Wind Energy Predictor for Power Plant

J. I. Chinthaka
Department of Computer Science and Engineering
University of Moratuwa
imesh.20@cse.mrt.ac.lk

M.P.H.P. Molligoda
Department of Computer Science and Engineering
University of Moratuwa
heshanm.20@cse.mrt.ac.lk

W.A.I.V. M. De Silva
Department of Computer Science and Engineering
University of Moratuwa
vimarshani.18@cse.mrt.ac.lk

Abstract - Wind power is one of the fast-growing renewable energy technologies in the world due to low cost and restriction of Co_2 emission but it also presents inherent challenges in some regions of the world. The data that has been used in this paper is obtained from Sri Lanka Sustainable Energy Authority based on the Pooneryn wind turbine site. This article studies more than 20 years of data to provide various important analyses on feasibility of building a wind turbine in Pooneryn. The paper also predicts and forecasts the wind and energy to be produced by the wind turbine in the future.

I. Introduction

Environmental concerns and supply uncertainties are driving many countries to rethink their energy mix and develop diverse sources of clean, renewable energy. Cost-effective energy that can be replaced without major negative environmental impacts has become the goal worldwide.

Wind power as a clean, renewable resource has become a global hot spot of government investment with its advanced technology within the last decade. Globally on-shore wind power has seen considerable growth in all grid systems. In the coming decade off-shore wind power is also expected to expand rapidly. Wind power is variable and intermittent over various time scales because it is weather dependent. Therefore wind power integration into traditional grids needs additional power systems and electricity market planning and management for system balancing. This extra system balancing means that there are additional system costs associated with wind power assimilation. Wind power forecasting and prediction methods are used by system operators to plan unit commitment, scheduling and dispatch and by electricity traders and wind farm owners to maximize profit. Accurate wind power forecasting and prediction has numerous challenges.

Numerical Weather Predictions (NWP) [1] are the basic input of wind power forecasting systems, which are only for several nodes of the grid, in the vicinity of wind farms, covering a certain area. To forecast wind power output at the wind farms, down scaling of NWP is usually conducted first to get wind speed forecasts, and then converted to wind power with specific power curves. Regional wind power forecasting is obtained by upscaling the wind farms' wind power output. Wind power generation is proportional to the cube of input wind speed, so accurate wind speed forecasting at grid points near wind farms is critical for advanced wind power forecasting systems.

In this paper, we analyse wind data of Pooneryn. The Pooneryn site is situated west of the village of Pooneryn in the northern tip of the Island. The site offers exceptional wind and good solar resources both attractive for renewable energy generation. The site also exhibits distinct environmental values, mainly a rich dune landscape and several seasonal water bodies that interact with the northern bordering lagoon to host a varied bird species population.

In section II we describe the Dataset what we used. Methodology of the Wind Energy Predictor for Power Plant is explained in section III under three subtopics; Descriptive Analytics, Diagnostic Analytics and Predictive Analytics.

II. Overview of the Data Set Used

The Ministry of Power and Renewable Energy of the Government of Sri Lanka [7] is responsible for collecting wind related data to analyze the feasibility of building a new Wind turbine in Pooneryn, Sri Lanka. These datasets are machine readable and well structured. Data and information used for all the analytics described in this paper belongs to the Government of Sri Lanka.

The data set that is used, contains mainly two types of data. First set is Merra Data, which contains wind and weather data collected by satellites for the past 20 years in the Pooneryn area. Second set is the data collected from sensors placed at Pooneryn at different heights to capture the wind related data during 2015 to 2107. We used data of both types to perform different types of analytics depending on the needs.

III. Methodology

A. Descriptive Analytics

The case study in this paper is the wind energy plant of Poonerin, which has been studied from 1997 to 2017 in each 1 hour period, the average temperature of the site is 27.96 ^{0}C through the period, and the wind speed is measured every 1 hour at the height of 50 meter.

The data that has been used in this paper is obtained from Sri Lanka Sustainable Energy Authority and applied in the site of Poonerin.

A wind power plant is set up to generate electricity for about 25 years. Therefore, a minimum of 20 years of data analysis is required to establish a wind power plant.

The average wind speed for each month was calculated and shown in a graph to see the variations of wind speed throughout 21 years.

Figure 1. The average wind speed for 21 years

Month / Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1997	7.284	6.342	5.772	4.457	8.714	9.493	10.069	9.398	9.364	4.885	5.825	7.036
1998	7.703	6.579	5.663	5.439	8.609	11.357	8.783	8.614	9.351	8.066	5.605	8.442
1999	8.220	7.577	4.504	8.779	9.978	10.493	10.372	9.242	8.835	7.129	7.792	8.295
2000	8.495	6.739	5.500	7.223	8.372	10.485	9.060	10.509	7.870	7.428	8.298	9.321
2001	8.319	5.382	5.975	4.346	8.882	9.905	9.265	9.200	9.048	7.060	6.738	8.377
2002	7.667	7.819	5.664	4.180	9.460	9.390	9.477	9.798	9.378	6.388	6.023	8.469
2003	8.467	6.964	5.506	5.077	9.586	10.096	8.745	8.507	8.265	8.596	7.924	8.279
2004	7.887	7.265	5.612	5.341	9.508	10.425	9.045	9.115	8.077	6.957	8.343	8.776
2005	7.338	6.955	5.452	5.571	8.044	10.704	10.660	8.813	9.594	6.831	7.372	7.455
2006	7.652	7.014	4.861	6.363	9.882	9.220	9.895	10.422	9.237	7.330	6.217	8.788
2007	8.256	6.877	5.600	4.627	9.076	9.695	9.717	9.554	9.590	7.484	6.559	8.621
2008	7.292	6.578	6.597	6.149	10.059	10.234	9.456	9.006	9.086	6.396	8.044	7.742
2009	7.562	6.172	5.153	6.418	9.035	11.345	10.264	8.851	9.796	8.426	6.534	7.887
2010	7.542	6.400	5.551	5.353	9.477	10.177	9.325	8.290	7.629	8.084	5.535	8.081
2011	8.585	7.487	6.162	5.267	9.054	10.171	9.539	9.589	9.106	5.996	7.311	7.919
2012	6.852	7.079	5.306	6.137	10.327	10.916	9.435	10.159	9.198	7.617	6.858	8.419
2013	7.929	7.004	5.897	5.613	9.380	10.574	10.085	8.107	9.941	7.831	6.895	8.029
2014	8.689	6.994	6.663	4.438	8.499	9.588	9.623	9.277	8.406	6.531	7.694	8.391
2015	7.045	7.446	5.658	4.379	8.207	9.874	9.421	9.125	9.107	6.740	6.423	7.521
2016	7.094	6.540	5.459	5.144	8.964	9.938	8.754	8.242	8.132	6.446	6.815	7.460
2017	7.992	7.162	5.043	6.645	8.545	10.443	8.784	8.569	8.788	7.737	7.354	7.435

By considering the on site data from 2015 to 2016 following graph fig 4. Obtained.

Southwest season (May - September)
Northeast season (December - March)

During the Northeast season wind blows from the Northeast direction. Northeast

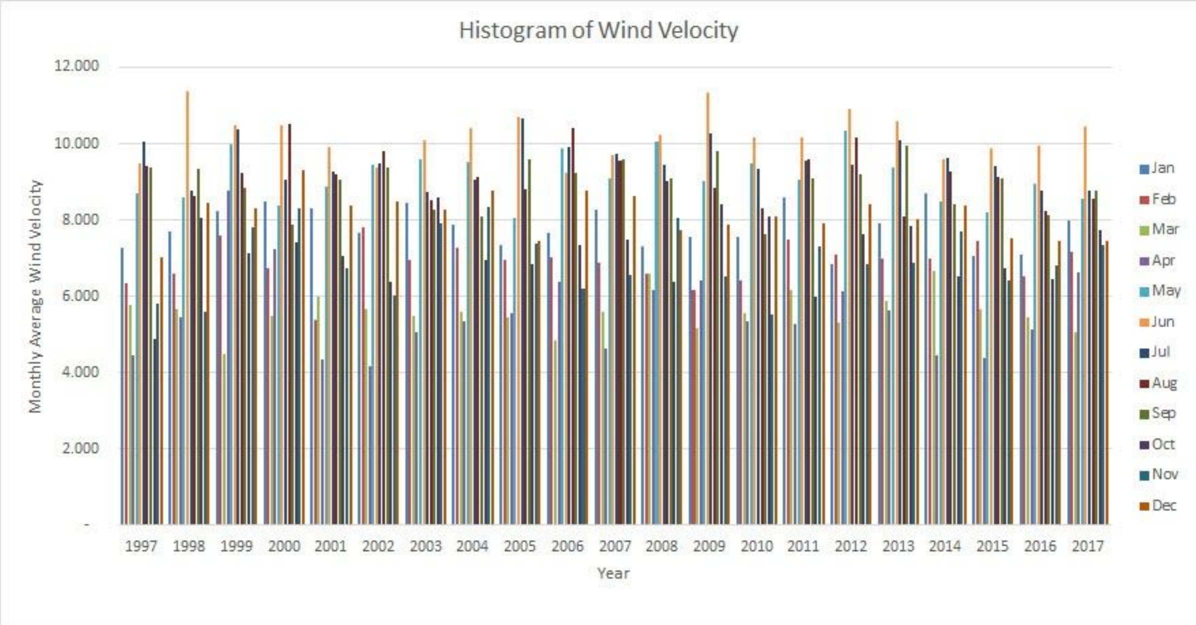
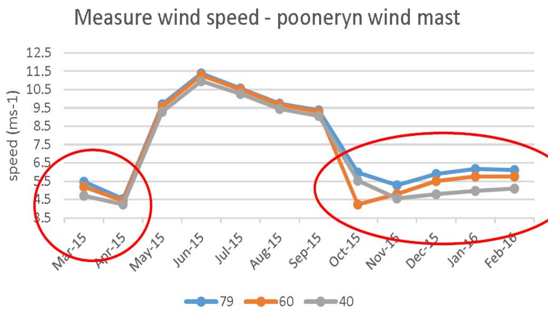


Figure 2. Histogram of Wind Velocity

Figure 3. Wind speed -pooneryn wind mast



side of the Pooneryn site is covered with land and the Southwest direction has a water surface. Thus wind speed during the Northeast season is lower than the Southwest season. In addition to that fig4. Shows that wind speed increases with the height from the ground level.

Taking one year is not accurate for the data analysis as if there is high wind throughout the year, analysis will be over estimated and if there is low wind throughout the particular year it will be an under estimation. To avoid such circumstances long term correction is done. The data consists of 184,079 Samples. That included timestamp, wind speed at the height of 50m, wind direction and temperature. We calculated the average annual wind speed, annual power and annual energy for each year. According to the calculations total energy of 190,317.81MWh(190,317,810.82 kWh) could be generated for 21 years from 1997 to 2017 at the site of Pooneryn.

Figure 4. Pooneryne

Twenty one years mean annual wind speed variations at 50m level and produced power and energy calculated are tabulated below.

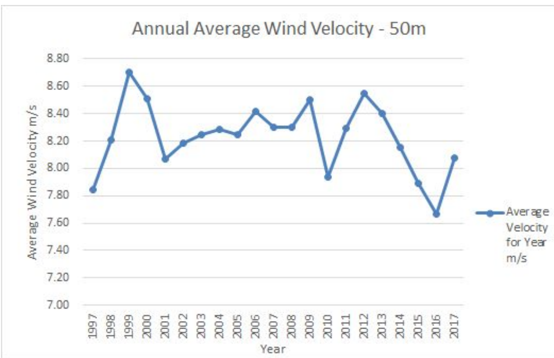
Estimation of wind power can be obtained from the formula;

$$P = \frac{1}{2}\rho * A * v^{3} * A \ [W]$$

ρ : density of air		1.23	kg/m^3	
Rotor diameter		112	m	
A : area swept by the blades		9,852.03	m²	
Betz limit		59%		
Yield losses :		51.556500%		
Blades		67%	between 0,2 and 0,85	
Gear box		90%	between 0,7 and 0,98	
Generator		90%	between 0,8 and 0,98	
Transformer :		95%	between 0,85 and 0,98	

Year	Average Velocity for Year m/s	Annual Power / kW	Annual Energy /MWh
1997	7.85	893.94	7,830.94
1998	8.21	1,023.82	8,968.69
1999	8.71	1,221.56	10,700.87
2000	8.51	1,141.60	10,027.83
2001	8.07	971.13	8,507.10
2002	8.18	1,014.40	8,886.10
2003	8.24	1,037.22	9,086.07
2004	8.28	1,052.27	9,243.14
2005	8.25	1,038.76	9,099.58
2006	8.41	1,102.99	9,662.21
2007	8.30	1,058.65	9,273.79
2008	8.30	1,058.71	9,299.71
2009	8.50	1,136.87	9,958.99
2010	7.94	925.54	8,107.74
2011	8.29	1,056.26	9,252.81
2012	8.55	1,156.93	10,162.47
2013	8.40	1,097.16	9,611.12
2014	8.16	1,004.20	8,796.75
2015	7.89	909.38	7,966.17
2016	7.67	834.69	7,331.92
2017	8.08	975.32	8,543.82
Total Energy for 21 years			**190,317.81**

Figure 5. Annual Average Wind Velocity



The variations of annual mean wind speeds for 21 years are depicted in the fig.3.

Figure 6. Annual General Power

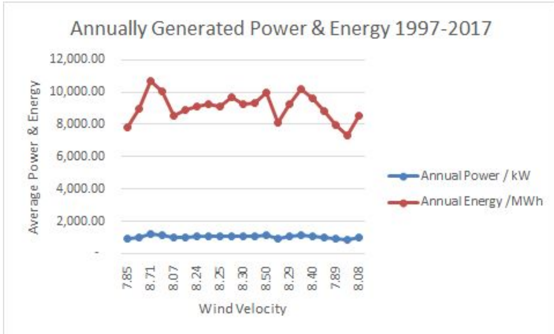
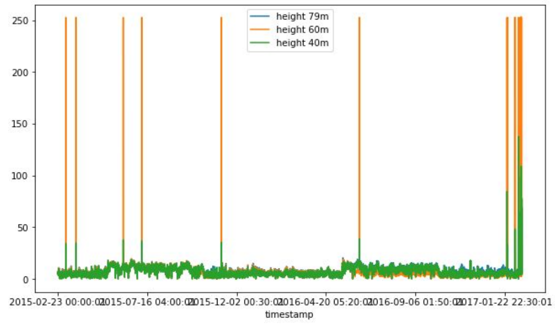


Figure 7. Average wind speeds against height



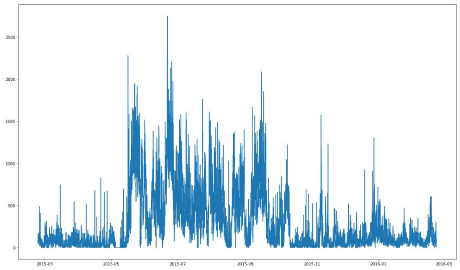
B. Diagnostic Analytics

According to IBM analytics [6], Diagnostic Analytic is finding reasons as to why it happened. Diagnostic Analytics can be broadly categorized into 3 sections such as identify anomalies, discover anomalies, determine hidden relationships. Sometimes analysts have to look into patterns outside existing data sets and also analyze data from external sources to identify anomalies. Identifying anomalies can lead to discovery of Hidden casual relationships as well.

When we draw the calculated energy based on the wind speeds collected through the sensors placed at Pooneryn agins the dates, it is clear that more energy was created during May to September than the other months.

During May and September Sri Lanka has the South-West monsoon season which brings a lot of wind into the country[8]. So one of the reasons behind the increment of energy during that period is caused by monsoon winds.

Figure 8. Energy generated on each day



C. Predictive Analytics

In order to provide predictive analytics for the data set first approach we took was time series analysis.

A time series is a sequence of data points, measured typically at successive times spaced at uniform time intervals.

First, we applied an ARMA model to the data set using 80% of data as training data. Then we used the ARIMA model to predict 10000 steps with the parameters (0,1,2) for (p,d,q). Both of the models were unable to provide a good result, when compared with the test data set.

Figure 9. ARMA results

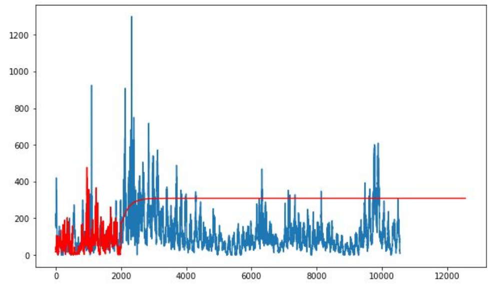
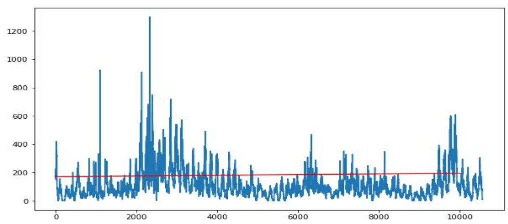


Figure 10. ARIMA results



In conclusion, time series forecasting methods were unable to provide better predictions because mostly they are not applicable in non-linear models.

Then, according to our Descriptive analytics, we predict Wind Speed because wind speed is proportional to the wind power. Then we predict the Wind Direction. With this predicted wind direction, we can excel the generated wind power by changing the direction of the wind turbine.

For predictions, we tried with both Machine Learning Models and Deep Learning Models.

For Both of these predictions we use MERRA dataset. First we rearranged the MERRA [2] dataset manually. Then divide the data set to three data frames; Train Data, Validation Data and Test Data, using python to process Data..

First we tried Machine Learning Models, Linear Regression and Random ForestRegressor using the Scikit-Learn Python library.

Mark Holmstrom et al [3] have concluded that, both linear regression and functional regression were outperformed by professional weather forecasting services, although the discrepancy in their performance decreased significantly for later days, indicating that over longer periods of time. Linear regression proved to be a low bias, high variance model whereas functional regression proved to be a high bias, low variance model. Linear regression is inherently a high variance model as it is unstable to outliers, so one way to improve the linear regression model is by collection of more data. Functional regression, however, was high bias, indicating that the choice of model was poor, and that its predictions cannot be improved by further collection of data. This bias could be due to the design choice to forecast weather based upon the weather of the past two days, which may be too short to capture trends in weather that functional regression requires.

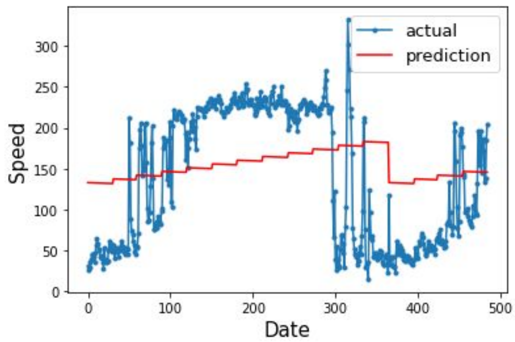
In most web resources [4] [5], Neural Networks with data processing is recommended for weather data prediction. Neural networks can take different features as input variables to find nonlinear relationships between input and output.

We used LSTM Recurrent Neural Networks in Python with Keras library. And we had the best results for prediction because our MERRA data set is big enough to fit the Neural Network. All results are in the Result section.

IV. Results

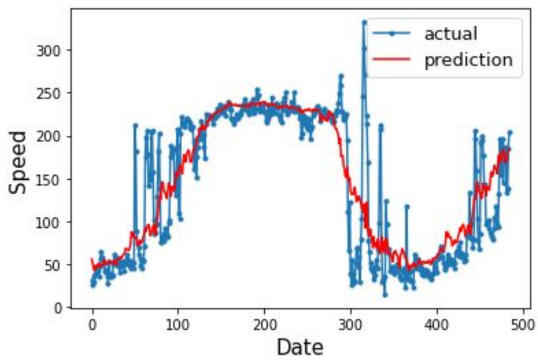
A. Wind Velocity Prediction

Figure 11. The wind velocity prediction with LinearRegression



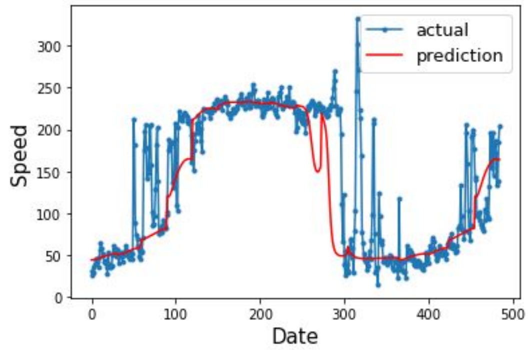
Confidence = 0.07243916265021422

Figure 12. The wind velocity prediction with RandomForestRegressor



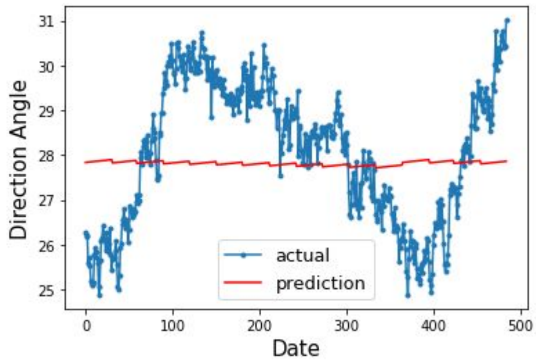
Confidence = 0.7739204029335223

Figure 13 The wind velocity prediction with LSTM RNN



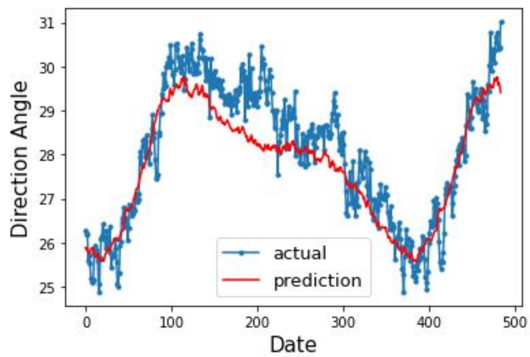
B. Wind Direction Prediction

Figure 14. The wind Direction prediction with LinearRegression



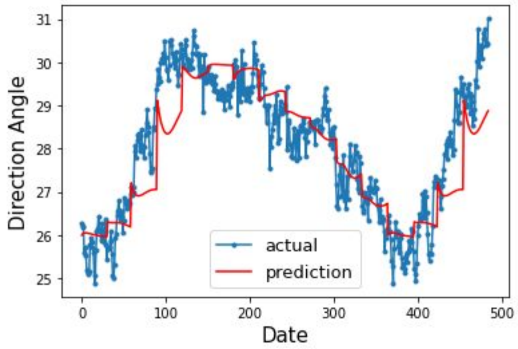
Confidence = -0.034839730783395195

Figure 15. The wind Direction prediction with RandomForestRegressor



Confidence = 0.8083416377008348

Figure 16. The wind Direction prediction with LSTM RNN



V. Conclusion

When comparing the results, we can conclude our predictive analysis; LSTM RNN is given a better prediction than the Regression models. Both linear regression and functional regression were outperformed by professional weather forecasting services, although the discrepancy in their performance decreased significantly for later days, indicating that over longer periods of time.

With our all analysis, in the future, generated wind power can be exceeded by predicting to change the wind turbine direction and height according to the predicted data. And can manage the power storages according to the predicted power using wind speed predictions.

References

[1] https://en.wikipedia.org/wiki/Numerical_weather_prediction

[2] https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/

[3] Mark Holmstrom, Dylan Liu, Christopher Vo, "Machine Learning Applied to Weather Forecasting", 2016

[4] https://www.researchgate.net/post/Which_machine_learning_algorithm_are_suitable_for_weather_prediction

[5] https://www.kaggle.com/questions-and-answers/27537

[6]https://www.ibm.com/blogs/business-analytics/diagnostic-analytics-101-why-did-it-happen/

[7]http://www.energy.gov.lk/en/renewable-energy/new-renewable-energy

[8]http://www.meteo.gov.lk/index.php?option=com_content&view=article&id=94&Itemid=310&lang=en#2-southwest-monsoon-season-may-september