Analysis for the Prediction of Solar and Wind Generation of India using ARIMA, Linear Regression and Random Forest

MD Wazid Ansari (2K19/AE/037)[1], Sahil Yadav (2K19/AE/050)[2]

Delhi Technological University(DTU) Shahbad Daulatpur, Main Bawana Road, Delhi, 110042, Delhi

Abstract

In this paper, we introduced Solar energy and Wind energy prediction, where the technologies of renewable energy are clean sources of energy i.e., their impact on the environment is much lower than the technologies of conventional energy. However, renewable energy like solar and wind sources are generated regularly therefore we don't know how much the resource size of energy storage is required in microgrids in order to utilize these energies. Therefore, in this project, we are predicting the renewable energy (solar and wind) which can help to ensure the adequate resource size of energy storage in microgrids. We have introduced various machine learning algorithms such as logistic regression and random forest and the ARIMA (Autoregressive integrated moving average) time series algorithms in order to predict renewable energy. We predicted the renewable energies 1 year ahead in the future with each algorithm and the performance of each algorithm is evaluated using the mean absolute error (MAE), mean squared error (MAE), root mean squared error (RMSE) and mean absolute percentage error (MAPE). The Lower the values of these matrices will be, the better will be the performance of the algorithm. The MAE value for the ARIMA (0.06 and 0.20) model for solar and wind energy is very less as compared to Random Forest (15.65 and 61.73) and Linear Regression (15.78 and 54.65) of solar and wind energy. Same with MSE and RMSE, the MSE and RMSE value for the ARIMA of solar energy model obtained is 0.01 and0.08 and wind energy is 0.07 and 0.27 respectively, which is very low as compared to the MSE, RMSE values for Random Forest and Linear Regression for both solar and wind energy. But the MAPE value for ARIMA (32.07 and 21.42) is relatively higher than the MAPE value for Random Forest (0.12 and 0.40) and Linear regression (0.12 and 0.35). After comparing all of these matrices of each algorithm for both the dataset, we concluded that the ARIMA model is best fit for the forecasting of renewable energy (solar energy and wind energy).

Keywords: Renewable energy, Solar energy, Wind energy, Machine Learning, Linear Regression, Random Forest, Time series, ARIMA, MAE, MSE, RMSE, MAPE.

INTRODUCTION

Electricity is generated from a variety of sources, including hydropower, nuclear power, and renewable energy. Coal, oil, or natural gas are used to produce power, account for 1/3rd of global greenhouse gas productions. It is vital to raise citizens' quality of life by supplying clean and efficient electricity. To achieve the country's economic development goals, India's energy demand is increasing A growing supply of energy is a basic requirement for a nation's economic success. The National Electricity Plan of the Ministry of Power has created a 10-year plan of action with the purpose of offering power across the nation, and also provide an idea to ensure that power is supplied to residents efficiently and at an affordable price. In this way, the govt will accelerate and globalize the transition to RE technology in order to attain imperishable development. Renewable energy can be considered as having capacity to reach the demand of power and also minimize the pollution.

To save energy residents have motivated to use wind, solar as these are inexhaustible in nature. It goes without saying that sustainable energy is fewer costlier and hazardous. India has an aim of achieving 175 GW energy from renewable by 2022. Solar accounting for 100 GW, 10, 60 and 5 GW accounts for biomass, wind and hydro respectively. Renewable energy comes from natural processes that are renewed on a regular basis. The majority of renewable energy sources are environmentally beneficial and help to reduce carbon emissions, which in turn helps to battle global warming.

Renewable energy is seen as the most encouraging replacement of fossil fuel since it is clean, green, and regenerated over a large geographic region; yet, it also introduces unplanned uncertainty, endangering energy reliability and stability, particularly when it comes to large-scale renewable energy integration [1]. In the latest decades, the electricity market has moved its attention to renewable energy sources to minimize its greenhouse discharge during power production. Solar and wind have regularly been used as a combination due to their RE variety and reliability. Several approaches for predicting RE have been grown over time, all of which have emphasized on the efficiency of estimation techniques with no or small concern for the environmental conditions. Renewable energy can efficiently cut fossil energy use, minimize pollution, and promote the healthy growth of the social economy [2]. Due to the variable character of power production from solar and wind, the operators must regulate and maintain power system properly. The electricity production of photovoltaic- wind must be forecasted to schedule the power transmission for long and short term [3].

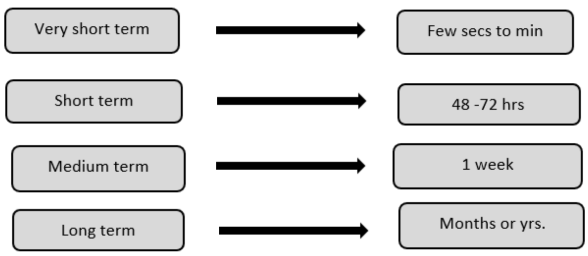


Fig 1. Horizon and Timeframe of Prediction

Solar Energy Prediction: Because of its reliance on a variety of elements such as pressure, temperature, speed of wind etc. solar radiation is most stimulating constraints. As these are unpredictable, the machine learning techniques is used for forecasting. It is the process of collecting and examining data in order to estimate solar power generation across a period of time with an aim of reducing solar intermittency's impact. Solar power forecasting is used to operate the electric grid more efficiently and for power trading.

Wind Energy Prediction: for the fluctuated nature of wind, it is necessary to do prediction. Due to fluctuation in speed of wind, forecast is not an easy task. To increase wind prediction efficacy, several ML techniques have been used. There are various techniques used in machine learning for the prediction such as

1. LR: Linear Regression
2. RF: Random Forest
3. ARIMA: Autoregressive integrated moving average
4. AR: Autoregressive etc. … Machine learning (ML) approaches are now widely used in a variety of renewable energy-related applications, including the growth of energy and unification, utilization, and forecasting [5]. The remaining part of this study is presented as follows: A proposed plan is conducted in Part 2. Study design and methodology is described in Part 3. At last, Part 4 and 5 brings the paper to an end with results and conclusion.

I. PROPOSED PLAN

In this study, for the prediction of Solar and Wind Energy, we are using three machine learning algorithms.

1. ARIMA
2. Logistic Regression
3. Random Forest

After applying these algorithms, we are checking the performance of each algorithms using following matrices:

1. Mean Absolute Error
2. Mean Squared Error
3. Root Mean Squared Error
4. Mean Absolute Percentage Error

Below flowchart shows the working flow for this study

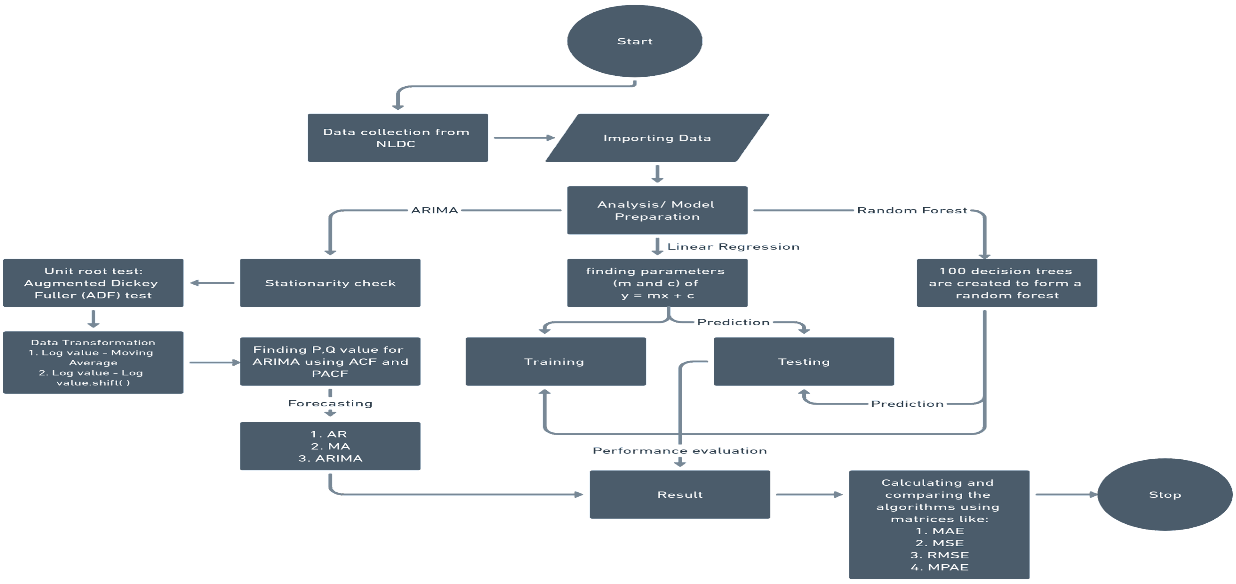


Fig 2. Flowchart for proposed model

1. ARIMA

ARIMA is used for time series analysis in which observations are collected in series at specified time intervals. With the help of these series, we can predict the next values on the basis of past data. Generally, in time series, we have only 2 variables - Time and the feature that we want to forecast. ARIMA is a combination of two models – a. AR (Autoregressive), b. MA (Moving Average)

ARIMA has three hyperparameters that need to be optimized in order to achieve better performance.

i. p (Autoregressive lags), ii. d (Order of differentiating), iii. q (Moving Average lags)

Before applying any time series model, stationarity of data is evaluated. Stationary means that over different period of time data should have –

i. Constant mean, ii. Constant variance and standard deviation, iii. Auto-covariance.

2. Linear Regression (LR)

Linear regression is used to predicts the value based on some input features. Training is done at the beginning with input and output features. Linear regression is of two types.

I. Simple linear regression

In this case, only one input feature is taken into consideration with one output variable.

$$= m + \qquad (1)$$

Where,

is the output variable of input

is the parameter that need to be optimized in order to find best fit line

is the input of the column

is the bias that need to be optimized in order to find best fit line

II. Multiple linear regression (MLR)

It have multiple input variables with one output variable. The equation for multi regression is:

$$= {}^{1} + {}^{2} + {}^{3} + ... + \quad (2)$$

is the output variable of the input

, , ... are the parameter that need to be optimized in order to find best fit line

^{1}, ^{2}, ^{3} are the inputs of the column ${}_{1}$, ${}_{2}$ and ${}_{3}$ respectively.

is the bias that need to be optimized in order to find best fit line.

3. Random Forest

It builds decision trees based upon the samples in classification. One of the key features of this algorithm has to be the ability of handling data sets having continuous variables for the cases of regression and categorical variables for classification. Having said that, this algorithm gives better results for classification problems as compared to regression (Torgo, 1996).

To be able to understand the working of random forest methods, a concept called the ensemble technique should be known. Ensemble utilizes two types of methods:

Bagging: As mentioned above, a different training subset is generated from the sample training data and majority voting determines the final output data. Random forest uses this method for classification.

Boosting: In this method of ensemble, the final model created gives the maximum accuracy. This is done by generating a sequential model upon combining weak learners with strong learners. Examples of this include ADA BOOST, XGBOOST.

4. **Mean Absolute Error**

MAE is used for measuring the absolute error difference between predicted values and actual values.

Absolute error equation:

$$(\quad) = |{}_{\text{predicted}} - {}_{\text{actual}}| \quad (3)$$

Mean absolute error equation:

$$MAE = | \quad (4)$$

Where:

$\sum$ is the summation notation, is the total number of observations, is the value of prediction of observation and is true value of observation.

5. Mean Squared Error

MSE is a statistical method used to find the squared error between predicted values and actual values. Less the value of MSE, better will be the performance of the machine learning model.

$$=)^{2} \quad (5)$$

6. Root Mean Squared Error

If the value of RMSE is less, then the performance of machine learning will be better.

$$= \quad (6)$$

7. **Mean Absolute Percentage Error**

MAPE is used to calculate the performance of machine learning algorithms but instead of accuracy, it measures the accuracy percentage of machine learning algorithms.

Equation for MAPE is:

= (7)

III. STUDY DESIGN AND METHODOLOGY

Materials

I. **Data collection** - For this study, we have collected solar and wind energy generation data in MU (million units) on a daily basis from the National Load Dispatch Centre. Total 1912 observations are collected for both solar and wind energy generation starting from 2017-01-01 to 2022-03-27.

II. **Tools used**

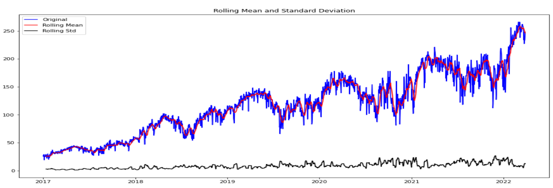
Following tools will be used for the implementation of this study: For machine learning implementation, python programming language is used because of its simplicity, compactness and better readability.

Analysis Descriptive summary of data

S. No.	Energy	Count	Mean	Std	Min	25%	50%	75%	Max
1.	Solar	1912	117.74	55.19	19	77	116.53	157.25	266
2.	Wind	1912	151.65	100.66	5.0	80.75	114	207	541

Table 1. Descriptive summary of solar and wind energy generation

As we can see from the above table, the total number of observed values for both the data is 1912. The maximum unit energy generated for solar energy is 266 MU while for wind energy is 541 MU. Standard deviation of solar energy is 55.19 while for wind is 100.66 i. e, there is much variance in wind data so it would be difficult for machine learning algorithms to perform well for wind data.



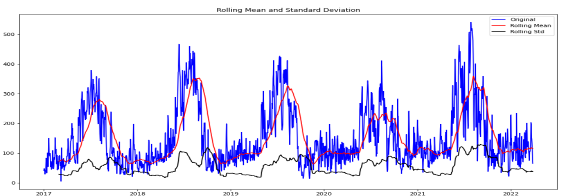


Fig 5. Rolling mean and standard deviation of solar energy generation (MU

Fig 6. Rolling mean and standard deviation of wind energy generation (MU)

Fig 5. and Fig 6. shows the rolling mean and rolling standard deviation of solar energy and wind energy generation. Both rolling mean and standard deviation are changing over time for solar energy i.e., they are not constant therefore, the solar energy are non-stationary in nature i.e., the values of solar energy generation are changing with time. However, the wind generation data looks somewhat non-stationary in nature. Let us use a unit root test to check whether the data is stationary or not. In this study, we are using Augmented Dickey Fuller test for stationarity.

AUGMENTED DICKEY FULLER TEST

S. No.	Energy	Test Statistics	p-value	Lags Used	No. of observations used
1.	Solar	0.542	0.883	23	1888
2.	Wind	3.986	0.001	12	1899

Table 2. Augmented dickey fuller test result

From Table 8. The ADF test result for both the energies are evaluated and compared. The p-value for solar energy obtained is 0.883 while for wind energy is 0.001. We are selecting p-value = 0.05 as our threshold value i.e., if p-value of any of the data of ADF test exceeds 0.05 then that data would be non-stationary and vice versa. From the table, we can say that solar energy data is non stationary in nature and wind energy is stationary. In order to use time series models like AR, MA, ARIMA, we have to use stationary data. To make our data stationary, we will take logarithmic value for both the data and then we will check the stationary using augmented dickey fuller test.

DATA TRANSFORMATION TO ACHIEVE STATIONARITY

To make the data stationary, first we'll take the log of our data then we will use a differencing method to achieve stationarity. For the accuracy purpose, we will take logarithmic values for wind data as well even though it is already stationary. To make this stationary, we will use a differencing method in which we will take the difference between the log values and moving average values of solar energy generation and wind energy generation.

Log scale(L) = stationary part(L1) + trend (LT)

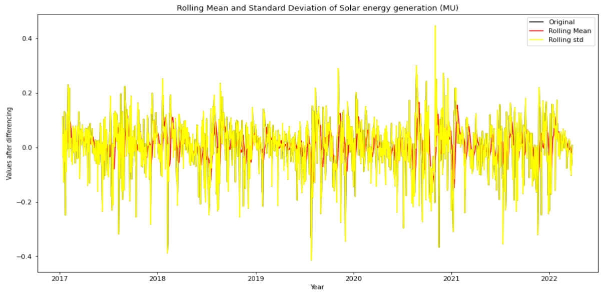
moving avg of log scale(A) = stationary part (A1) + trend (AT)

result series (R) = L − A = (L1 + LT) − (A1 + AT) = (L1 − A1) + (LT − AT)

Now, both log scale(L) and moving av of log scale(A) is part of series and moving average therefore, the trend will be almost the same i.e., LT - AT = 0 and trend component will be almost removed.

R = L1 – A1

Below graph (**Fig 9** and **Fig 10**) shows the representation of each energy after removal of trend.



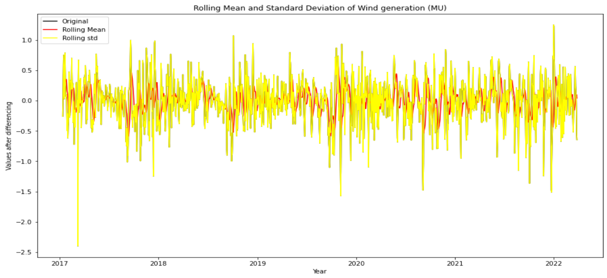


Fig 9. Rolling mean and standard deviation of solar energy generation (MU)

Fig 10. Rolling mean and standard deviation of wind energy generation (MU)

Black line of above graphs represents the original value i.e., result series (r) which is L1 – A1. Above graph seems stationary in nature i.e., their mean and standard deviation after some interval is almost same as any particular interval. Let's check the stationary of these graphs using unit root test (ADF).

ADF TEST

S. No.	Energy	Test Statistics	p-value	Lags Used	No. of observations used
1.	Solar	11.79	9.73e-22	21	1879
2.	Wind	12.16	1.46e-22	17	1883

Table 3. Augmented dickey fuller test result

From the above table (**Table 3**), the p-value of both the graphs are much smaller than the threshold p-value which is 0.05. The number of lags used by both the graphs are 21 and 17 respectively. Therefore, we have transformed the non-stationary data into stationary after taking the difference between log values and moving averages.

ACF AND PACF

I. Solar Energy

Fig 13. ACF plot for solar energy

Fig 14. PACF plot for solar energy

From the above graphs (**Fig 13** and **Fig 14**), for ACF graph of solar energy, the curve touches y = 0.0 line at x = 1 and for PACF graph, the curve touches y = 0.0 line at x = 1 as well therefore, for ACF, Q = 1 and for PACF, P = 1.

II. Wind Energy

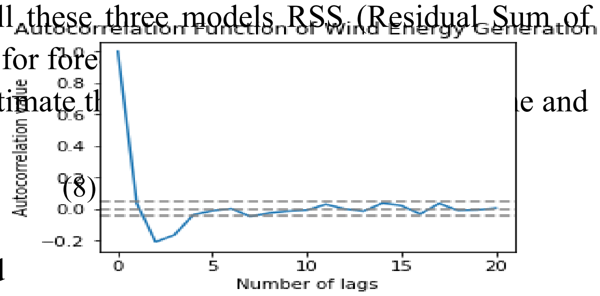


Fig 15. ACF plot for wind energy

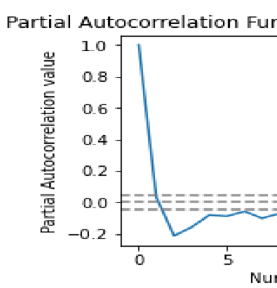


Fig 16. PACF plot for wind energy

Similarly, for the ACF graph of wind energy (**Fig 15**), the curve touches the y= 0.0 line at x = 1 and for the PACF graph (**Fig 16**), the curve touches the y = 0.0 line at x = 1 as well therefore, for ACF, Q = 1 and for PACF, P = 1.

BUILDING TIME SERIES MODEL

In this study, we are using time series models like AR, MA and ARIMA for forecasting the solar and wind energy generation 1 year ahead in future. Thus, to get the better model, we will compare all these three models RSS (Residual Sum of Squares) value to get the better model for fore... Residual Sum of Squares is used to estimate t... e and the regression function.

RSS= ^{2} (8)

Where:

= value of variable to be predicted

= predicted value ofterm

I. Solar Energy

Time series model 1: Auto Regressive Model (AR) (**Fig 17**) ARIMA (1,1,0)

Here, 1,1,0 denotes the absence of MA (Moving Average)

Time series model 2: Moving Average (MA) (**Fig 18**) ARIMA (0,1,1)

Here, 0,1,1 represents the absence of AR (Auto Regressive)

Time series model 3: Autoregressive Integrated Moving Average (**Fig 19**) (ARIMA (1,1,1))

Here,1,1,1represents the presence of both AR and MA with lag of 1 and differencing of order 1.

II. Wind Energy

Time series model 1: Auto Regressive Model (AR) (**Fig 20**) ARIMA (1,1,0)

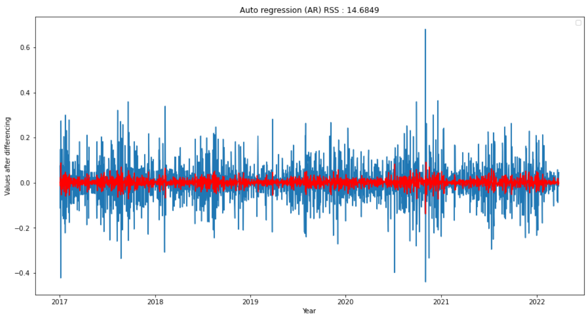
Here, 1,1,0 denotes the absence of MA (Moving Average)

Time series model 2: Moving Average (MA) (**Fig 21**) ARIMA (0,1,1)

Here, 0,1,1 represents the absence of AR (Auto Regressive)

Time series model 3: Autoregressive Integrated Moving Average (ARIMA) (**Fig 22**) (ARIMA (1,1,1))

Here, 1,1,1 represents the presence of both AR and MA with lag of 1 and differencing of order 1.



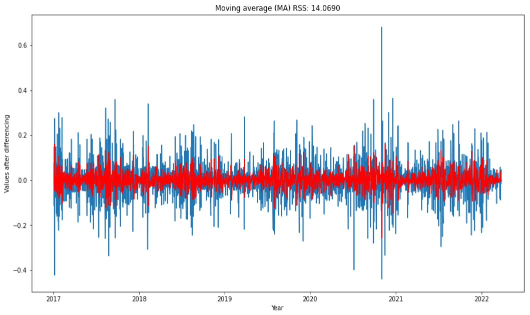
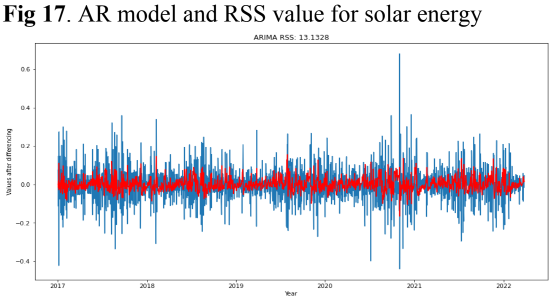


Fig 17. AR model and RSS value for solar energy **Fig 18**. Moving Average model and RSS value for wind



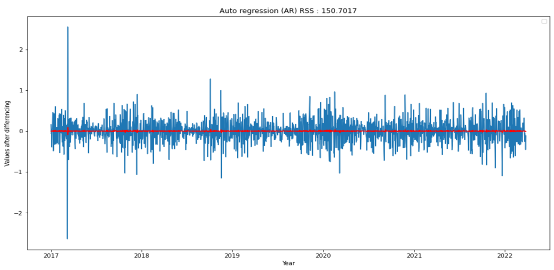
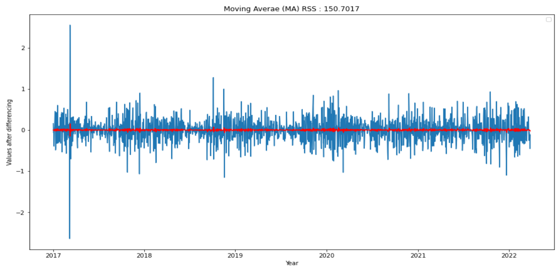


Fig 19. ARIMA model and RSS value for solar energy **Fig 20**. Autoregressive model and RSS value for wind energy



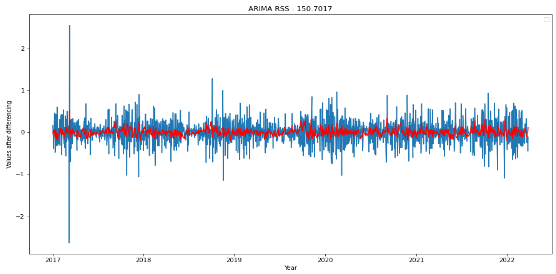


Fig 21. MA model and RSS value for wind energy

Fig 22. ARIMA and RSS value for wind energy

RSS VALUE

I. Solar Energy

S. No.	Algorithm	RSS VALUE
1.	AR (Auto Regression)	14.6849
2.	MA (Moving Average)	14.0690
3.	ARIMA (Autoregressive Integrated Moving Average)	13.1328

Table 5. RSS Value of time series model for solar energy

From the above table, the RSS value of ARIMA is lower than the RSS value of both AR and MA therefore, ARIMA will perform better for forecasting the solar energy generation as compared to AR and MA.

II. Wing Energy

S. No.	Algorithm	RSS VALUE
1.	AR (Auto Regression)	14.6849
2.	MA (Moving Average)	14.0690
3.	ARIMA (Autoregressive Integrated Moving Average)	13.1328

Table 6. RSS Value of time series model for wind energy

Similarly, here also the RSS value of ARIMA is lower than the RSS value of both AR and MA therefore in this case also ARIMA will perform better for forecasting the wind generation as compared to AR and MA.

Linear Regression Model

Now that we have prepared our model for the ARIMA model, let's see how linear regression is working on this problem. For linear regression inputs and outputs, we selected our input as the "Solar energy generation (MU)" and output as Average of "Solar energy generation (MU) after 365 days". Since, we have only two variables in this problem therefore, this is a simple linear regression problem. The equation for simple linear regression is -

= +

is the predicted value of value

is the slope of line that need to be optimized

is intercept of line that need to be optimized

I. Solar Energy

After fitting the linear regression model to our data, the obtained values of m and c for solar energy are 0.88 and 46.87 respectively. Therefore, the final equation of linear regression for solar energy generation will be-

= 0. 88 + 46. 87

Thus, from the above equation, as the value of is increasing, the value of is also increasing with a rate of 0.88

II. Wind Energy

Similarly, for wind energy generation, the obtained m and c values are 0.704 and 53.49 respectively.

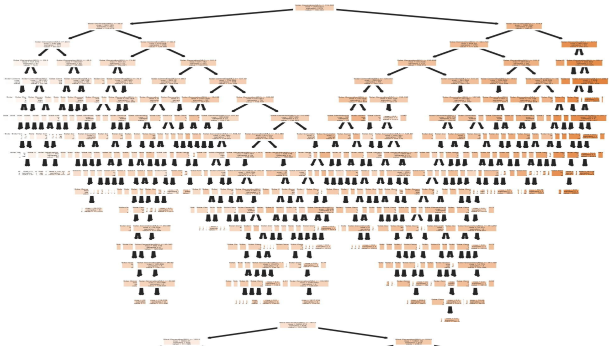
Therefore, equation of wind energy generation for linear regression will be -

= 0. 0. 704 + 53.49

Here also the value of is increasing as increases.

Random Forest Model

Next, we are using the Random Forest machine learning model. Random forest is a very popular and powerful machine learning algorithm which works on combination of multiple decision trees. For this study, we are using a combination of 100 decision trees for preparation of a random forest model. One such tree for both solar and wind energy is shown below in **Fig. 23** and **Fig. 24.**



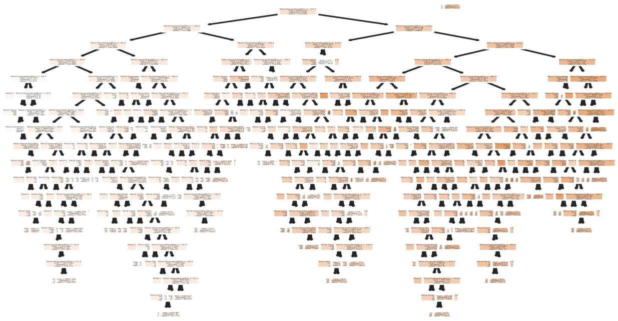


Fig 23. Decision tree of Random Forest for first estimator of solar energy

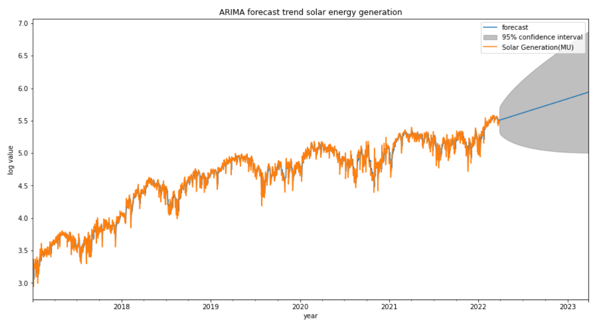
Fig 24. Decision tree of Random Forest for first estimator of wind energy

IV. RESULTS AND DISCUSSION

Visualizations

Now that we have prepared all the algorithms, we are now predicting or forecasting the result of energy generation for 1 year ahead in future. Collected data last date is 2022-03-27 and 1 year prediction i.e., on 2023-03-27, the graph for this prediction is shown below using each algorithm for both the energies generation.

1. ARIMA



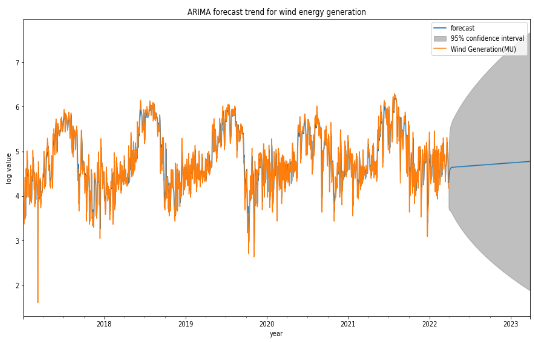
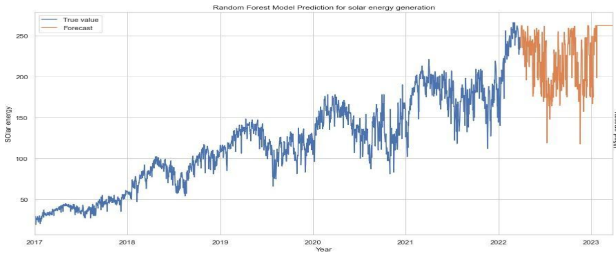


Fig 25. ARIMA forecast trend of solar energy generation generation

Fig 25. ARIMA forecast trend of wind energy

Above figures show the confidence interval of predicted values i.e, the forecasting results will range in shaded regions with the probability of 95%.

2. Random Forest



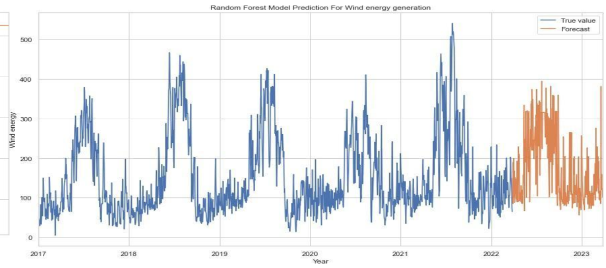
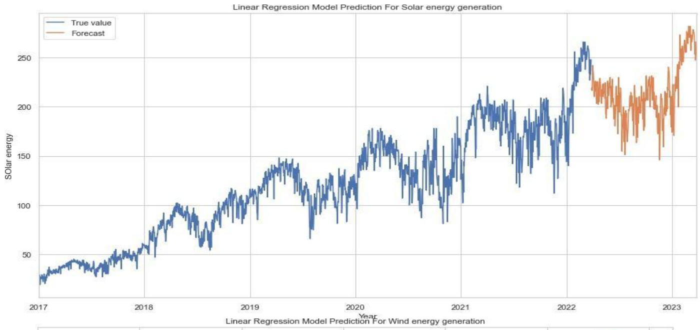


Fig 27. Random Forest prediction for solar energy generation generation

Fig 28. Random Forest prediction for wind energy

3. Linear Regression



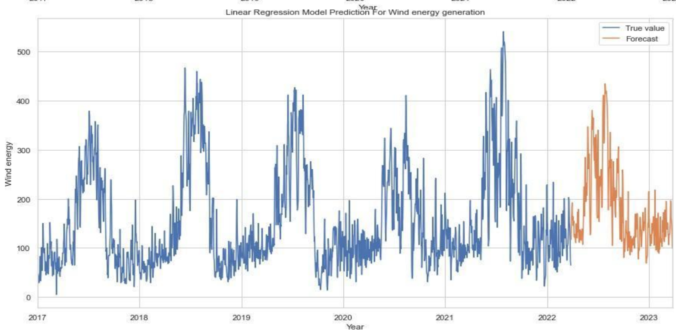


Fig 29. Linear Regression model prediction for solar energy generation **Fig 30.** Linear Regression model prediction for wind generation

Prediction Results

The value of solar and wind energy generation according to ARIMA, Random Forest and Linear regression on 202-03-27 is shown in below table (**Table 7**).

S. No.	Date	Energy Type	Value	Mean	Count
		ARIMA			
1.	2023-23-27	Solar Energy	379.02	30..98	38.37
2.	2023-23-27	Wind Energy	118.52	110.41	5.45
		RANDOM FOREST			
3.	2023-23-27	Solar Energy	262.14	225.44	33.26
4.	2023-23-27	Wind Energy	97.87	175.26	86.69
		LINEAR REGRESSION			
5.	2023-23-27	Solar Energy	266.32	215.14	28.03
6.	2023-23-27	Wind Energy	99.29	180.31	81.11

Table 7. Predicted values, Mean and Std for solar and wind energy using ARIMA, RF and LR

From the above table, we have compared predicted result of each model for both the energies on 2023-03-27. But the values of each model are varying therefore, we have to find the best model for the prediction of solar and wind energy generations. For that we are going to compare all these algorithms by using some useful metrics like MAE, MSE, RMSE, MAPE.

S. No.	Date	Energy Type	MAE	MSE	RMSE	MAPE
			ARIMA			
1.	2023-23-27	Solar Energy	0.06	0.01	0.08	32.07
2.	2023-23-27	Wind Energy	0.20	0.07	0.27	21.42
			RANDOM FOREST			
3.	2023-23-27	Solar Energy	15.65	439.95	20.97	0.12
4.	2023-23-27	Wind Energy	61.73	7093.15	84.22	0.40
			LINEAR REGRESSION			
5.	2023-23-27	Solar Energy	15.78	407.91	20.20	0.12
6.	2023-23-27	Wind Energy	54.65	5911.85	76.89	0.35

Table 8. Evaluating the performance of ARIMA, Random Forest and Linear Regression using metrics like MAE, MSE, RMSE and MAPE

V. CONCLUSIONS AND FUTURE WORK

In this study, we have proposed various machine learning and time series algorithms in order to predict the Solar and Wind generations which can help to ensure the adequate resource size of energy storages in microgrids. Therefore, our main objective for this study was to forecast or predict the renewable energy generations for 1 year ahead in future, and for that we have used ARIMA (Autoregressive integrated moving average) model, Linear regression and Random Forest. Then, we have compared these algorithms using certain performance metrics like MAE, MSE, RMSE, MAPE (see Table 8). The Lower the values of these matrices will be, the better will be the performance of the algorithm. From Table 8. The MAE value for the ARIMA (0.06 and 0.20) model for solar and wind energy is very less as compared to Random Forest (15.65 and 61.73) and Linear Regression (15.78 and 54.65) of solar and wind energy. Same with MSE and RMSE, the MSE and RMSE value for the ARIMA of solar energy model obtained is 0.01 and0.08 and wind energy is 0.07 and 0.27 respectively, which is very low as compared to the MSE, RMSE values for Random Forest and Linear Regression for both solar and wind energy. But the MAPE value for ARIMA (32.07 and 21.42) is relatively higher than the MAPE value for Random Forest (0.12 and 0.40) and Linear regression (0.12 and 0.35). After comparing all of these matrices of each algorithm for both the dataset, we concluded that the ARIMA model is best fit for the forecasting of renewable energy .In future, for the prediction or forecasting of solar and wind, we can use other machine learning time series models like VAR (Vector Autoregression), SARIMAX, (Seasonal Autoregressive Integrated Moving-Average with Exogenous Regressors) and Ensemble learning like gradient boosting, voting classifier, XGBoost classifier, ADA boost classifier in order to get the minimum values for MAE,MSE,RMSE i.e, to achieve better performance of the model.

REFERENCES

1. Ahmad, T., H. Zhang, and B. Yan. 2020a. “A review on renewable energy and electricity requirement forecasting modelfor smart grid and buildings.” *Sustainable Cities and Society*. Elsevier Ltd.

2. AlKandari, M., and I. Ahmad. 2019. “Solar power generation forecasting using ensemble approach based on deep learning and statistical methods.” *Applied Computing and Informatics*. Elsevier B.V.

3. Bakhtiari, H., J. Zhong, and M. Alvarez. 2021. “Predicting the stochastic behavior of uncertainty sources in planning a stand-alone renewable energy-based microgrid using Metropolis–coupled Markov chain Monte Carlo simulation.” *Applied Energy*, 290. Elsevier Ltd. https://doi.org/10.1016/j.apenergy.2021.116719.

4. Cao, Y., A. Raise, A. Mohammadzadeh, S. Rathinasamy, S. S. Band, and A. Mosavi. 2021. “Deep learned recurrent type-3 fuzzy system: Application for renewable energy modeling/prediction.” *Energy Reports*, 7: 8115–8127. Elsevier Ltd. https://doi.org/10.1016/j.egyr.2021.07.004.

5. Chen, Y., T. Li, C. Zhao, and W. Wei. 2021. “Decentralized Provision of Renewable Predictions within a Virtual Power Plant.” *IEEE Transactions on Power Systems*, 36 (3): 2652–2662. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/TPWRS.2020.3035174.

6. Corizzo, R., M. Ceci, H. Fanaee-T, and J. Gama. 2021. “multi-aspect renewable energy forecasting.” *Information Sciences*, 546: 701–722. Elsevier Inc. https://doi.org/10.1016/j.ins.2020.08.003.

7. Faraji, J., A. Abazari, M. Babaei, S. M. Muyeen, and M. Benbouzid. 2020. “Day-ahead optimization of prosumer considering battery depreciation and weather prediction for renewable energy sources.” *Applied Sciences (Switzerland)*, 10 (8). MDPI AG. https://doi.org/10.3390/APP10082774.

8. Fathollahzadeh, M. H., A. Speake, P. C. Tabares-Velasco, Z. Khademian, and L. L. Fight. 2021. “Renewable energy analysis in indigenous communities using bottom-up demand prediction.” *Sustainable Cities and Society*, 71. Elsevier Ltd. https://doi.org/10.1016/j.scs.2021.102932.

9. Gaigalis, V., and V. Katinas. 2020. “Analysis of the renewable energy implementation and prediction prospects in compliance with the EU policy: A case of Lithuania.” *Renewable Energy*, 151: 1016–1027. Elsevier Ltd. https://doi.org/10.1016/j.renene.2019.11.091.

10. Hosein, G., P. Hosein, S. Bahadoorsingh, R. Martinez, and C. Sharma. 2020. “Predicting renewable energy investment using machine learning.” *Energies (Basel)*, 13 (17). MDPI AG. https://doi.org/10.3390/en13174494.

11. Ikram, M. 2021. “Models for Predicting Non-Renewable Energy Competing with Renewable Source for Sustainable Energy Development: Case of Asia and Oceania Region.” *Global Journal of Flexible Systems Management*, 22: 133–160. Springer. https://doi.org/10.1007/s40171-021-00285-7.

12. Jana, R. K., I. Ghosh, and M. K. Sanyal. 2020. “A granular deep learning approach for predicting energy consumption.” *Applied Soft Computing Journal*, 89. Elsevier Ltd. https://doi.org/10.1016/j.asoc.2020.106091.

13. Jia, P., H. Zhang, X. Liu, and X. Gong. 2021. “Short-Term Photovoltaic Power Forecasting Based on VMD and ISSA-GRU.” *IEEE Access*, 9: 105939–105950. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2021.3099169.

14. Khan, P. W., Y. C. Byun, S. J. Lee, D. H. Kang, J. Y. Kang, and H. S. Park. 2020. “Machine learning-based approach to predict energy consumption of renewable and nonrenewable power sources.” *Energies (Basel)*, 13 (18). MDPI AG. https://doi.org/10.3390/en13184870.

15. Kim, S. G., J. Y. Jung, and M. K. Sim. 2019. “A two-step approach to solar power generation prediction based on weather data using machine learning.” *Sustainability (Switzerland)*, 11 (5). MDPI. https://doi.org/10.3390/SU11051501.

16. Lai, J. P., Y. M. Chang, C. H. Chen, and P. F. Pai. 2020. “A survey of machine learning models in renewable energy predictions.” *Applied Sciences (Switzerland)*. MDPI AG.

17. Lee, J., W. Wang, F. Harrou, and Y. Sun. 2020. “Wind Power Prediction Using Ensemble Learning-Based Models.” *IEEE Access*, 8: 61517–61527. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2020.2983234.

18. Li, L. L., S. Y. Wen, M. L. Tseng, and C. S. Wang. 2019. “Renewable energy prediction: A novel short-term prediction model of photovoltaic output power.” *Journal of Cleaner Production*, 228: 359–375. Elsevier Ltd. https://doi.org/10.1016/j.jclepro.2019.04.331.

19. Liu, C., W. Z. Wu, W. Xie, and J. Zhang. 2020. “Application of a novel fractional grey prediction model with time power term to predict the electricity consumption of India and China.” *Chaos, Solitons and Fractals*, 141. Elsevier Ltd. https://doi.org/10.1016/j.chaos.2020.110429.

19. Long, Y., and R. Zhang. 2020. “Short-term Wind Speed Prediction with Ensemble Algorithm.” *Proceedings - 2020 Chinese Automation Congress, CAC 2020*, 6192–6196. Institute of Electrical and Electronics Engineers Inc.

20. Makade, R. G., S. Chakrabarti, and B. Jamil. 2019. “Prediction of global solar radiation using a single empirical model for diversified locations across India.” *Urban Climate*, 29. Elsevier B.V. https://doi.org/10.1016/j.uclim.2019.100492.

21. Mohy-Ud-Din, G., K. M. Muttaqi, and D. Sutanto. 2020. “Adaptive and Predictive Energy Management Strategy for Online Optimal Power Dispatch from VPPs with Renewable Energy and Energy Storage.” *2020 IEEE Industry Applications Society Annual Meeting, IAS 2020*. Institute of Electrical and Electronics Engineers Inc.

22. Mosavi, A., M. Salimi, S. F. Ardabili, T. Rabczuk, S. Shamshirband, and A. R. Varkonyi-Koczy. 2019a. “State of the art of machine learning models in energy systems, a systematic review.” *Energies (Basel)*. MDPI AG.

23. Nam, K. J., S. Hwangbo, and C. K. Yoo. 2020. “A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea.” *Renewable and Sustainable Energy Reviews*, 122. Elsevier Ltd. https://doi.org/10.1016/j.rser.2020.109725.

24. Qadir, Z., S. I. Khan, E. Khalaji, H. S. Munawar, F. Al-Turjman, M. A. P. Mahmud, A. Z. Kouzani, and K. Le. 2021. “Predicting the energy output of hybrid PV–wind renewable energy system using feature selection technique for smart grids.” *Energy Reports*, 7: 8465–8475. Elsevier Ltd. https://doi.org/10.1016/j.egyr.2021.01.018.

25. Qiao, L., S. Chen, J. Bo, S. Liu, G. Ma, H. Wang, and J. Yang. 2019. “Wind power generation forecasting and data quality improvement based on big data with multiple temporal-spatual scale.” *Proceedings - IEEE International Conference on Energy Internet, ICEI 2019*, 554–559. Institute of Electrical and Electronics Engineers Inc.

26. Rahman, M. M., M. Shakeri, S. K. Tiong, F. Khatun, N. Amin, J. Pasupuleti, and M. K. Hasan. 2021. “Prospective methodologies in hybrid renewable energy systems for energy prediction using artificial neural networks.” *Sustainability (Switzerland)*. MDPI AG.

27. Rahul, A. Gupta, A. Bansal, and K. Roy. 2021. “Solar energy prediction using decision tree regressor.” *Proceedings - 5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021*, 489–495. Institute of Electrical and Electronics Engineers Inc.

28. Somu, N., G. Raman M R, and K. Ramamritham. 2021. “A deep learning framework for building energy consumption forecast.” *Renewable and Sustainable Energy Reviews*, 137. Elsevier Ltd. https://doi.org/10.1016/j.rser.2020.110591.

29. Thukral, M. K. 2020. “Solar power output prediction using multilayered feedforward neural network: A case study of Jaipur.” *Proceedings - 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security, iSSSC 2020*. Institute of Electrical and Electronics Engineers Inc.

30. Ullah, I., R. Ahmad, and D. H. Kim. 2018. “A prediction mechanism of energy consumption in residential buildings using hidden markov model.” *Energies (Basel)*, 11 (2). MDPI AG. https://doi.org/10.3390/en11020358.

31. Wang, H., Z. Lei, X. Zhang, B. Zhou, and J. Peng. 2019. “A review of deep learning for renewable energy forecasting.” *Energy Conversion and Management*. Elsevier Ltd.

32. Wang, J., S. Chung, A. AlShelahi, R. Kontar, E. Byon, and R. Saigal. 2021. “Look-ahead decision making for renewable energy: A dynamic ‘predict and store’ approach.” *Applied Energy*, 296. Elsevier Ltd. https://doi.org/10.1016/j.apenergy.2021.117068.

33. Wang, Z., and L. Jia. 2020. “Short-Term Photovoltaic Power Generation Prediction Based on LightGBM-LSTM Model.” *2020 5th International Conference on Power and Renewable Energy, ICPRE 2020*, 543–547. Institute of Electrical and Electronics Engineers Inc.

34. Xia, M., H. Shao, X. Ma, and C. W. de Silva. 2021. “A Stacked GRU-RNN-Based Approach for Predicting Renewable Energy and Electricity Load for Smart Grid Operation.” *IEEE Transactions on Industrial Informatics*, 17 (10): 7050–7059. IEEE Computer Society. https://doi.org/10.1109/TII.2021.3056867.

35. Yu, Q., Z. Jiang, and M. Zhao. 2020. “Research of photovoltaic prediction compensation control strategy of hybrid renewable energy microgrid.” *International Conference on Signal Processing Proceedings, ICSP*, 658–663. Institute of Electrical and Electronics Engineers Inc.