# Novel Hybrid ARIMA-BiLSTM Model for Forecasting of Rice Blast Disease Outbreaks for Sustainable Rice Production

**M Varsha** ( mvarsha16@gmail.com )
Bapuji Institute of Engineering and Technology
**B Poornima**
Bapuji Institute of Engineering and Technology
**MP Pavan Kumar**
Jawaharlal Nehru National College of Engineering
**S Basavarajappa**
G.M. Institute of Technology

**Research Article**

**Keywords:** Statistical ARIMA, Autocorrelation, partial autocorrelation, BiLSTM, Novel hybrid ARIMA-BiLSTM

**Posted Date:** September 29th, 2022

**DOI:** https://doi.org/10.21203/rs.3.rs-2103058/v1

# Novel Hybrid ARIMA-BiLSTM Model for Forecasting of Rice Blast Disease Outbreaks for Sustainable Rice Production

**Varsha M[*]., Poornima B[1]., Pavan Kumar M.P[2]., Basavarajappa S[3]**

[*]Bapuji Institute of Engineering and Technology, Davangere India.
[1]Bapuji Institute of Engineering and Technology, Davangere India.
[2]Jawaharlal Nehru National College of Engineering, Shimoga India.
[3]G.M. Institute of Technology, Davangere India.

**Abstract**

In recent years, the application of artificial intelligence (AI) in agriculture has grown to be the most important research domain. The proposed work focuses on forecasting of rice blast disease outbreaks in paddy crop. Disease management in the farm fields is the most difficult problem on the planet. There are variety of reasons for this, first lack of farmers experience in diagnosing diseases, second experts experience in detecting diseases visually, third unfavourable climate. Recent days, researchers have offered variety of time series techniques in different applications. This study adds time series techniques to the field of agriculture by forecasting crucial rice blast disease outbreaks in paddy crop of Davangere region based on daily weather data obtained from KSNDMC. The statistical time series technique called ARIMA is trained by employing real data of blast disease outbreaks in Davangere region from the period of 2015-2019. Meanwhile deep BiLSTM model is trained by employing real weather data and blast disease outbreaks of Davangere region. Both the models are evaluated by performance metrics such as mean squared error and mean absolute error. The proposed research is focused on hybrid model ARIMA-BiLSTM which is a combination of statistical ARIMA model and deep BiLSTM model. Seasonal component of rice blast disease outbreak feature is extracted from additive decompose function used in ARIMA model and fed as dependent feature for BiLSTM model. According to the results obtained, the hybrid approach has the ability to successfully forecast blast disease outbreaks in paddy crop with mean squared error 0.037 and mean absolute error 0.028 compared to statistical ARIMA and deep BiLSTM model.

**Key words:** Statistical ARIMA, Autocorrelation, partial autocorrelation, BiLSTM, Novel hybrid ARIMA-BiLSTM

## 1. Introduction

Knowledge of agricultural sectors is critical for contributing to nation building. Agricultural production is a yet another source of wealth that enhances farmer development. Agriculture development serves as both an essential and a requirement in the global market for a successful country. The world's population is increasing at a rapid rate, creating large agricultural production during the next 50 years. Data about various crop types and pathogens that occur at every level, as well as their early analysis, play an important and wider approach in the agriculture sector [1]. The main issue for a farmer is the occurrence of different diseases on their crop production. Disease categorization and analysis are critical concerns for agriculture's maximum food yield. Food safety is a major concern because of a lack of equipment and technology, so plant disease classification and recognition must be favoured in the near term. These are considered necessary for crop yield, food production, and disease prevention. The detection and classification of pathogens is a significant area of study because it has the potential to monitor large fields of crop production and predicts disease symptoms as they appear on plant leaf. As an outcome, discovering a rapid, cost-effective method for determining plant disease occurrences is essential.

Artificial intelligence (AI) aids agriculture significantly, which boosts a country's overall gross domestic product (GDP) primarily throughout this sector. Climate science, availability of labour, unpredictable rainy seasons, natural catastrophes, and diseases on plant leaves are all significant problems in farming. There is a new breakthrough with various deep learning models that helps to overcome the challenging tasks. The effectiveness of thermal imaging, visible and infrared imaging applications helps in strengthening agriculture productivity and processes by providing farmers and farming management with critical data on the factors determining crop condition and growth [2]. The above technology is extensively used in a broad range of agricultural applications, which includes sustainable agriculture.

The rice (Magnathaporthe Oryza) is one of the world's major food crops, and it has contributed significantly to the resolution of food shortage issues. In comparison to other crops [3], the rice requires less manpower and mechanization. Thus, a trend has emerged in which the area of cultivation has increased. The Pyricularia grisea fungus causes rice blast disease [4]. In the Karnataka state, blast disease has become a serious problem. Controlling blast disease becomes more difficult due to rapid development of pathogen and its region dependency. Blast disease has the ability to rapidly break down even in variety resistance hence, superior cultivars that are resistant to blast disease will become vulnerable after widespread sowing for 2-3 successive sowing season [5]. The external environment factor [6], or climatic factor, are some of the causative agents of blast disease. The function of wind speed, for example, is critical for the spreading of micro-organisms or the light severity that impacts the infection penetrating process. Hence, it is necessary to develop the accurate forecasting model which helps farmers to take critical actions before the severity of the rice blast disease increases.

Prediction systems plays a crucial role in managing and controlling a major disease such as rice blast. Disease forecast models can help farmers as well as other users to decide strategically about the amount and timescale of fungicide applications. In terms of genetics, prediction systems were being based on assumptions about the disease interactions with the host and the climate, which are commonly referred to as the disease triangle. The existence of reliable and dependable early detection systems would allow for the prevention of the pathogen explosive nature through the prompt and effective implementation of preventive actions. This might cause a decrease in reduction in crop loss and fungicide applications, thereby reducing the environmental impact of rice production.

Major contribution of this paper can be summarized as follows:

- To present existing statistical time series techniques and machine learning techniques for the forecasting of rice blast disease outbreaks.
- To discuss key enablers of novel hybrid ARIMA-BiLSTM model for the forecasting of rice blast disease outbreaks and its key efficiency compared to statistical ARIMA model and deep BiLSTM model.
- To present state-of-the-art frameworks for understanding occurrence of rice blast disease. Wherein we discuss about different climate variables that were studied to understand relationship of climate variables with rice blast disease outbreaks across the world.
- To present and evaluate statistical ARIMA model and deep BiLSTM model in performing long term forecasting of rice blast disease outbreaks using climate data of Davangere region.
- To develop and evaluate novel hybrid ARIMA-BiLSTM model for the forecasting of rice blast disease in Davangere region and compare performance of novel hybrid model with statistical ARIMA and deep BiLSTM model using MAE and MSE allowing results to be more accurate and closer to the real-world scenarios.

The remaining section of the research article is structured as follows. Section 2 presents the significant literature works related to the Disease predication and Forecasting of Rice Blast Disease Outbreaks. Section 3 presents the technical details of the proposed methodology. Comprehensive experiments and result analysis are stated in the Section 4 followed by futurity actions, section 5 presents comparative analysis and performance evaluation. Conclusions are stated in section 6.

## 2. Related Work

Due to the severity of rice blast impacts on paddy cultivation, several researchers have devised early detection of rice blast predictive model. Researcher [7] examined 52 rice blast forecasting methods and determined that the parameters used were ambient temperature (T, 67.3%), humidity levels (RH, 57.7%), rainfall (55.8%), leaf temperature (34.6%), sunshine (30.8%), wind velocity (30.8%), and dewpoint (15.4%). The variables that were most frequently combined were minimum and maximum temperature (T) and humidity levels (RH). Furthermore, the disease emergence was found to be positively related to surface temperatures under the phase structure of the rice blast disease. [8] revealed that rice blast occurrence was minimum lowest when minimum temperature at 27.8 C and highest when minimum temperature at 20 C. According to [9], resistance will increase as both atmosphere and soil temperature increases. The above forecasting models may be employed to determine which period are favourable as well as whether fungicide usage is cost-effective or costly within these scenarios. The several nations have created both observational and comprehensible simulation models for rice blast forecasting using regression analysis [7]. However, because climate change has a significant impact on the rice field atmosphere, the traditional forecasting model may lose predictive accuracy [11].

In recent decades, very few researchers have thought the rice blast framework as a dynamic system and have developed rice blast prediction model using machine learning (ML), particularly artificial neural networks

(ANNs), which is considered as intelligent problem-solving methodologies. Researcher collected weekly weather data in India for the development of a cross location and cross years forecasting model. The model included neural network, regression approach and support vector machine[12]. In the feature analysis process rainfall found to be a most influential parameter for the disease occurrence. [13] developed a early rice blast occurrence prediction models for four Korean regions using a long short-term memory (LSTM) . Climatic data such as temperature, humidity levels, and sunlight were collected in June and July from 2003 to 2016. Researcher obtained accuracy of about 79% from LSTM model. [14] compared process-based models such as WARM and YOSHINO with machine learning models such as M5Rules and RNN. The result showed that LSTM model obtained that accuracy of 70%.

In comparison to the conventional REG model, the above ML-based rice blast forecasting frameworks, such as BPNN, GRNN, SVM, and RNN, have evolved an accurate forecasting result based on various climate factors like temperature, relative humidity, and sunlight. However, it has been discovered that rainfall affects blast pathogen development and spread; additionally, blast occurrence is known to be related to heavy rainfall, which have yet to be shown in appropriate applied research. As a result, the impact of rainfall as one of the factors for prediction model development was examined in this study.

Each of the above-mentioned efforts using neural network approach are critical, and they have proven to be quite efficient in forecasting short-term time series. Traditional time series models and deep learning models were both found as powerful tools for modelling any type of application. To the best of our knowledge, no substantial study has been presented for forecasting rice blast disease outbreaks using historical climatic data of the Davangere region. As per our priori research experience and recent literature, an attempt has made to develop statistical ARIMA model to see if it could produce accurate projections in the case of forecasting of blast disease. As described above, models based on deep neural networks outperforms compare to traditional ARIMA model in the case of time-series long-term prediction problems. An experiment is carried out to investigate the performance of deep leaning forecasting systems in the event of blast disease epidemic. An examination is conducted to check the usage of BiLSTM architecture for forecasting long-term blast disease outbreaks. The learning process is carried out using a multivariate input structure. Furthermore, the performance of deep BiLSTM is compared to the classic ARIMA technique. In addition to that, we developed a ARIMA-BiLSTM hybrid model, by accessing seasonal component recovered via additive decomposition to use in the input sequence of deep BiLSTM model in order to forecast the occurrence of rice blast disease in Davangere district. The performance of the proposed novel hybrid model is compared with the classic ARIMA technique and the deep BiLSTM model.

## 3. Proposed Framework

In the proposed study, forecasting models for blast disease outbreaks is designed using statistical time series model, deep BiLSTM model and novel hybrid ARIMA-BiLSTM model. The diagrammatic representation of the proposed model is shown in Figure 3.

### 3.1 Time Series Dataset

Selection of dataset in disease management of precision agriculture is very important because performance of models depends on the accuracy of the dataset. In the proposed study, as the dataset is not available publicly, dataset which is proposed in this work is the real time dataset. To forecast rice blast disease, it is important to understand relationship of climate parameters with the occurrence of rice blast disease. In the literature survey we understood that computerized forecasting model EPIBLA [15] was developed in India to analyze rice cultivars such as IR50 and IR20 for prediction of disease incidence and simulation model suggests that minimum temperature, higher rainfall and maximum relative humidity influence disease incidence. Another experiment [16] was conducted in Kangra District of Himachal Pradesh to find important weather features of rice blast disease severity. During the experiment conduction minimum temperature and high relative humidity influenced disease development. Many experiments were conducted in different countries across the world [17]–[20]to analyze the most important weather parameters that influences rice blast disease occurrence and progression. Results states that minimum and maximum temperature, relative humidity, rainfall are the important factors that influence disease occurrence and along with above stated weather factors wind speed is another important factor that influence in rice blast disease progress. Table 1 and Table 2 describes experiments and prediction models that are available to analyze most important weather factors towards rice blast disease occurrence and progress. The proposed work is focused on forecasting rice blast disease occurrence hence, the climate dataset is collected for Davangere district (Karnataka, India) as the region is considered as the high rice-producing district in the state of Karnataka. The compiled and authenticated data is acquired from the Karnataka State Natural Disaster and Monitoring Centre Bengaluru using robust telemetric weather stations across Davangere district. The dataset contains 9130 instances of day weather data from the year 2015 to 2019. The proposed work is focused on forecasting of rice blast disease outbreaks. Therefore, extensive literature survey is conducted to decide important weather features of rice blast disease occurrence in Davangere district. Survey reveals that minimum temperature, maximum temperature, relative humidity and rainfall are the important features which decides occurrence of rice blast disease in specific region. Hence, the proposed dataset has seven input climate attributes and one numerical outcome variable. The dataset comprises of information such as minimum temperature, maximum temperature, temperature difference, maximum relative humidity, minimum relative humidity, humidity difference, and rainfall.

The study area is influenced by aforementioned climate variables in rice blast disease occurrence. However, dataset used in the proposed study invites a big challenge to the research scholar' for understanding sensitive information in improving accuracy and effectiveness of the statistical, deep learning and hybrid models. The dataset entails determining the probability of occurrence of rice blast disease in the region based on climate variables.

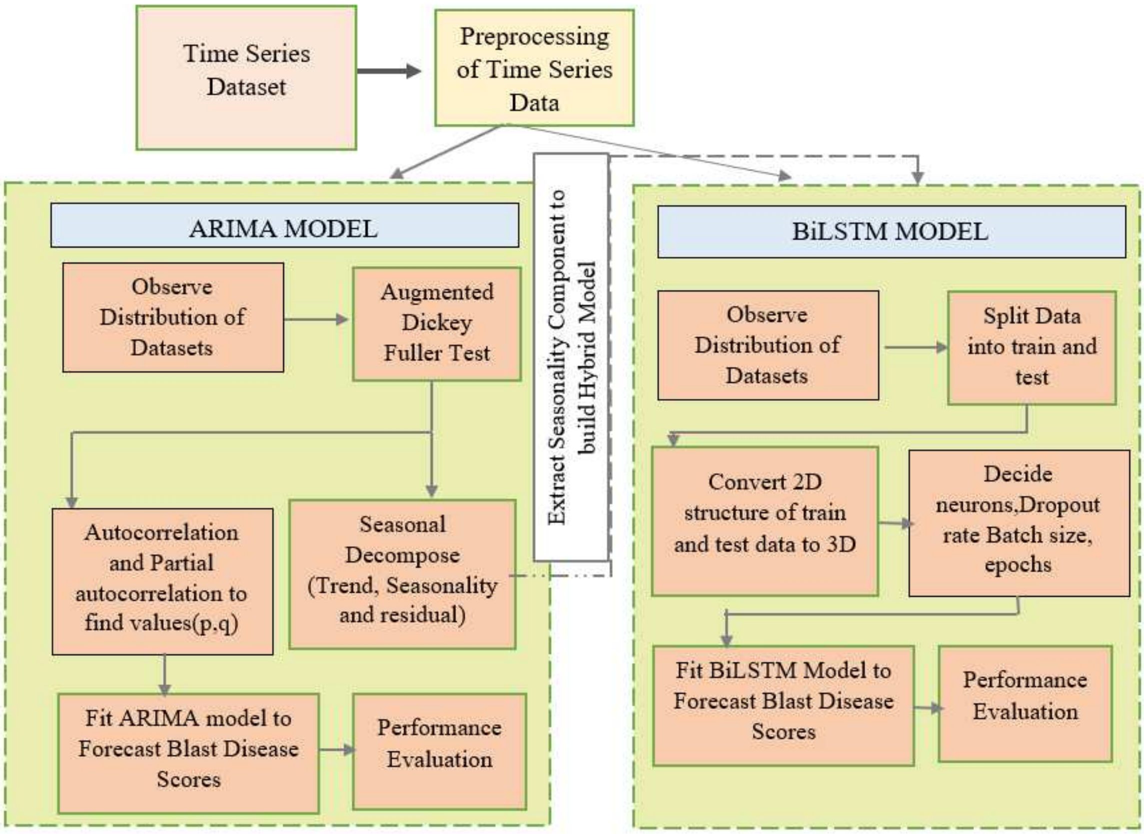


**Figure 1: Novel Hybrid Model for Sustainable Rice Production**

### 3.2 ARIMA Statistical Time Series Model

Statistical approaches for time-series data analysis and to construct the ARIMA are presented in the following subsection.

#### 3.2.1 ADF Test

To test the null-hypothesis for the existence of unit root in the time-series samples the proposed Augmented Dicky Fuller (ADF) test has to be performed [21]. The Dicky Fuller test is done to evaluate whether or not a time-series sampling has random walk. This is achieved by the following Equation (1).

$$\Delta y_t = y_t - y_{t-1} = \propto + \beta t + \gamma y_{t-1} + e_t \tag{1}$$

Stationary time series [26] implies the absence of any trend or seasonal influences in time-series data that makes easier for the predictions. The enhanced Dicky Fuller test is a simplified form of Dicky Fuller examination that allows increased regression process of the form $\Delta y_{t-p}$ where $1 \leq p < t$.

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \ldots \tag{2}$$

According to the null hypothesis defined by ADF test, the p value of dataset is examined to confine that time series data acquired is stationary or not. If p value is less than 0.05 for any time series data the null hypothesis gets rejected and data is said to be stationary.

### 3.2.2 Seasonal Decomposition

The statistical process of decomposing a time-series into the trend, seasonal and residual are known as time series decompose. A trend is general movement of data over time. Seasonal is the behaviour of data observed in individual seasonal periods. Last, residual is the data which is not observed by trend and seasonal components.

There are two techniques for decomposition of time series data namely additive decomposition and multiplicative decomposition.

A time series that follows an additive model is mathematically described as follows:

$$y_t = \tau_t + c_t + s_t + \epsilon_t \tag{3}$$

A multiplicative model on the other hand is mathematically represented as follows:

$$y_t = \tau_t \times c_t \times s_t \times \epsilon_t \tag{4}$$

Where $c_t$ $s_t$, $\epsilon_t$ , $\tau_t$ are the components of cyclical, trend, irregular (noise), and seasonality.

### 3.2.3 Autocorrelation and Partial Autocorrelation

The correlation among two values in a time series is known as autocorrelation. To put it another way, to check the correlation of time series attributes autocorrelation function is used and the word "lags" is used to describe these relationships. Time data series is generated by making an assessment of a feature at regular period such as daily, monthly, or annually. The lag is the frequency of intervals between two moments.

The observation at $y_t$ and $y_{t-k}$ are spaced by k units of time. The lag is denoted by the letter k. According to the nature of the data, the lag might be days, week, or years. When k=1, describes evaluating observations that are close together. The autocorrelation function with lag k is mathematically written as:

$$r_k = \frac{\sum_{j=1}^{M-k}(Y_j-\bar{Y})(Y_{j+k}-\bar{Y})}{\sum_{j=1}^{M}(Y_j-\bar{Y})^2} \tag{5}$$

The partial autocorrelation function is identical to the ACF, it shows the correlation between two data points that is not explained by the shorter delays. For example, the partial autocorrelation of lag 3 is the association that is not explained by lag 1 and lag 2. In other words, the partial correlation is the distinct correlation between two measurements for each lag. The autocorrelation function, is clearly stated in above section, aids in determining the qualities of a time series data whereas the Partial Autocorrelation Function (PACF) is useful in the definition phase of regression analysis. To explain regression techniques for time-series dataset analyst employs partial autocorrelation plots.

## 3.3 ARIMA Model for Blast Disease Outbreaks

The ARMA algorithms stated in the section 3.2 can only be used with static time series data. In practise, most of the time series such as data related to socioeconomics and industry, exhibits non-stationary behaviour. Time - series data which includes trend and seasonal patterns are called as non-stationary. As a result, ARMA theories are insufficient for adequately describing non-stationary time series. Thus, the ARIMA model is proposed, which is an elaboration of the ARMA model that includes non-stationarity time sequence data.

The Auto Regressive Integrated Moving Average (ARIMA) proposed by Box and Jenkins. It is a repetitive and successful method for analysing time data. An ARIMA model assumes data as a linear mixture of previous values, past mistakes, and future values. The ARIMA model is a strong traditional method that incorporates the Auto Regressive and Moving Average models through a differential process which renders the series stationary. A non-stationary time-series can be converted to stationary using ARIMA models by implementing a bounded differencing to the time-series data samples.

Auto regression use the regression equation to anticipate the value for the future time step based on data from prior time steps. To render the sequence stationary, the auto regression employs differencing of raw observations. Differencing is a method for turning non-stationary time- series into stationary. A pure Auto-Regressive model is one in which $X_t$ is solely dependent on its lags which is described in Equation 6.

$$X_t = \alpha + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_p X_{t-p} + \epsilon_1 \quad (6)$$

Where $X_t$ is the first lag in the series, $\beta_1$ is the estimated lag1 coefficient by the model and α is the phrase "intercept"

A pure Moving average model on the other hand is one in which the $X_t$ is determined by the lagged prediction error as described in Equation 7.

$$X_t = \alpha + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \quad (7)$$

Where the erroneous terms represent the errors of the different lags in autoregressive models. $\epsilon_t$ and $\epsilon_{t-1}$ are the errors from the equations:

$$X_t = \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_0 X_0 + \epsilon_t \quad (8)$$
$$X_{t-1} = \beta_1 X_{t-2} + \beta_2 X_{t-3} + \cdots + \beta_0 X_0 + \epsilon_{t-1} \quad (9)$$

When combined the Auto-Regressive and Moving Average model, we get the following Equation 10

$$X_t = \alpha + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_p X_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \quad (10)$$

Equation 11 represents ARIMA model and is a stationary time series model in which it is differenced at least once. The usual notation used is ARIMA (p, d, q) the bracketed parameters are replaced with integer values, indicating the exact values for implementing ARIMA. The following are the definition of the settings:

p is the frequency of lags features in the model;

d is the quantity of differencing procedures required for converting non stationary data to stationary time series data.

q is the size of the window for moving averages.

The autoregressive model and the simple moving model are effectively connected to produce the ARMA type of time-series models, which is a broad and useful. By differencing the data series, the same class of models may be expanded to build ARIMA model for handling non-stationary time sequence. As non-stationary sequence is unpredictable to obtain consistent and dependable findings, non-stationary data must be turned into stationary data.

ARIMA model can be guessed to some extent based on three things: Time series plot, Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The experimental results of these are presented in section 4.

**Steps followed to build ARIMA model**

The following procedures are used to forecast Blast Disease occurrence using the statistical ARIMA model:

1. Observe the structure of the distributed time series using analytical procedures, and pre-process the information in accordance with the model's requirements.
2. Find the differential order d using the Augmented Dickey–Fuller (ADF) test, then evaluate the p and q values, by applying the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF).
3. After ARIMA parameters are determined, experiment is conducted to discover the optimum values for p and q.
4. Design a comprehensive ARIMA model to forecast the occurrence of rice blast disease and assess the performance of the model using performance measures like mean square error and mean absolute error.

**Algorithm of Statistical ARIMA Model**

**ALGORITHM : ALGORITHM OF STATISTICAL ARIMA MODEL TO FORECAST BLAST DISEASE OUTBREAKS**

***Input:*** *Data set of historical blast disease score*

***Output:*** *Performance of ARIMA model in terms of MAE and MSE*

***Sizeofdata***←*length(series) * 0.80*

***training*** ← *series[0...Sizeofdata]*

***testing*** ← *series[Sizeofdata...length(Series)]*

*hist* ← *training*

*prediction* ← *NULL*

```
for value of t1 in range (length of(testing)) do
    modelofARIMA ← ARIMA(hist, orders=(3, 0, 1))
    modelofARIMAfit ← modelofARIMA.fit()
    modelofARIMAhat ← modelofARIMAfit.forecast()
    prediction.appends(modelofARIMAhat)
    observations ← testing[t1]
    hist.appends(observations)
end
MSE_MODEL = mean_square_error(testing, prediction)
MAE_MODEL = mean_absolute_error(testing,prediction)
Return (MSE_MODEL, MAE_MODEL)
```

### 3.4 Bidirectional LSTM for Blast Disease Outbreaks

The notion of memory cells with regulating gates assists LSTMs in overcoming the challenges of long-term reliance and disappearing gradients related with RNNs. Backpropagation is used in the learning process to determine and update the weights. Researcher [22] has proposed LSTM variant in which a unique LSTM model layers are designed which consists of two blocks of LSTM to analyse time input in opposite directions at the same time. Each time instance's output is a concatenation of each LSTM block's outputs. As demonstrated in Figure 2, the architecture of LSTM may be extended to create a bidirectional LSTM (BiLSTM).

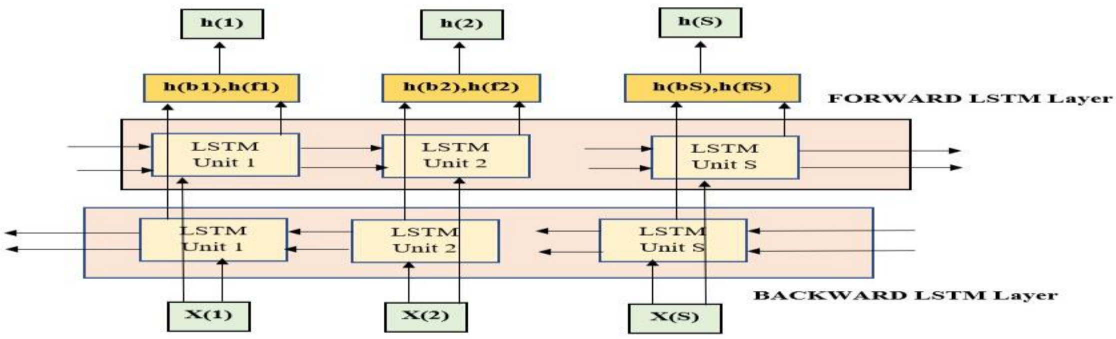


**Figure 2: Bidirectional Structure Applied to LSTM Structure**

The bidirectional LSTM model is made up of two separate LSTM cell, one processing the sequence of input from left to right and the other cell processing the input from right to left. As illustrated in Figure 2, the LSTM design enables the model to learn the input data both from the forward and backward phases. The interpretations of outputs from both the directions are coupled to yield the relative value at the current time interval. In the recent days, the BiLSTM model is widely used to solve the problems of time-series data forecasting. The key principle is that the forecasted points are created by combining the input data with past output values. Furthermore, time series data samples and its reversed copies are used to make predictions that can offer additional context to the model performance, resulting in faster processing and more effective problem learnings. It is a closed-loop model that may be used for sequential prediction or regression analysis. It is highly adaptable since it is predicated on actual and simple real-world data rather than complicated mathematical models. BiLSTM has fully connected layer. Figure 3 depicts fully connected BiLSTM model.

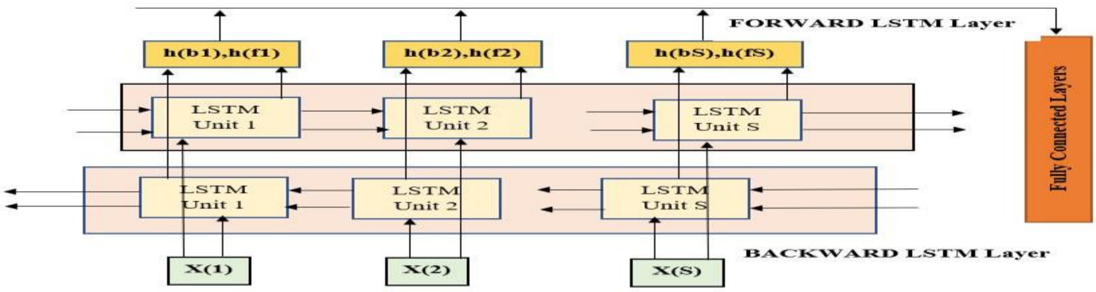


**Figure 3: Fully Connected BiLSTM Model**

**Algorithm of Bidirectional LSTM Model**

**ALGORITHM : FORECASTING OF BLAST DISEASE OUTBREAKS USING BILSTM MODEL**

```
Input: Time Series weather data and blast disease score data
Output: MSE, MAE of the proposed BiLSTM model
sizeofdata := length(series) * 0.80
training := series [0...sizeofdata]
testing := series[sizeofdata...length(Series)]
set random. Seed(7)
procedure fitting_BiLstm(training, epochs, neurons) do
        x1 := training
        y1 := training - x
        model1 := sequentialmodel()
        model1.adding (Bidirectional (LSTMNW (neurons,statefulness=True))
        model1.compilemodel(loss='mean_square_error', optimizers='adam')
        for each of i in ranges(epochs) do
                model1.fitting(X_train, y_test, epochs=30, shuffle=False)
                model1.reset_state()
        End
return model1
procedures forecasting_BiLstm(model1, X) do
        yha := model1.predict(X)
return yha
epochs:=30
neurons:=64,32
prediction := NULL
Bilstm_model := fitting_BiLstm(training,epochs,neurons)
BiLstm_model.prediction(training)
for each of i in the range(length(testing)) do
        X:= testing[i]
        yha := forecasting_BiLstm(BiLstm_model, x)
        prediction.appending(yha)
        expectations := testing[i]
end
MSE := mean_squarederror(expectations, prediction)
MAE:= mean_absoluteerror(expectations ,prediction)
return(MSE,MAE)
```

### 3.5 Proposed Novel Hybrid ARIMA-BiLSTM Model

So far, in the section 3.3 it has been demonstrated two primary components mainly the trend and cyclical components, which are critical to analyse behaviour of time series data. The first denotes continuous movement, whereas the second denotes periodic oscillations. The ARIMA model is used to estimate, isolate, and delete certain components from the series and it uses maximum likelihood to estimate the model coefficients. Seasonality is another important component of time series. ARIMA model does not consider seasonal component of data sequence as it restricts statistical ARIMA model to learn trend in data sequence. Thus, seasonal component is excluded from data sequence which makes ARIMA model smoother. In the proposed novel hybrid ARIMA-

BiLSTM model seasonal component of data sequence is extracted from additive seasonal decompose and included in data sequence of deep BiLSTM model for further computations. Hence, the developed model is named as ARIMA-BiLSTM. Figure 4 depicts architecture of the hybrid model. The initial phase of ARIMA model is an additive decomposition and stated as $y_t = s_t + \tau_t + \epsilon_t$ where, $s_t$ is trend, $\tau_t$ is seasonal component and $\epsilon_t$ is noise. $s_t$ and $\epsilon_t$ are retained to reproduce trend and noise. $\tau_t$ is seasonality decomposition which is used in the data sequence of BiLSTM model for achieving the better performance.

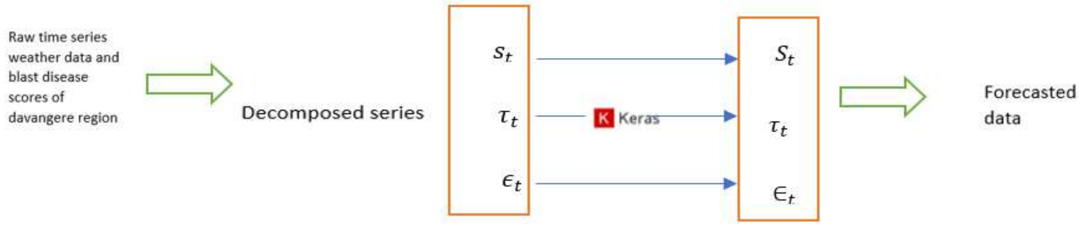


**Figure 4: Data pipelines for the Hybrid Model**

**Algorithm of Novel Hybrid ARIMA-BiLSTM Model**

**ALGORITHM: FORECASTING OF BLAST DISEASE OUTBREAKS USING HYBRID ARIMA-BILSTM MODEL**

```
Input: Time Series weather data and blast disease score data
Output: MSE, MAE of the proposed ARIMA-BiLSTM model
Model:=seasonal_decompose
Result:=model(blast disease scores, additive, period=30)
Series:=drop.column['BlastDiseaseScore']
Result1:=series+result
sizeofdata := length(Result1) * 0.80
training := Result1[0...size]
testing := Result1[sizeofdata...length(Series)]
set random.seed(7)
procedures fitting_ARIMA-Bilstm(training, epochs, neurons) do
    | X := training
    | y := training – X
    | modelling_hybrid := sequentialmodel()
    | modelling_hybrid.adds (Bidirectional (LSTMmodel (neurons,stateful=True))
    | modelling_hybrid.compilation(loss='mean_square_error',
    | optimization='adam')
    | for each of i in the range of(epochs) do
    |         | modelling_hybrid.fit(X, y, epochs=30, shuffling=False)
    |         | modelling_hybrid.reset_states()
    | end
return model
procedures forecasting_hybridBiLstm(modelling_hybrid, x) do
    | yha := modelling_hybrid.prediction(X)
return yha
epochs:=30
neurons:=64,32
prediction := empty
hybridBiLstm_modelling := fitting_hybridBiLstm(training,epochs,neurons)
hybridBiLstm_modelling. prediction(training)
for each of i in the range of(length(testing)) do
    | X := testing[i]
    | yha := forecasting_BiLstm(hybridBiLstm_model, X)
    | prediction.append(yha)
    | expected := testing[i]
end
MSE := mean_squarederror(expected, prediction)
MAE:= mean_absoluteerror(expected,prediction)
return(MSE,MAE)
```

## 4. Experimental Results and Comparative Analysis

The results of the research study are reported in this section which includes a brief discussion of long-term trends predicted by the models. The investigation was conducted on a testbed that included a 6C/ 12T Ryzen 5 3600 CPU running at 3.6 GHz, 16 GB 3000 Mhz DDR4 RAM. Because of the existence of strong high-end libraries like **numpy, tensorflow, statsmodels, and scikit learn**, the development code in the proposed **study is written in Python**. This helped to reduce the overall complexities of the code without affecting efficiency and performance.

### 4.1 Dataset Description

To achieve proposed objective on forecasting of blast disease occurrence in Davangere district, proposed study has used the historical climate data and blast disease occurrence data from the year 2015-2019. Dataset obtained from Karnataka State Natural Disaster and Monitoring Centre (KSNDMC) Bengaluru, which has designed a robust Real-time Weather Monitoring Program through deploying computerized, solar-powered, GPRS-enabled 5,978 Telemetry data Rain gauges at Grampanchayath and BBMP levels, as well as 747 Telemetric Weather Stations at Hobli level and taluk level across Karnataka. Figure 5 explains daily recorded climate data and blast disease occurrence data from the year 2015-2019. Dataset has seven important climate features and one feature on occurrence of blast disease in Davangere district. Seven weather features are minimum temperature, maximum temperature, temperature difference, minimum humidity, maximum humidity, humidity difference and rainfall.

| RECORDED_DATE | MINTEMP | MAXTEMF | TEMP_DIF | MINHUMI | MAXHUMI | HUMIDITY | RAINFALL | Blast Disease Occurrence |
|---|---|---|---|---|---|---|---|---|
| 01-01-2015 | 20.63333 | 32.41667 | 11.78333 | 37.75 | 91.43333 | 53.68333 | 0.083333 | 1 |
| 02-01-2015 | 16.51667 | 31.26667 | 14.75 | 38.33333 | 97.31667 | 58.98333 | 0 | 0 |
| 03-01-2015 | 16.85 | 32.08333 | 15.23333 | 26.21667 | 95.06667 | 68.85 | 0 | 0 |
| 04-01-2015 | 15.86 | 31.26667 | 15.4 | 39.18333 | 95.48333 | 56.3 | 0 | 0 |
| 05-01-2015 | 17.25 | 32.53333 | 15.28333 | 28.08333 | 89.43333 | 61.35 | 0 | 0 |
| 06-01-2015 | 16.53333 | 31.45 | 14.91667 | 37.96667 | 90.4 | 52.43333 | 0 | 0 |
| 07-01-2015 | 17.38333 | 31.4 | 14.01667 | 37.56667 | 88.36667 | 50.8 | 0 | 0 |
| 08-01-2015 | 17.1 | 31.86667 | 14.76667 | 32.53333 | 90.66667 | 58.13333 | 0 | 0 |
| 09-01-2015 | 18.01667 | 31.73333 | 13.71667 | 32.06667 | 89.93333 | 57.86667 | 0 | 0 |
| 10-01-2015 | 14.48333 | 31 | 16.51667 | 35.16667 | 82.63333 | 47.46667 | 0 | 0 |
| 11-01-2015 | 10.23333 | 30.78333 | 20.55 | 22 | 88.8 | 66.8 | 0 | 0 |
| 12-01-2015 | 10.63333 | 30.73333 | 20.1 | 16.13333 | 74.46667 | 58.33333 | 0 | 0 |
| 13-01-2015 | 11.01667 | 29.6 | 18.58333 | 16.93333 | 79.73333 | 62.8 | 0 | 0 |
| 14-01-2015 | 10.55 | 30.8 | 20.25 | 16.43333 | 77.11667 | 60.68333 | 0 | 0 |
| 15-01-2015 | 13.31667 | 31 | 17.68333 | 23.21667 | 88.76667 | 65.55 | 0 | 0 |
| 16-01-2015 | 12.53333 | 30.6 | 18.06667 | 26.61667 | 76.36667 | 49.75 | 0 | 0 |
| 17-01-2015 | 15.15 | 31.2 | 16.05 | 27.78333 | 83.75 | 55.96667 | 0 | 0 |
| 18-01-2015 | 15.45 | 31.3 | 15.85 | 31.58333 | 84.96667 | 53.38333 | 0 | 0 |
| 19-01-2015 | 15.51667 | 32.65 | 17.13333 | 29.38333 | 78.13333 | 48.75 | 0 | 0 |
| 20-01-2015 | 17.51667 | 32.25 | 14.73333 | 28.08333 | 79.6 | 51.51667 | 0 | 0 |

**Figure 5: Dataset used in the Proposed Study**

### 4.2 Performance Evaluation Metrics

The proposed work examined statistical, deep learning and hybrid models for the occurrence of blast disease in Davangere district from 3.4 to 3.6. The next major concern is off-course implementations, or the use of these algorithms to forecasting. When applying a modelling to a real or simulated time series, the raw data is first separated into two parts, the training dataset and the testing dataset. The desired model is built using the data from the training set. The validation set is a tiny subset of the training set that is preserved for validation. Pre-processing might include normalising data or applying logarithmic or other transformations. After constructing a model, it is applied to generate forecasts. The test set observations are retained in order to see how well the fitted model forecast the tested values. The forecasted values are subjected to an inverse transformation to return to their original scale. The relative performance of multiple models on the test dataset is taken into account when assessing the forecasted accuracy of a particular model or assessing and comparing between the models. Because time series forecasting is so important in so many real-world scenarios, appropriate attention should be made while choosing a model. As a result, many performance metrics for estimating forecast accuracy and comparing different models have been developed in the literature. These are sometimes referred to as performance measures. Each of these measurements is a function of the time series' actual and forecasted quantities.

**Summary of numerous Forecasting performance Metrics**

In the current section performance measures used in the proposed study and their important properties are described. In each of the forthcoming definitions, $y_t$ is the actual value, $f_t$ is the forecasted value, $e_t = y_t - f_t$ is the forecast error and n is the size of the test set. Furthermore, $\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$ is the test mean and $\sigma^2 = \frac{1}{n-1}\sum_{t=1}^{n}(y_t - \bar{y})^2$ is the test variance.

**The Mean Absolute Error (MAE)**

The Mean Absolute Error is defined as MAE=$\frac{1}{n}\sum_{t=1}^{n}|e_t|$ Its features are:

- It calculates the average absolute difference between forecasted and actual values.
- It is also called as the Mean Absolute Deviation (MAD).
- It depicts the magnitude of the overall inaccuracy caused by predicting.
- In MAE, positive and negative error will not cancel out each other.
- Unlike MFE, MAE does not offer any evidence of error direction.
- The estimated MAE must be as low as feasible for a successful forecast.
- MAE is also influenced by measurement scale and data processing.
- MAE does not panelise extreme forecast inaccuracies.

**The Mean Squared Error (MSE)**

Mathematical definition of this measure is MSE=$\frac{1}{n}\sum_{t=1}^{n} {e_t}^2$ Its properties are:

- It is a measure of average squared deviation of forecasted values.
- The opposite signed errors do not offset one another, MSE gives an overall idea of the error occurred during forecasting.
- It panelises extreme errors occurred while forecasting.
- MSE emphasizes the fact that the total forecast error is in fact much affected by large individual errors, i.e. large errors are much expensive than small errors.
- MSE does not provide any idea about the direction of overall error.
- MSE is sensitive to the change of scale and data transformations.
- Although MSE is a good measure of overall forecast error, but it is not as intuitive and easily interpretable as the other measures discussed before.

Two important performance measures are explained for judging forecast error of a fitted models. Each of these measures has some unique properties which are different from others. In experiments of time series forecasting, it is better to consider more than one performance criteria which will help to obtain a reasonable knowledge about the amount, magnitude and direction of overall forecast error. For this reason, time series analysts usually use more than one measure for judgment.

### 4.3 Experimental Results

Experimental results of statistical ARIMA model, Deep BiLSTM model and novel hybrid ARIMA-BiLSTM model for forecasting of rice blast disease outbreaks for sustainable rice production is accomplished using testing stationary, seasonal decompose and forecasting of Blast Disease Outbreaks. More detail description of aforementioned approaches are as follows.

#### 4.3.1 Testing Stationarity

In the current section testing Stationarity of the data is presented. Augmented Dicky Fuller (ADF) test is used to screen for stationarity. The ADF test is a stationarity test, which is a statistical test. The unit test's aim is to determine how significantly a time series is influenced by a trend.

- Null-Hypothesis (H0): A non-stationary unit root can be used to represent the time series.
- Alternative-Hypothesis (H1): The time series remains unchanging where the p-value and its interpretation are defined.

  P value > 0.05: Accept the Null Hypothesis

  P value < 0.05: Reject the Null Hypothesis

From the Table 1 it is observed that p-value < 0.05 hence, reject the null hypothesis. Since data found to be stationary lags are not applied.

**Table 1: Results of ADF Test and its Critical Values**

| **Results of ADF Test** | |
|---|---|
| ADF | -7.07490442766327 |
| P-VALUE | 4.830110167116215e-10 |
| NUM OF LAGS | 18 |
| NUMBER OF OBSERVATIONS | 9111 |
| **CRITICAL VALUES** | |

| | |
|---|---|
| 1% | -3.431067939122514 |
| 5% | -2.8618572829761098 |
| 10% | -2.56693888467881 |

### 4.3.2 Seasonal Decompose

The "seasonal decompose ()" function in python is used to estimate the trend component and seasonal component of a time - series data which may be explained using an additive or multiplicative decompose functions. According to the time series blast disease occurrence data in the Davangere region, the occurrence of blast disease is seasonal, reaching a peak during kharif season (July-November) of the crop cultivation and troughing every summer. Estimated residual and trend components of data sequences are shown in Figures 6 to 7. The seasonal pattern for the blast disease data sequence is identical from the beginning of the series to the pattern at the ending. Seasonal variables recovered from additive seasonal decomposition are used to build a unique ARIMA-BiLSTM hybrid model, some seasonal values derived from additive seasonal decompose is shown in Table 2

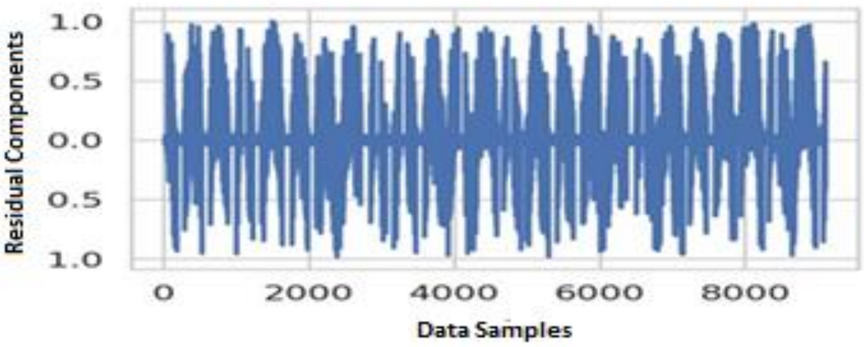


**Figure 6: Residual Component of Occurrence Blast Disease Score from Seasonal Additive Decompose**

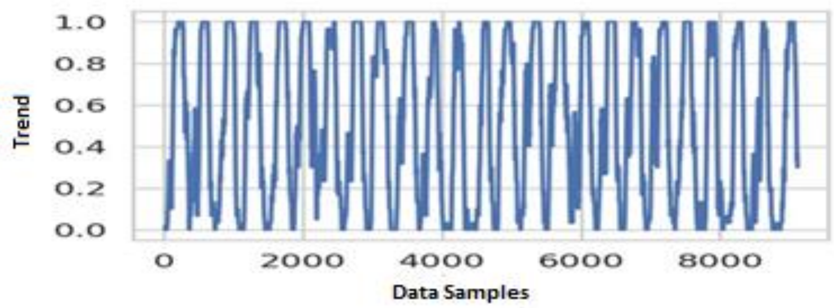


**Figure 7: Trend Component from Seasonal Additive Decompose**

**Table 2: Seasonal Component of Blast Disease Scores**

| | | | | | |
|---|---|---|---|---|---|
| 0.007746 | -0.012111 | -0.002265 | -0.012166 | 0.020837 | -0.005565 |
| -0.015466 | -0.028667 | 0.007636 | -0.005565 | 0.005123 | 0.008726 |

Figure 8 and 9 shows plots of autocorrelation and partial autocorrelation. It can be seen that data is stationary, hence differencing is not necessary. For forecasting future values using ARIMA model three values are important namely p, d and q. Since, differencing is not applied d is considered as 0. For selecting values for p and q, plots of autocorrelation and partial autocorrelation are considered. From partial autocorrelation we can derive AR which is the value for p and from autocorrelation plot MA can be derived which is the value for q. From the Figure 9 shows that after lag 3, successive values are shut off near to 0 hence, p can be considered as 3. Similarly, from

Figure 8 it is observed that values from lag 1 to 25 are almost same. Hence, it can be derived that q equal to 1. Forecasting the occurrence of blast disease for the Davangere distirct can be achieved using ARIMA(p=3,d=0,q=1) model.

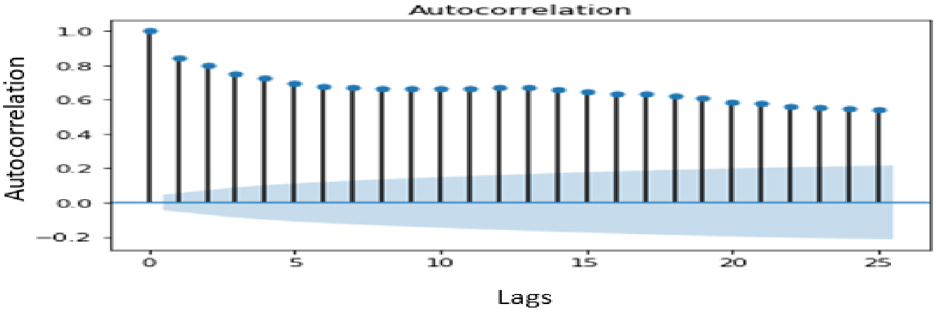


**Figure 8: Autocorrelation Plots of Daily Occurrence of Rice Blast Disease**

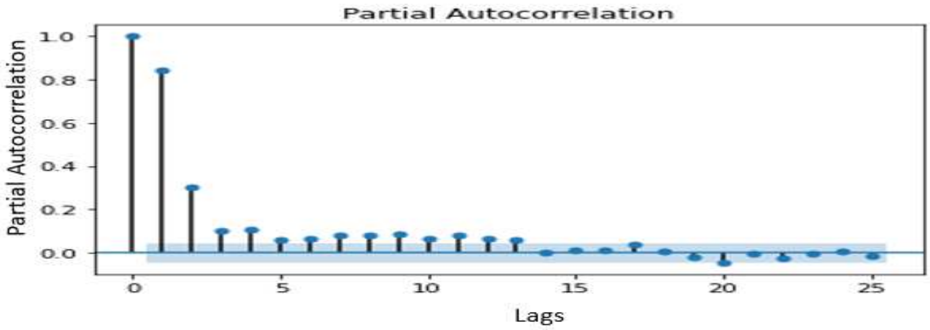


**Figure 9: Partial Autocorrelation Plots of Daily Occurrence of Rice Blast Disease**

We have used Keras open-source deep learning package with TensorFlow at the backend to develop the proposed BiLSTM and novel hybrid ARIMA-BiLSTM model in Python. The dataset is initially screened to find null values. Further, using min-max normalisation, the data is normalised between 0 and 1. The data is divided into two sections: training and testing. The first 80% of the data is set aside for training, with the remaining 20% set aside for testing. Train data is converted to 3 dimensional considering timesteps=20. The 80 percent of training data is used to train both BiLSTM and novel hybrid model. Table 3 provides data variable that are initialized to develop proposed BiLSTM and ARIMA-BiLSTM hybrid models.

**Table 3: Summary of BiLSTM and hybrid Model**

| Summary of BiLSTM and hybrid Model | |
|---|---|
| Epochs | 30 |
| Batch size | 32 |
| Hidden neurons | 64,32 |
| Dropout rate | 0.25 |
| Dense | 1 |
| Trained shape | 3dimensional |

### 4.3.3 Forecasting of Blast Disease Outbreaks

The traditional ARIMA model, deep BiLSTM model and last hybrid ARIMA-BiLSTM model are used to forecast occurrence of rice blast disease in Davangere region. Performance of these models are decided based on the values of MSE and MAE. Table 4 also describes the performance of each forecasting model it can be observed that MAE and MSE values of hybrid ARIMA-BiLSTM is very less and near to 0 compared to the performance of other two models. Thus, it can be stated that novel hybrid model has achieved good performance in forecasting occurrence of rice blast disease in Davangere region. The actual vs predicted values of BiLSTM and ARIMA-BiLSTM are represented in Figures 11, 12 and 14, 15. Figure 15 reveals that proposed novel hybrid ARIMA-BiLSTM has achieved best model performance compared to other two models. Forecasted values of blast disease occurrence are shown in Figure 11 and 14, respectively, for deep BiLSTM learning model and novel hybrid ARIMA-BiLSTM model where the portion shaded in green colour represents training data, portion shaded in red colour represents the forecasted values, and portion shaded in blue colour represents actual values. Thus, forecast values obtained serves as an indicator of the arrival of blast disease population causing attacks on rice commodities.

After the training and verification of the deep BiLSTM neural network and novel hybrid ARIMA-BiLSTM, loss function values are obtained. The loss function values of the BiLSTM and ARIMA-BiLSTM models are shown in Figure 10 and 13, where $x$ axis represents the number of epochs and the $y$ axis represents the loss values of the model. It can be concluded from the Figures 10 and 13 that the loss function values from the model's training set is decreased with an increase in the number of epochs. From the Figure 10 it can be observed that the loss function values are fluctuated slightly at epoch 4 of the deep BiLSTM model, and the loss function values generated by the hybrid ARIMA-BiLSTM model found relatively stable and consistent as represented in Figure 13. The overall loss function values of both the models are decreased and converged, there is no gradient explosion or dispersion phenomenon, and the network convergence speed found faster.

### 4.4 Comparative Analysis and Performance Evaluation

The optimal model is found by comparing MAE and MSE values. As indicated in Table 4, the MAE and MSE values of all the three models are compared to choose the optimal model for the forecasting of occurrence of rice blast disease. Table 4 reveals that the deep BiLSTM model and the novel hybrid ARIMA-BiLSTM model outperformed than the classic ARIMA model. To look more closely, results described in the Table 4 also reveals that the novel ARIMA-BiLSTM is better suited in forecasting occurrence of rice blast disease compared to deep BiLSTM model.

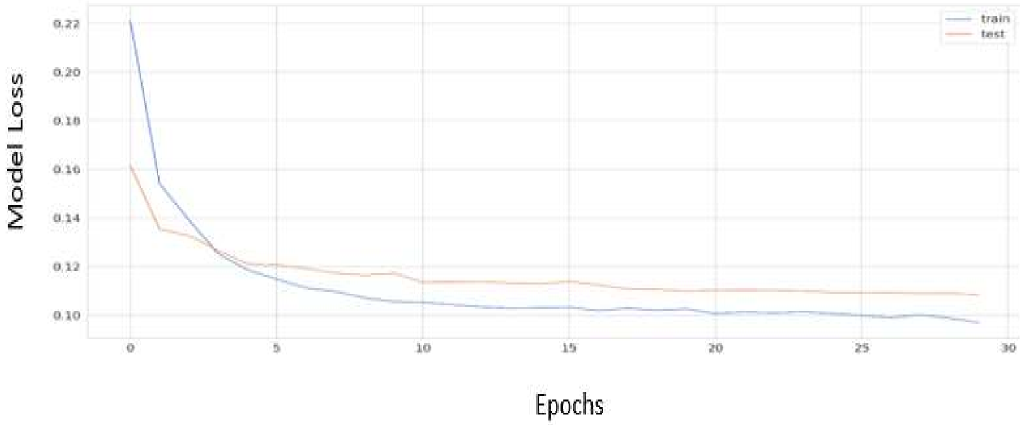


**Figure 10: BiLSTM Model Loss for Training and Testing data**

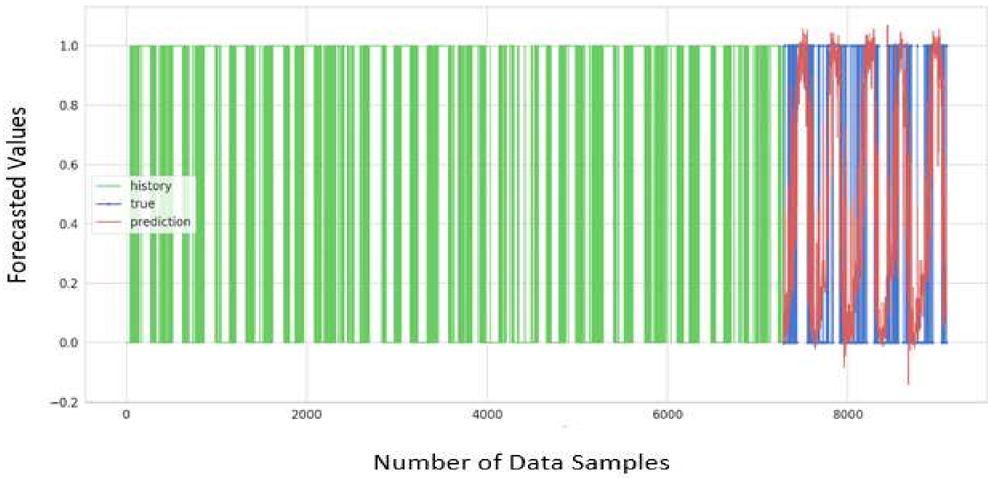

**Figure 11: Forecast Plots of Daily Occurrence of Rice Blast Disease for BiLSTM Model with Shaded Region in Green Colour Representing Training Data**

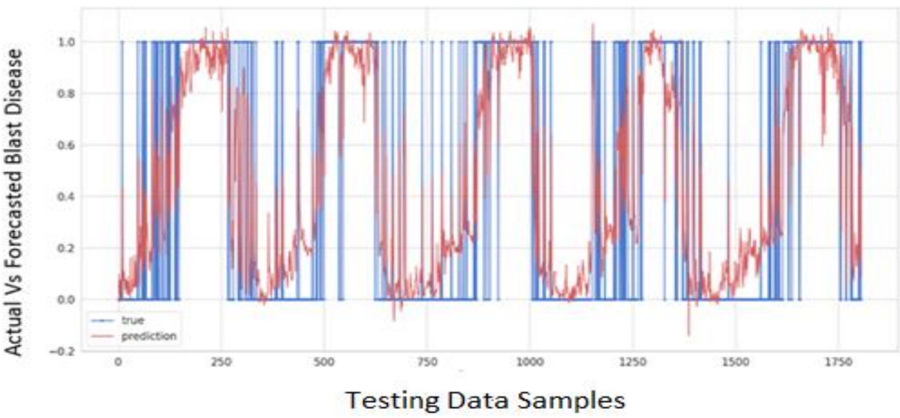


**Figure 12: Forecast Plots of Daily Occurrence of Rice Blast Disease for BiLSTM Model with Shaded Region in Blue Colour Representing True Values and Red Colour Predicted Values**

**Table 4: MAE and MSE Values for Comparison of Models.**

| | **MAE** | **MSE** | **Average Rank** |
|---|---|---|---|
| ARIMA | 0.492316 | 0.496358 | 3 |
| BiLSTM | 0.1979 | 0.2018 | 2 |
| Hybrid | **0.02815** | **0.03727** | 1 |

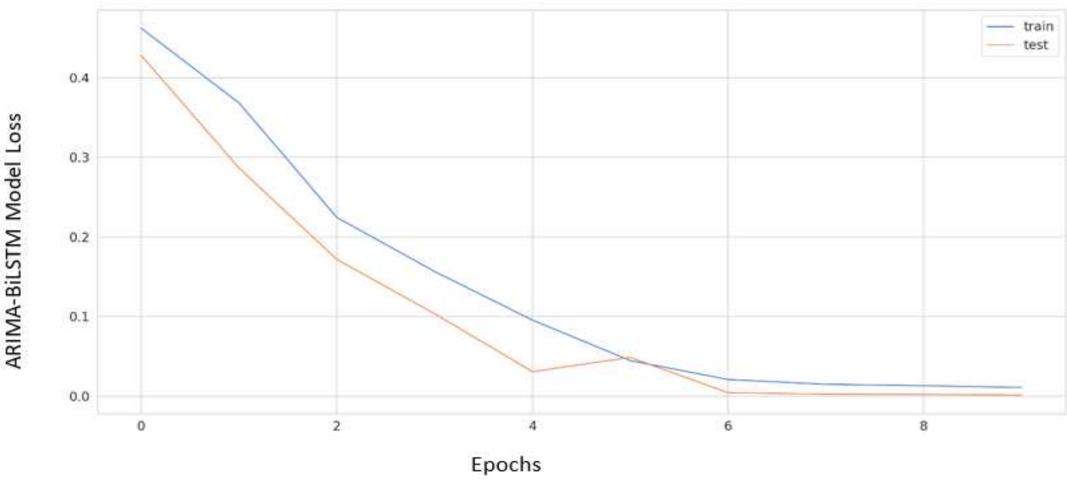


**Figure 13: ARIMA- BiLSTM Model Loss Training and Testing Data**

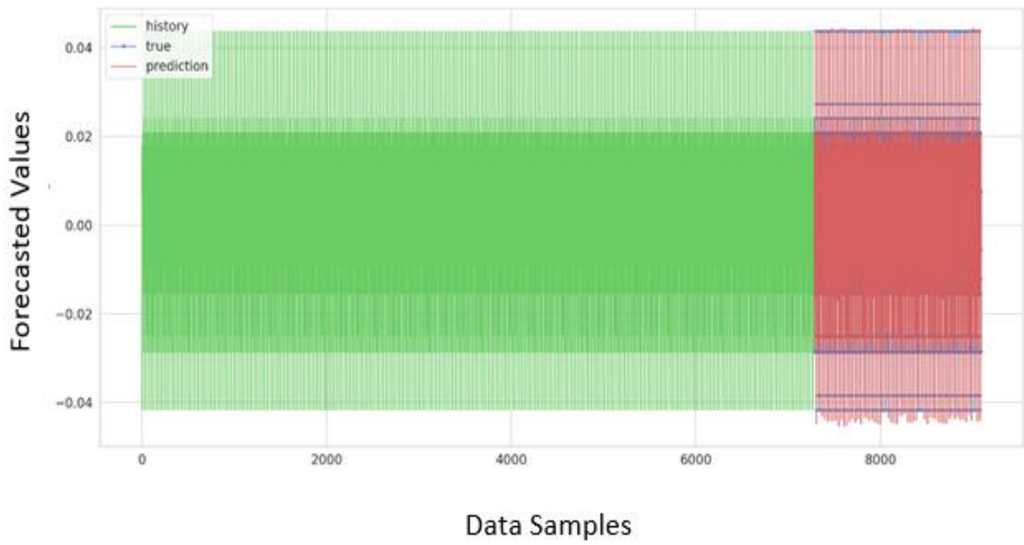


**Figure 14: Forecast Plots of Daily Occurrence of Rice Blast Disease for ARIMA-BiLSTM Model with Shaded Region in Green Colour Representing Training Data**

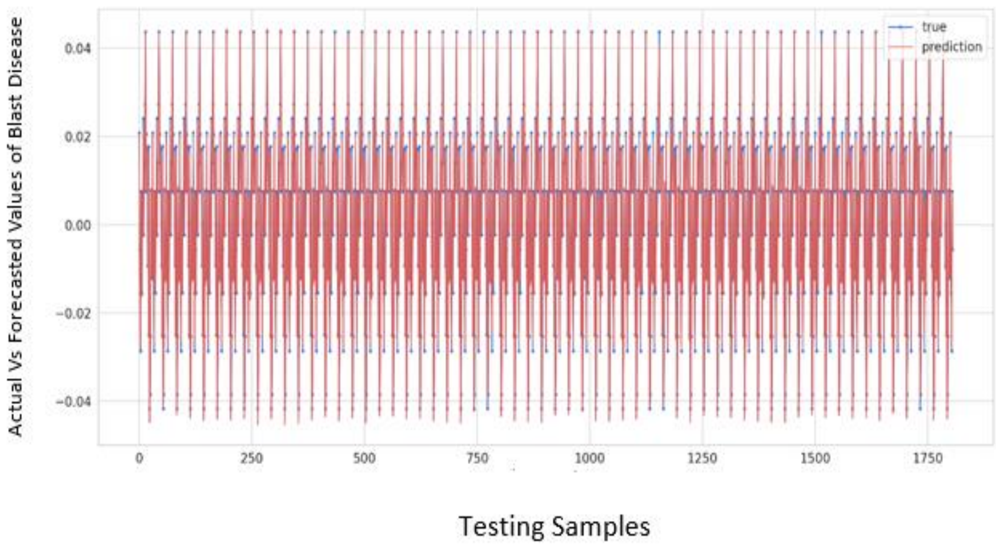


**Figure 15: Forecast Plots of Daily Occurrence of Rice Blast Disease for ARIMA-BiLSTM Model with Shaded Region in Blue Colour Representing True Values and Red Colour Predicted Values**

## 5. Conclusion

The rice blast disease of paddy crop is exponentially spreading in the paddy growing regions because of varying climate conditions, and the agricultural systems adopted in some high impacted regions of various states, such as West Bengal, Tamil Nādu, Madhya Pradesh and the Karnataka. Accurately forecasting the occurrence of rice blast disease provides pertinent information to governments and agriculture scientists about the expected situation and the needed measures to impose. Thus, forecasting information can be useful for motivating the wider public farmers to consider the imposed measures for down slowing the spread of the disease. In the proposed study, statistical time series model ARIMA and NN-based models including, BiLSTM, and ARIMA-BiLSTM have been applied to the real-time weather data to forecasts daily occurrence of rice blast disease in four regions of Davangere district. The choice is highly motivated by the extended capacity of deep learning models in capturing process of nonlinearity and their flexibility in modelling time-dependent data. The performance of each model has been verified in terms of MAE and MSE. Results demonstrates that the novel hybrid model ARIMA-BiLSTM has achieved better forecasting performance in comparison to ARIMA and BiLSTM models.

**Acknowledgements**

We would like to thank Director of Karnataka State Natural Disaster and Monitoring Centre Bengaluru and his colleagues for providing climate data of Davangere region.  We also thank Agriculture and Horticulture Research Station Kathalgere, Davangere District and Taralabalu Krushi Vigyana Kendra for their constant support in understanding occurrence of Pathogen

**Declaration**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**References**

[1] C. -J. Chen, Y. -Y. Huang, Y. -S. Li, Y. -C. Chen, C. -Y. Chang and Y. -M. Huang, "Identification of Fruit Tree Pests With Deep Learning on Embedded Drone to Achieve Accurate Pesticide Spraying," in *IEEE Access*, vol. 9, pp. 21986-21997, 2021, doi: 10.1109/ACCESS.2021.3056082.

[2] Hiary, Hazem, S. B. Ahmad, M. Reyalat, Malik Braik and Zainab Al-Rahamneh. "Fast and Accurate Detection and Classification of Plant Diseases." *International Journal of Computer Applications* 17 (2011): 31-38.

[3] R. Hegazy, A. Schmidley, E. Bautista, D. Sumunistrado, M. Gummert, "Mechanization in rice farming: lessons learned from other countries," *Asia Rice Found.*, pp. 1–49, 2013, [Online]. Available: https://www.researchgate.net/publication/258951878_Mechanization_in_rice_farming_lessons_learned_from_other_countries.

[4] I. Setiawan, I. Made Sumertajaya, and F. Mochammad Afendi, "Forecasting The Broad Proportion Attack of Rice Blast Disease in Indonesia," *IJCSN -International J. Comput. Sci. Netw.*, vol. 6, no. 65, 2017.

[5] O. Singh, J. Bathula, and D. Singh, "Rice Blast Modeling and Forecasting," *Int. J. Chem. Stud.*, vol. 7, no. 6, pp. 2788–2799, 2019.

[6] L. W. Liu, S. H. Hsieh, S. J. Lin, Y. M. Wang, and W. S. Lin, "Rice blast (Magnaporthe oryzae) occurrence prediction and the key factor sensitivity analysis by machine learning," *Agronomy*, vol. 11, no. 4, 2021, doi: 10.3390/agronomy11040771.

[7] Katsantonis, D.; Kadoglidou, K.; Dramalis, C.; Puigdollers, P. Rice blast forecasting models and their practical value: A review. Phytopathol. Mediterr. 2017, 56, 187–216.

[8] Abe, T. On the influence of soil temperature upon the development of the blast disease of rice. Forsch. Gebiet Pflanzenkr. 1933, 2, 30–54.

[9] Hashioka, Y. Studies on the rice blast disease in the tropics. IV. Influence of temperature of air and soil upon the resistance of the rice plants to the blast disease. Jpn. Soc. Trop. Agric. 1944, 15, 163–176.

[10] Asibi, A.E.; Chai, Q.; Coulter, J.A. Rice Blast: A Disease with Implications for Global Food Security. Agronomy 2019, 9, 451.

[11] Calvero, S.; Coakley, S.; Teng, P. Development of empirical forecasting models for rice blast based on weather factors. Plant Pathol. 1996, 45, 667–678.

[12] Kaundal, R.; Kapoor, A.S.; Raghava, G.P. Machine learning techniques in disease forecasting: A case study on rice blast prediction. BMC Bioinform. 2006, 7, 485.

[13] Kim, Y.; Roh, J.-H.; Kim, H.Y. Early forecasting of Rice blast disease using long short-term memory recurrent neural networks. Sustainability 2018, 10, 34.

[14] Nettleton, D.F.; Katsantonis, D.; Kalaitzidis, A.; Sarafijanovic-Djukic, N.; Puigdollers, P.; Confalonieri,

R. Predicting rice blast disease: Machine learning versus process-based models. BMC Bioinform. 2019, 20, 514.

[15] K. Manibhushanrao and P. Krishnan, “Epidemiology of Blast (Ebipla) - a Simulation-Model and Forecasting System for Tropical Rice in India,” *Rice Blast Model. Forecast.*, pp. 31–38, 1991.

[16] R. Prasad, A. Sharma, and S. Sehgal, “Influence of weather parameters on occurrence of rice blast in mid hills of Himachal Pradesh,” *Himachal J. Agric. Res.*, vol. 41, no. 2, pp. 132–136, 2015, [Online]. Available: http://www.hillagric.ac.in/research/HJAR/publication/HJAR_Vol/Volume-41(2)/5.HJAR-41(2)-132-136-2015-Prasad-etal.pdf

[17] C. K. Kim, “Epidemiology of Rice Blast Disease in Korea,” *Major Fungal Dis. Rice*, pp. 119–129, 2001, doi: 10.1007/978-94-017-2157-8_9.

[18] D. Babula, M. Kaczmarek, P. A.Zi´, Ołkowski, and J. Adowski, “Genome Mapping and Molecular Breeding in Plants,” *Brassica oleracea*, pp. 227–324, 2007.

[19] Y. Uehara, M. Imoto, and Y. Sakai, “Studies on the forecasting of the rice blast development using weather data from AMeDAS,” *Bull. Hiroshima Prefer. Agric. Exp. Stn.*, vol. 51, 1988.

[20] K. Ishiguro and A. Hashimoto, “Recent advances in forecasting of rice blast epidemics using computers in Japan.,” *Trop. Agric. Res. Ser.*, no. No. 22, pp. 153–162, 1989, [Online]. Available: https://www.cabdirect.org/cabdirect/abstract/19922322495.

[21] L. Castrejón, K. Kundu, R. Urtasun, and S. Fidler, “Annotating object instances with a polygon-RNN,” *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 4485–4493, 2017, doi: 10.1109/CVPR.2017.477.

[22] Q. Tao, F. Liu, Y. Li, and D. Sidorov, “Air Pollution Forecasting Using a Deep Learning Model Based on 1D Convnets and Bidirectional GRU,” *IEEE Access*, vol. 7, pp. 76690–76698, 2019, doi: 10.1109/ACCESS.2019.2921578.