumption with forecasted consumption using LSTM is given in Figure 12, where the predicted power consumption pattern deviates from actual consumption.

4.5 Comparison among the SARIMA, Exponential Smoothing, RF and LSTM method

The month-wise forecast of power consumption for agriculture is given in Table 4. The proposed 4 predictive models (SARIMA, ES, RF and LSTM) validate from Jan-2020 to Nov-2020 where SARIMA and ES outperform than other forecasting techniques (see Table 5). RMSE, RMSPE and MAE errors are least for statistical models (SARIMA and Holt-Winters). The result shows peak consumption happens in March.

Table 3. Seasonal parameter's coefficients using Holt-Winters

Parameters					
a	20565.09	b	51.83	S1	11889.77
S2	15912.37	S3	19475.89	S4	-16390.9
S5	-13676.2	S6	-12675.1	S7	-7218.14
S8	-4514.08	S9	1980.70	S10	1315.53
S11	4990.68	S12	9854.38		

Figure 11. Forecasted electricity consumption using Random Forecast a) Month-wise b) Year-wise

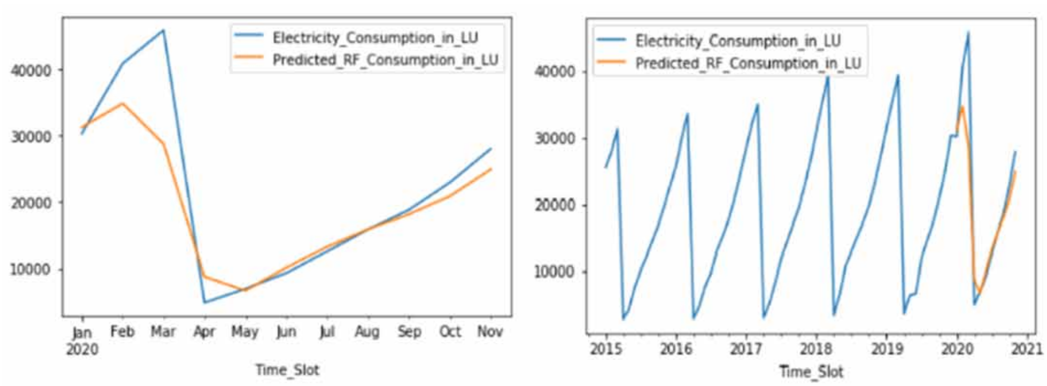


Figure 12. Forecasted electricity consumption using LSTM a) Month-wise b) Year-wise

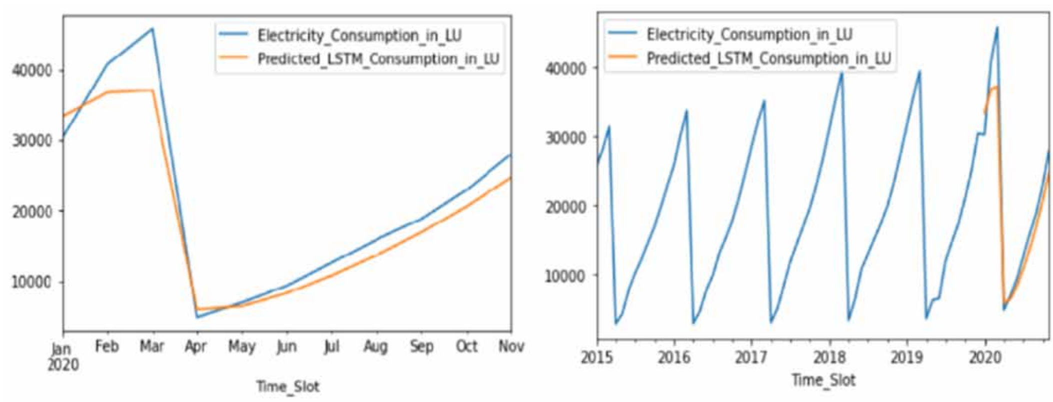


Table 4. Actual v/s forecasted power consumption from January-20 to November-20 with various forecasting techniques

Month	Actual	SARIMA	ES	RF	LSTM
Jan-20	30245.96	34850.36	32506.71	31198.37	33366.98
Feb-20	40777.96	38691.08	36581.14	**34794.27**	36758.74
Mar-20	**45793.18**	**42704.76**	**40196.50**	28699.55	**37125.02**
Apr-20	04860.68	04665.73	04381.52	08737.55	05956.65
May-20	06946.40	08013.04	07148.04	06679.84	06439.53
Jun-20	09295.96	09864.81	08200.92	10112.75	08245.23
Jul-20	12580.66	14148.79	13709.81	13285.45	10763.47
Aug-20	15871.37	16246.67	16465.71	15891.75	13646.93
Sept-20	18820.06	18633.01	19050.93	18157.72	16896.3
Oct-20	22926.32	21966.34	22399.01	20873.69	20569.78
Nov-20	27987.78	25829.45	26126.00	24898.55	24698.24

The root mean square error, root mean square percentage error and mean absolute error are evaluated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(Actual_i - Predicted_i\right)^2}{N}} \tag{20}$$

$$RMSPE = \sqrt{\frac{1}{N}\frac{\sum_{i=1}^{N}\left(Actual_i - Predicted_i\right)^2}{\sum_{i=1}^{N} Actual_i^2}} \tag{21}$$

$$MAE = \frac{\sum_{i=1}^{N}\left|Actual_i - Predicted_i\right|}{N} \tag{22}$$

In Rajasthan, the maximum temperatures are noted in May and June, where power consumption is minimal or in-depth as these are the months of Zaid season. Minimum temperature noted in January, December while electricity consumption is at its peak in these months. Rainfall begins in June and it is highest in July and August month that is the Kharif season. The electricity consumption is less due to it is the monsoon crops and irrigation supplement by rainwater, so less amount of electricity is required. In mid-September, monsoon weakens. Rainfall decreases in October and November (Gurjar, 2015) so electricity consumption increases and this is the Rabi season where most of the farming is done using irrigation. The conclusion is that the Zaid season requires less amount of electricity consumption, Kharif crops require a moderate amount of electricity as most of the farming performed using rainwater and Rabi crops require a considerable amount of electricity and peak occur in December- March.

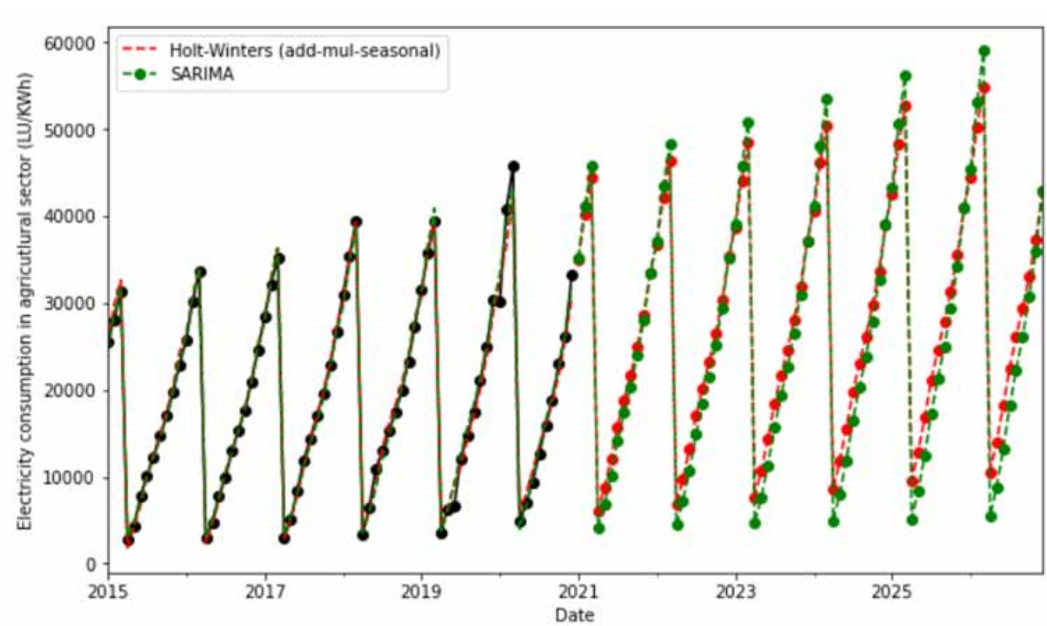
The authors evaluated the statistical and intelligent methodologies and determined five years ahead load in the agriculture industry. Figure 13 depicts the future consumption in the agriculture sector till 2026 using the best models seasonal ARIMA and exponential smoothing. The results show that the performance of SARIMA and exponential smoothing is more satisfactory than that of RF and LSTM.

As per the author's knowledge, the agriculture sector is untouched for forecasting the season-wise load consumption. There exists no work related to the agriculture domain that needs to be compared with the existing work, so the author has excluded it and compared the result with real-time survey analysis.

Table 5. Comparative analysis of selected models using RMSE, RMSPE and MAE

Model	RMSE(LU)	RMSPE (%)	MAE(LU)
SARIMA	2018.28	8.69	1532.63
Holt-Winters	2353.76	7.98	1652.13
Random Forest	5715.72	11.03	3229.02
LSTM	3463.04	13.50	2733.94

Figure 13. Forecasted Electricity Consumption in Agricultural Sector using Seasonal-ARIMA and Holt-winters Method



4.6 Survey at the Bagol Panchayt, nathdwara, Rajsamand

The authors visited the Rajsamand district to collect information and met farmers of the Nathdwara. The main crops of the region are Jwar, Bajra, Maize and Wheat. Farmers reported that there is no rain in April, May and June, and there is no harvest season, so they need a small amount of water and electricity. These months are the Zaid season in which the authors predict non-peak demand for electricity.

Wheat is a Rabi season crop, its time period is November to March, it requires more amount of water and farmers need more electricity. At present they are given electricity for 5 hours in a day but during Rabi season they require 8 hours of electricity in a day.

Maize is a Kharif season crop with duration of July to October which is the monsoon crop and most of the irrigation is being met by rainwater. If monsoon is delayed and rainfall is less then farmer needs to fulfil the requirement from the electricity. Thus, they require average power demand in Kharif season. They require electricity for 3 hours a day during the rainy season. The proposed work results along with the farmers show that Rabi season requires peak demand for electricity, Kharif requires average demand and Zaid season requires off-peak power demand as shown in Table 6. In Rabi season (December- March) the peak month of electricity consumption is March in agricultural sector as highlighted in Table 4, and statistical model gives better results than intelligent approach.

5. CONCLUSION

Agriculture is a prominent sector which is heavily dependent on electricity and hence it is necessary to forecast the power demand based on monsoon or season. This article analysed the seasonal peak demand for power consumption in agriculture sector. Random Forest and LSTM techniques are used in the proposed work to predict power consumption patterns and compare these techniques with statistical methods. The statistical approach performs better among these models and identifies peak demand in March. The result shows the peak-demand occurs in the Rabi season from December-March as rainfall is less as compared to other seasons and the crop irrigation is not supplemented by rainwater. Thus, farmers need peak demand in the Rabi season, if adequate and cost-effective power is provided to them during this season, productivity can be improved. Irrigation is being met by rainwater in Kharif season, thus the average demand for electricity to the farmer is required, but Zaid season has non-peak demand and other sectors require higher demand as it is exposed to warmer temperatures so the consumption can be shifted to the other sectors. In Rabi season, there is a need to fulfil the requirement of peak demand in the agricultural sector.

The historical dataset is not satisfactory for accurate estimation of agricultural electricity consumption, but weather, groundwater levels and types of crops also influence. Thus, the authors can use regression techniques to estimate and plan future load consumption in the agricultural sector.

Table 6. Season/Crop wise electricity demand requirement

Types of crops	Kharif	Rabi	Zaid
Duration	June-October	November-March	April-June/July
Sown	Beginning of the first rain (June)	Winter (Mid November)	Between Rabi and Kharif (March/April)
Harvested	End of Monsoon (Sept/Oct)	Spring (March/April)	June/July
Known as	Monsoon crops (Rainwater required)	Winter crops	Summer crops/no need for rain
Demand of electricity	**Average demand**	**Peak Demand**	**Non-peak demand**

ACKNOWLEDGMENT

The work was supported by Genus Power Infrastructures Limited. The authors would like to thank Federation of Indian Chambers of Commerce & Industry (FICCI) for funding the work with the collaboration of Genus Power Infrastructures Limited under the Prime Minister's Fellowship for Doctoral Research (PMRF). The authors are also grateful to Jaipur Vidyut Vitran Nigam Limited (JVVNL) for providing the dataset.

COMPETING INTERESTS

All authors of this article declare there are no competing interest.

REFERENCES

Alvarez, F. M., Troncoso, A., Riquelme, J. C., & Ruiz, J. S. A. (2010). Energy time series forecasting based on pattern sequence similarity. *IEEE Transactions on Knowledge and Data Engineering*, *23*(8), 1230–1243. doi:10.1109/TKDE.2010.227

Arthapedia. (2015). *Cropping seasons of India- Kharif & Rabi*. http://www.arthapedia.in/index.php%3Ftitle%3DCropping%5C_seasons%5C_of%5C_India-%5C_Kharif%5C_%2526%5C_Rabi

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Cai, M., Pipattanasomporn, M., & Rahman, S. (2019). Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques. *Applied Energy*, *236*, 1078–1088. doi:10.1016/j.apenergy.2018.12.042

Cheng, Y.-Y., Chan, P. P. K., & Qiu, Z.-W. (2012). Random forest based ensemble system for short term load forecasting. *2012 International Conference on Machine Learning and Cybernetics*, *1*, 52–56. doi:10.1109/ICMLC.2012.6358885

Christiaanse, W. R. (1971). Short-term load forecasting using general exponential smoothing. *IEEE Transactions on Power Apparatus and Systems*, *2*(2), 900–911. doi:10.1109/TPAS.1971.293123

Components of RACP. (2012). *Rajasthan Agriculture Competitiveness Project, Environment Assessment and Management Framework, Department of Agriculture*. https://documents1.worldbank.org/curated/fr/571371468267617038/pdf/E29290REVISED00B0SAR0EA0P124614vol1.pdf

Conejo, A. J., Plazas, M. A., Espinola, R., & Molina, A. B. (2005). Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Transactions on Power Systems*, *20*(2), 1035–1042. doi:10.1109/TPWRS.2005.846054

Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, *18*(3), 1014–1020. doi:10.1109/TPWRS.2002.804943

Eapen, R. R., & Simon, S. P. (2019). Performance analysis of combined similar day and day ahead short term electrical load forecasting using sequential hybrid neural networks. *Journal of the Institution of Electronics and Telecommunication Engineers*, *65*(2), 216–226. doi:10.1080/03772063.2017.1417749

Energy Statistics. (2020). *National statistical office, Ministry of statistics and programme implementation*. Government of India.

Eskandarnia, E., Kareem, S. A., & Al-Ammal, H. (2018). A review of smart meter load forecasting techniques: Scale and horizon. *IEEE Conference*.

Galicia, A., Talavera-Llames, R., Troncoso, A., Koprinska, I., & Martínez-Álvarez, F. (2019). Multi-step forecasting for big data time series based on ensemble learning. *Knowledge-Based Systems*, *163*, 830–841. doi:10.1016/j.knosys.2018.10.009

Gardner, E. S. Jr. (2006). Exponential smoothing: The state of the art—Part II. *International Journal of Forecasting*, *22*(4), 637–666. doi:10.1016/j.ijforecast.2006.03.005

Gonzalez-Romera, E., Jaramillo-Moran, M. A., & Carmona-Fernandez, D. (2006). Monthly electric energy demand forecasting based on trend extraction. *IEEE Transactions on Power Systems*, *21*(4), 1946–1953. doi:10.1109/TPWRS.2006.883666

Günay, M. E. (2016). Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. *Energy Policy*, *90*, 92–101. doi:10.1016/j.enpol.2015.12.019

Gurjar, R. D. (2015). *Assessment of impact of urbanization on micro climate in jaipur urban complex based on satellite derived parameters*. University of Rajasthan. https://shodhganga.inflibnet.ac.in/bitstream/10603/148206/7/07_chapter 2.pdf

Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32*(3), 914–938. doi:10.1016/j.ijforecast.2015.11.011

JVVNL. (2020). *Jaipur Vidyut Vitaran Nigam Ltd*. Department of Energy, Government Of Rajasthan. https://energy.rajasthan.gov.in/content/raj/energy-department/jaipur-vidyut-vitran-nigam-ltd-/en/home.html#

Karvelis, P., Kolios, S., Georgoulas, G., & Stylios, C. (2017). Ensemble learning for forecasting main meteorological parameters. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 3711–3714.

Katara, S., Faisal, A., & Engmann, G. M. (2014). A time series analysis of electricity demand in Tamale, Ghana. *International Journal of Statistics and Applications*, *4*(6), 269–275.

Kaur, H., & Ahuja, S. (2017). Time series analysis and prediction of electricity consumption of health care institution using ARIMA model. *Proceedings of Sixth International Conference on Soft Computing for Problem Solving*, 347–358. doi:10.1007/978-981-10-3325-4_35

Kaytez, F., Taplamacioglu, M. C., Cam, E., & Hardalac, F. (2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power & Energy Systems, 67*, 431–438.

Kumar, M. D. (2005). Impact of electricity prices and volumetric water allocation on energy and groundwater demand management: Analysis from Western India. *Energy Policy*, *33*(1), 39–51. doi:10.1016/S0301-4215(03)00196-4

Martinez-Álvarez, F., Troncoso, A., Asencio-Cortés, G., & Riquelme, J. C. (2015). A survey on data mining techniques applied to electricity-related time series forecasting. *Energies*, *8*(11), 13162–13193.

Moon, J., Kim, Y., Son, M., & Hwang, E. (2018). Hybrid short-term load forecasting scheme using random forest and multilayer perceptron. *Energies*, *11*(12), 3283.

Moulik, T. K., Dholakia, B. H., & Shukla, P. R. (1990). Energy demand forecast for agriculture in India. *Economic and Political Weekly*, A165–A176.

Muangprathub, J., Boonnam, N., Kajornkasirat, S., Lekbangpong, N., Wanichsombat, A., & Nillaor, P. (2019). IoT and agriculture data analysis for smart farm. *Computers and Electronics in Agriculture*, *156*, 467–474.

Ou, S.-L. (2012). Forecasting agricultural output with an improved grey forecasting model based on the genetic algorithm. *Computers and Electronics in Agriculture*, *85*, 33–39.

Panapongpakorn, T., & Banjerdpongchai, D. (2019). 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)*, 86–89.

Powell, K. M., Sriprasad, A., Cole, W. J., & Edgar, T. F. (2014). Heating, cooling, and electrical load forecasting for a large-scale district energy system. *Energy*, *74*, 877–885.

Prayas (Energy Group). (2018). *Understanding the Electricity, Water, Agriculture Linkages - Volume 1*. http://www.indiaenvironmentportal.org.in/files/file/Understanding the Electricity%2C Water %26 Agriculture Linkages.pdf

Sadaei, H. J., Silva, P. C. de L., Guimarães, F. G., & Lee, M. H. (2019). Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. *Energy*, *175*, 365–377.

Saravanan, S., Kannan, S., & Thangaraj, C. (2012). Forecasting India's electricity demand using artificial neural network. *IEEE-International Conference on Advances in Engineering, Science And Management (ICAESM-2012)*, 79–83.

Saravanan, S., & Karunanithi, K. (2018). Forecasting of electric energy consumption in Agriculture sector of India using ANN Technique. *International Journal of Pure and Applied Mathematics*, *119*(10), 261–271.

Sina, A., & Kaur, D. (2021). An Accurate Hybrid Approach for Electric Short-Term Load Forecasting. *Journal of the Institution of Electronics and Telecommunication Engineers*, 1–16.

Singh, R. P. (1976). *Improved dryland agriculture for Western Rajasthan*. Central Arid Zone Research Institute. https://krishi.icar.gov.in/jspui/bitstream/123456789/18719/1/Improved dryland agriculture for western Rajasthan.pdf

Somu, N., MR, G. R., & Ramamritham, K. (2020). A hybrid model for building energy consumption forecasting using long short term memory networks. *Applied Energy*, *261*, 114131.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *The Journal of the Operational Research Society*, *54*(8), 799–805.

Tepedino, C., Guarnaccia, C., Iliev, S., Popova, S., & Quartieri, J. (2014). Time Series Analysis and Forecast of the Electricity Consumption of Local Transportation. *Recent Advances in Energy, Environment and Financial Planning", Proc. of the 5th Int. Conf. on Development, Energy, Environment, Economics (DEEE'14)*, 13–22.

Torriti, J., Hassan, M. G., & Leach, M. (2010). Demand response experience in Europe: Policies, programmes and implementation. *Energy*, *35*(4), 1575–1583.

Valasai, G., Uqaili, M. A., Harijan, K., & Associates. (2012). Forecasting Electricity Demand for Agricultural and Services Sector of Pakistan. In Energy, Environment and Sustainable Development (pp. 59–67). Springer.

Vera, I., & Langlois, L. (2007). Energy indicators for sustainable development. *Energy*, *32*(6), 875–882.

Voyant, C., Muselli, M., Paoli, C., & Nivet, M.-L. (2011). Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation. *Energy*, *36*(1), 348–359.

Wang, L., Mao, S., Wilamowski, B. M., & Nelms, R. M. (2020). Ensemble learning for load forecasting. *IEEE Transactions on Green Communications and Networking*, *4*(2), 616–628.

Zhang, L., Traore, S., Ge, J., Li, Y., Wang, S., Zhu, G., Cui, Y., & Fipps, G. (2019). Using boosted tree regression and artificial neural networks to forecast upland rice yield under climate change in Sahel. *Computers and Electronics in Agriculture*, *166*, 105031.

ENDNOTES

[1] Dataset collected from Jaipur Vidyut Vitran Nigam Ltd. and data accessed on the request basis.
[2] Testing and validation of the work are done with the corporation of the Genus Power Infrastructures Limited, Sitapura, Jaipur by visiting different farms in Rajasthan.

Megha Sharma received the BTech and MTech degrees in Computer Engineering from RTU, Kota, India in 2011 and 2016, respectively. She is pursuing PhD in Computer Science and Engineering at Malaviya National Institute of Technology Jaipur, India. Her current research areas are Load Forecasting, Demand Response and Artificial Intelligence.

Namita Mittal received the PhD degree in Computer Science and Engineering from Malaviya National Institute of Technology Jaipur, India. She is an Associate Professor in the Department of Computer Science and Engineering, Malaviya National Institute of Technology Jaipur, India. Her research interests are Data Science, Information Retrieval, Data Mining and NLP.

Anukram Mishra received his BE degree in Electronics and Communication from MBM, Jodhpur, India, MTech degree in Aircraft Production Engineering from IIT Madras, India and PhD in Computer Science and Engineering from Malaviya National Institute of Technology Jaipur, India. He is Chief Technology Officer (CTO, metering solution) in Genus Power Infrastructures Limited, Jaipur, India. His research areas are Smart Grid, Advance Metering Infrastructures and Demand Response.

Arun Gupta received his BE degree in Computer Science & Engineering. He is working as Head – Software Platforms with Genus Power Infrastructures Limited for over two years. He had been working in the Software Industries for over 20 years with 12 years dedicated to Advanced Metering Solutions, delivered across all continents. He enjoys designing scalable secure solutions and has patented his architecture for Integrated Head End Utility Metering System.