Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 54 (2014) 211 – 220

Energy Procedia

4th International Conference on Advances in Energy Research 2013, ICAER 2013

# Quantitative Performance Evaluation of a Wind Turbine Generator Cluster using Statistical Techniques

S. V. Joshi[a,*], Salil M. Sainis[a], M. D'Souza[b]

[a]School of Mechanical and Building Sciences, VIT University, Vellore, India
[b]Suzlon Energy Limited, One Earth, Pune, India

**Abstract**

This report dissects the analysis of annual performance of a cluster of wind turbine generators operated by Suzlon Energy Limited to troubleshoot shortfalls in the predicted generation of individual machines for the fiscal year of 2011-12. The first phase involved an estimation of the annual energy production using wind resource assessment on the commercial software Wind Atlas, Analysis and Application Program (WAsP). Comparison of the Annual Energy Production (AEP) with direct extrapolation of monthly generation to centum machine and grid availability highlighted negative inconsistency between the two in the case of five machines. Detailed study of their data histories indicated the reason to be coarse extrapolation and generation curtailment, although the latter is a minor contributor. The second phase consists of quantification of these losses for the concerned machines using statistical extrapolation and a variety of data approaches. A qualitative comparison of the methods is presented based on accuracy and utility. It is concluded that correlation of average hub-height wind velocities with the concerned machine's generation yields the most reliable and professionally useful results. Also, the advantages and shortcomings of the other methods have been discussed.


Selection and peer-review under responsibility of Organizing Committee of ICAER 2013

*Keywords:* Wind power, Generation Loss Quantification, Machine/Grid Availability, Data Analysis, Curve Fitting

* Corresponding author. Tel.: +91-763-956-7408
*E-mail address:* saumitra.vjoshi@gmail.com

1876-6102 
Selection and peer-review under responsibility of Organizing Committee of ICAER 2013
doi:10.1016/j.egypro.2014.07.264

## 1. Introduction

Power generation from wind has emerged as a major sector in the energy market expanding at annual rates of 25 to 35% [1]. It has seen large-scale investment and government support catalyzing its exponential growth from a research field in the 1950s to a full-fledged industry in the present age. Wind power companies now compete on a global level for energy market shares, and, along with technology development, focus with equal commitment on building a strong customer base. A dominantly major source of customer satisfaction in the wind industry is the generation of the product: the wind turbine generators (WTGs). Their performance is directly linked with the monetary benefit or loss to the customer. The two important factors that play a vital role in the performance of the site are the wind at site, and the machine availability (MA), i.e. the period for which a WTG will be capable of generation given the source of power exists.

Out of the two factors, MA is what the WTG companies have control over. This is due to the fact that wind at site is a natural phenomenon and has a high degree of improbability and variance. The wind energy company must therefore strive to ensure high MA, failing which it may be liable to pay pre-determined penalty which results in loss of customer faith, or worse, loss of precious global market share. In cases where the former does happen, it becomes necessary to accurately quantify energy losses due to unavailability of machines as it determines the value of monetary remuneration to the customer.

This project studies the different techniques of calculating losses caused due to machine unavailability and grid unavailability, as well as other factors in order to determine the best suited method based on accuracy of results as well as practical utility. The core method relies on statistical correlation, with an array of data approaches for analysis parameters. In this study, sources of correlation are mast wind data at hub-height, and average nacelle wind data of the given cluster of turbines. The project also quantifies the values of various losses for concerned turbines.

**Nomenclature**

| | |
|---|---|
| $A$ | Rotor cross-sectional area ($m^2$) |
| $C_p$ | Betz limit (-) |
| $E$ | Energy in wind over time $t$ (J) |
| $G_{act}$ | Actual energy generation (kJ) |
| $G_{expected}$ | Expected energy generation (kJ) |
| $i, j$ | Incremental variables |
| $Loss_{curtailed}$ | Energy lost over given time interval due to curtailment (kJ) |
| $Loss_{gen}$ | Total energy generation loss (kJ) |
| $NetLoss$ | Total energy loss calculated for annual period (kJ) |
| $T00x\_Gen$ | Actual energy generation for WTG '$x$' (kJ) |
| $v$ | Wind speed ($ms^{-1}$) |
| $y$ | Actual value of variable |
| $y_f$ | Fitted value of variable |
| $y_m$ | Average value variable |
| $\eta$ | Extraction factor (-) |
| $\rho$ | Air density ($kgm^{-3}$) |

## 2. Site Description

The WTG cluster under scrutiny is located in the Kuchchh region of Gujarat in India and is operated and maintained by Suzlon Energy Limited. It consists of ten wind turbine generators of model S82 (Rated capacity: 1.5 MW) sited as shown is Figure 1. The mast shown in the figure is the source for the mast data. The proximity of the mast to the site is to be noted. The region of study is flat terrain, with a slope of 0.03 degrees, which equals a rise of

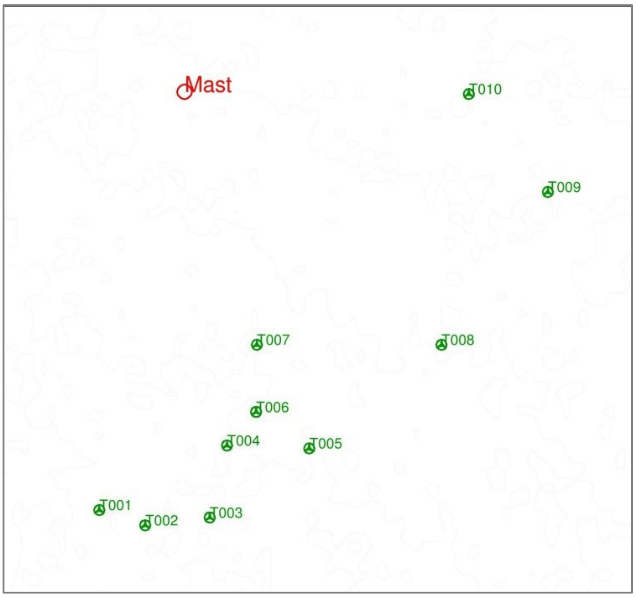


Fig. 1. Locations of cluster WTGs and mast

5 meters over a kilometer. This justifies the use of mast data in our analysis [2]. No significant obstacles to wind flow are found in the region, and the roughness of local terrain is considered.

## 3. Data Sources

The data used in this project is obtained from Suzlon Energy Group's Supervisory Control and Data Acquisition (SCADA) Monitoring Center (1-second data readings averaged over ten minutes) taken at the mast and each WTG for April 2011- March 2012. Mast data includes wind velocities and directions at 65m and 80m heights. WTG data includes the following parameters at hub height (80m): wind velocity, wind direction, power generation, pitch-angles, yaw angle (calculated indirectly), rotor RPM and generator RPM [3]. Mast data validity was checked and necessary corrections made. A few noteworthy points are:

- Mast data is missing from 17 Jan, 2012 (14:50) to 18 Jan, 2012 (00:00)
- Gaps in turbine data logs exist for periods of machine breakdown and grid unavailability.
- A few sets of data for T003 are missing in SCADA records (15248 to 19439). Hence the calculations for actual loss during this period have not been done.

### *3.1. Methods for data completion*

In order to obtain a complete data set, two methods are employed:

#### *3.1.1. Measure-Correlate-Predict (MCP) Method:*

MCP Method involves scattering wind velocity measured at the WTG's nacelle (henceforth 'nacelle velocity') against wind speed recorded at the hub height at mast. MATLAB is used to calculate the best fitting polynomial function relating nacelle velocity to mast velocity. Using this function, the missing velocities are calculated. This velocity set is used for calculation of power generation loss due to lack of both MA and Grid Availability (GA) as the mast data is available for almost the entire period of study. MCP is fairly valid in this study as the terrain is relatively undulating, and the distance between turbine and mast is within 10 km. MCP data is used in section 6.2.

Table 1. Comparison of WAsP Results and Actual Observed Power Generation

| WTG | Gross Output (lakh units) | Wake Losses (per cent) | $AEP_{G,M}$ (lakh units) | GA, MA Correction | Estimated Output (lakh units) | Actual Output (lakh units) |
|---|---|---|---|---|---|---|
| T001 | 35.39 | 1.53 | 34.85 | 86% | 30.0 | 29.2 |
| T002 | 35.62 | 2.65 | 34.68 | 84% | 29.3 | 24.1 |
| T003 | 35.36 | 3.27 | 34.2 | 86% | 29.5 | 26.2 |
| T004 | 34.88 | 6.63 | 32.56 | 86% | 28.1 | 29.7 |
| T005 | 34.53 | 5.84 | 32.52 | 82% | 26.7 | 22.7 |
| T006 | 34.51 | 8.99 | 31.41 | 87% | 27.3 | 28.1 |
| T007 | 34.02 | 4.23 | 32.58 | 85% | 27.6 | 29.3 |
| T008 | 34.24 | 3.93 | 32.89 | 87% | 28.7 | 25.8 |
| T009 | 33.63 | 3.72 | 32.38 | 87% | 28.3 | 26.9 |
| T010 | 34.00 | 1.44 | 33.51 | 87% | 29.2 | 27.3 |

### *3.1.2. Cluster Nacelle Velocity Average (CNVA):*

In this method, nacelle velocity data of all WTGs in the cluster are averaged at every time instant to give a data collection that is exhaustive as far as machine availability is concerned. Sets without a reading indicate periods of lack of GA, which are 324 sets (0.61% of 52704). Hence, this data set is used for calculation of generation loss due to lack of MA only. Such calculations are of more practical importance to a wind turbine generator company, since all its remuneration is calculated in terms of MA shortfall. This is used in calculations in section 6.3.

## 4. WAsP Results and Power Loss Identification

### *4.1. WAsP*

Using mast and terrain data, and the WTG locations, WAsP is used to calculate the gross annual energy output for each WTG. WAsP incorporates wake modeling in its code, and hence wake losses are included in the calculation of the net output ($AEP_{G,M}$) [4]. This is the predicted AEP that assumes centum machine and grid availability (Table 1). Based on the monthly averaged grid and machine availability for every WTG, correction factors are applied to the $AEP_{G,M}$ to yield the predicted output corrected to grid and machine availability conditions (henceforth 'estimated output'). A comparison of actual generation to the estimated output is shown in Table 1. It can be seen that machines T001, T002, T005, T008 and T010 show significant negative deviation from the estimated output and hence are selected for a detailed study. The next step is the determination of the reasons for underperformance of identified WTGs. Hence, a root cause analysis of shortcoming is carried out.

### *4.2. Root Cause Analysis*

Upon inquiry at the Contract Management department, it was found that T008 and T010 had undergone commissioning in the month of March 2011, and were in stabilization mode for two months of the period of this study. As a result, the data recorded by the controller for this period is not an eligible reference for actual performance of these machines, and hence, T008 and T010 are excluded from this study.

A look at the monthly MA data of the WTGs (Figure 2 (a): X-axis is in fiscal month order) indicates that WTGs T002, T003 and T005 all have relatively low values of MA during the high wind (HW) months of June, July, August and September. The theoretical limit of energy from wind of velocity v by a rotor of area *A* in time *t* is:

$$E = 0.5\rho C_p A v^3 \quad (1)$$

Hence, the loss of power due to lack of MA and GA for a time period t in the HW season will be much more than the loss for the same t period in the low wind (LW) season [5].

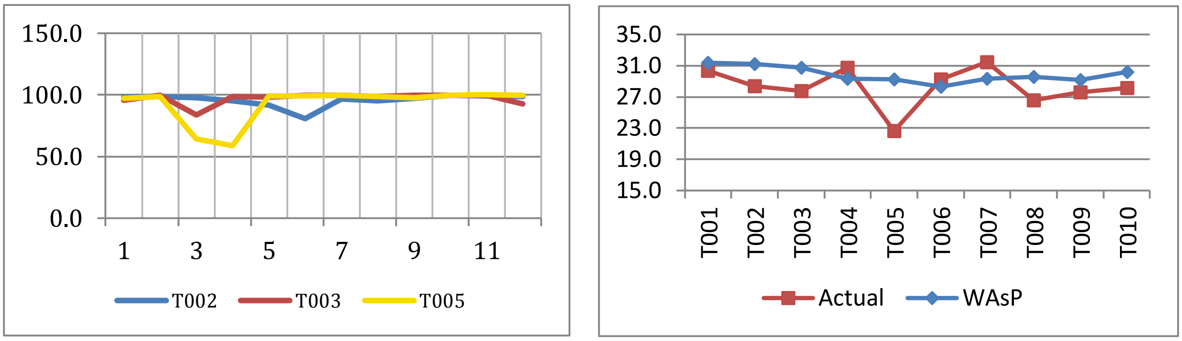


Fig. 2. (a) Machine Availability monthly variation for T002, T003 and T005, (b) WAsP-predicted and actual generation at centum MA and GA

During calculation of power generation for centum MA and GA for T002, T003 and T005, it is seen that the net power generation falls much below WAsP predictions (Figure 2 (b)). This can be explained by the fact that the extrapolated data is of the LW period, and hence is not a correct base for estimating power generation for HW months. The same inference applies on a diurnal scale. Hence, the difference in estimated and extrapolated values of wind turbine generator performance arises due to coarse extrapolation. There is an uncertainty in extrapolation of power generation data on a monthly as well as daily basis.

On plotting the pitch angles 1, 2 and 3 of the rotor blades of T002 versus the wind speed, a trend is noticed in the scatter. This trend indicates high pitching (greater pitch angles than standard) during higher wind speeds, indicating that the machine is being run at reduced power. This is known as power curtailment, and is carried out when (a) the grid cannot bear full power generation, and (b) when certain turbine component parameters (such as temperature) reach high values and are in danger of causing the turbine to trip. This factor is not included in the estimated wind power, and hence contributes to over-evaluated losses of power generation. Hence, root cause analysis identifies three major reasons for difference in estimated and actual power generation of the cluster:

- Turbine being in its stabilization period, due to which the data from the WTG is not eligible for reference during calculation of power generation losses.
- Coarse extrapolation of power generation data. Difference in generation during high wind and low wind is significant enough to cause a difference in actual and estimated generation.
- Power curtailment of turbines.

## 5. Statistical Correlation Technique

This technique of prediction is based on determining the weightage of points (xi, yi) on the argand plane to find the best-fitting polynomial through that set. The conditions for the application of this technique are: 1. A sufficiently large, period-exhaustive, finite and verified data set, and 2. The existence of a pattern in the data set. The SCADA data satisfies the two basic conditions required for statistical correlation in the sense that it is a large, fairly consistent, finite data set over the period of interest of this study. The basic steps in statistical correlation are:

- Identification of the independent variable $x$. The independent variable must be available in sufficiently large quantity. This, in our case, is the wind speed. Wind speed can be obtained from MCP as well as CNVA methods as described in section 3.
- Choice of the dependent parameter $y$ from available parameters such that: $y = f(x) + c$.
- Obtaining the spatial distribution of $y$ with $x$ to carry out a weight analysis on the argand plane.
- Expression of $y$ as a polynomial function of $x$ with respect to the best weightage accountability.
- Calculation of $y_j$ at conditionally selected points $x_j$.

Non-linear regression coefficient ($R^2$) is calculated to determine the goodness of fit for $y(x)$ for all statistical correlations proposed in this study [6]. It is calculated as follows:

$$R^2 = 1 - [\sum_{i=1}^{N} (y_i - yf_i)^2 / \sum_{i=1}^{N} (y_i - \bar{y})^2] \quad (2)$$

## 6. Power Loss Calculation

### *6.1. Curtailment Losses*

Data analysis of the three quarantined WTGs for pitch angle variations with wind speed shows that curtailment exists for T002 only. A scatter plot of actual WTG power generation versus wind speed is plotted for sets excluding the sets where curtailment exists. Filter condition for this in MatLab code is set as:

```
if ((T00x_Gen > 10) && (WindSpeed < 9.8) || ((1050 < T00x_Gen) || (T00x_Gen < 900)))
```

$$y = -0.07256x^9 + 0.48027x^8 - 0.20266x^7 - 1.5435x^6 - 5.6741x^5 + 3.5955x^4 + 7.9122x^3 + 69.827x^2 + 402.22x + 433.62 \quad (3)$$

Using statistical correlation, a ninth-degree, centered-and-scaled data polynomial is calculated (Eqn. (3)) that relates actual T002 power generation to wind speed. It is to be noted from Figure 3(a) that this equation is valid only for points above 3.5 and below 12 m/s wind speeds. For speeds greater than 12 m/s, power generation of 1500 kW is assumed.

$$G_e = y(x_j) \quad (4a)$$

$$Loss_{curtailed} = G_e - G_{act}(x_j) \quad (4b)$$

$$NetLoss_{curtailed,new} = NetLoss_{curtailed,old} - NetLoss_{curtailed} \quad (4c)$$

Hence, a value is obtained for the annual generation loss due to curtailment for T002. This value is calculated to be 0.33 Lakh units of electricity for 2011-12. This is a small fraction of the total generation losses of the WTG (7.15%) which are calculated in the forthcoming sections.

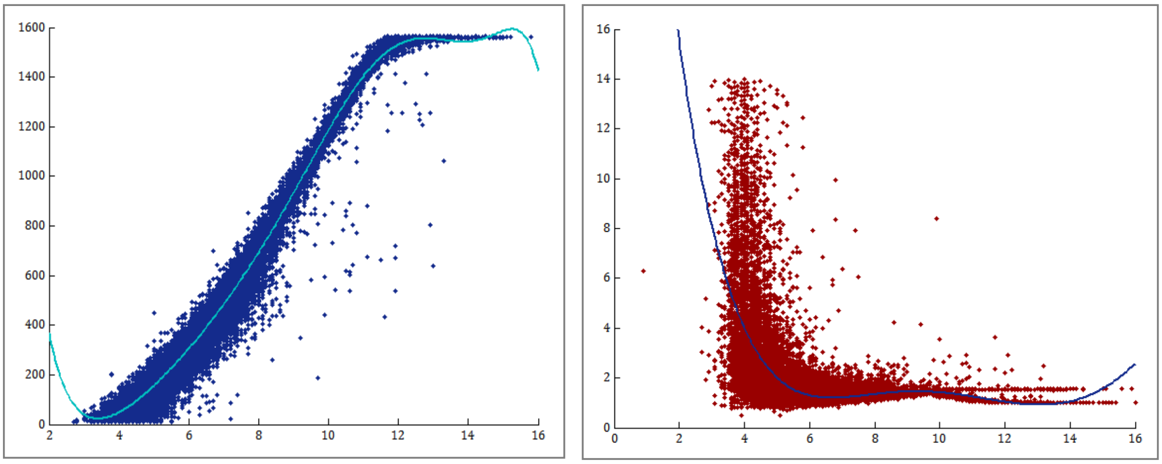


Fig. 3. (a) Actual Power Generation for T002 versus Wind Speed showing fitted polynomial, (b) Plot of $\eta$ versus Wind Speed for T002

### 6.2. Power Generation Losses due to Lack of MA + Lack of GA

#### 6.2.1. Extraction Factor Method (EFM)

The core process in this method is to find the relationship between Ideal Power (IP; the maximum theoretical power limit) and the Actual Power (AP; recorded power generation) at that wind speed using all data points $i$ where the power generation of that WTG exists. An extraction factor $\eta$ is calculated as the ratio of IP to AP at all data points $i$. The filter condition for $i$ in the MatLab code is:

```
if((T00x_gen > 0) && (T00x_gen ~= NotANumber))
```

$$\eta_i = 0.5\rho A C_p v_i^3 / G_i \tag{5}$$

It is to be noted that, the value of $\eta$ is higher at lower wind speeds. Hence a lower fraction of the IP is extracted at lower wind speeds. A scatter plot of the factor $\eta$ versus wind speed is plotted for all $i$ using MATLAB. From this scatter (Figure 3(b)), using statistical correlation, a characteristic polynomial is calculated for every quarantined turbine. This polynomial is then used to predict $\eta$ at all points $j$ where power generation is less than or equal to zero or if we have a lack of GA. The velocity set used is the one obtained via MCPM. Using this $\eta_j$ we find power generation loss $G_j$. By summing up the losses ($G_j$) over all $j$ points, the estimated power lost is calculated. Shown here is the process for T002. Appropriate scaling of points is done according to equation (6a).

$$x = (v_j - 6.7681)/2.2486 \tag{6a}$$

$$\eta_j = -0.0080342x^5 + 0.10635x^4 - 0.38775x^3 + 0.3666x^2 + 0.053408x + 1.2653 \tag{6b}$$

$$Loss_{gen_j} = 0.5\rho A C_p v_j^3 / G_j \tag{6c}$$

$$NetLoss_{gen,new} = NetLoss_{gen,old} + NetLoss_{gen} \tag{6d}$$

It must be noted, that the polynomial obtained is valid only for velocities between 4 and 12 m/s. For velocities greater than 12 m/s, rated generation of 1500 kW is assumed, and the power generation for wind speeds below 4 m/s is taken as zero according to turbine model definition.

#### 6.2.2. Monthly Correlation Method (MCM)

The MCM correlates wind speed with WTG power generation *distinctly* for each month. For a month $n$, data points $i$ are identified where power generation exists (`T00x_Gen`$_i$ `> 0`) and a scatter plot of power generation versus wind speed of these $i$ is generated. From this, a statistically weighted relation is derived such that:

$$G_{expected} = f(v) \tag{7}$$

Using (7), loss of power generation is calculated using MCP data for points $j$ where (T00x_Geni <= 0):

$$Loss_{gen_j} = f(v_j) \tag{8a}$$

$$NetLoss_{gen,new} = NetLoss_{gen,old} + NetLoss_{gen} \tag{8b}$$

Another major difference between this and the other method is that this is done on a monthly basis to obtain a temporally more versatile relation between power and wind speed. It is found that monthly calculation brings a significant change in the calculation from the annual trend.

### 6.3. Power Generation Loss due to Lack of MA

#### 6.3.1. Turbine Specific Power Curve Method (TSPCM)

In this method, a scatter plot between power generation versus wind speed is created for a WTG for every point i where power generation exists. From this, a statistically weighted relation is derived such that:

$$G_{expected} = f(v) \tag{9}$$

Using (10a, 10b), loss of power generation is calculated using CNVA data for points *j* where ($T00x\_Gen_j <= 0$) as follows:

$$Loss_{gen_j} = f(v_j) \tag{10a}$$

$$NetLoss_{gen,new} = NetLoss_{gen,old} + NetLoss_{gen} \tag{10b}$$

In the present study, the following WTG-specific forms of (9) are obtained (plotted in Fig. 4).

$$y_{T002} = -0.16584x^8 + 1.2634x^7 - 0.06920x^6 - 13.491x^5 + 6.2171x^4 + 18.953x^3 + 65.686x^2 + 398.62x + 415.7 \tag{11a}$$

$$y_{T003} = -0.00017735x^8 + 0.014105x^7 - 0.45578x^6 + 7.805x^5 - 77.098x^4 + 443.34x^3 - 1382.3x^2 + 1941.8x + 467.3 \tag{11b}$$

$$y_{T005} = -0.14884x^8 + 1.1499x^7 - 0.075434x^6 - 13.563x^5 + 10.249x^4 + 21.907x^3 + 47.132x^2 + 363.62x + 424.5 \tag{11c}$$

#### 6.3.2. Rated Power Curve Method (RPCM)

In this method, the relation between WTG power generation and wind speed is obtained directly from the WTG Power Curve adapted to the density of the site. This curve, for this particular site, is given by:

$$x = (v_i - 11.95)/4.6909 \tag{12a}$$

$$y = -4.4x^{10} - 20.1x^9 + 42.8x^8 + 159.8x^7 - 175x^6 - 473.1x^5 + 436.4x^4 + 582.9x^3 - 775.3x^2 + 246.8x + 1483.3 \tag{12b}$$

This density adaptation is done by WAsP. Power loss calculation is done in a manner similar to TSPCM (10a, b) using CNVA wind data.

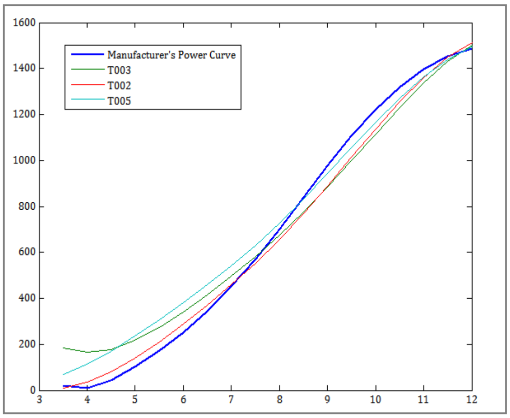


Fig. 4. Power Curves for RPCM (Power Generation in kW, Wind Speed in ms-1)

## 7. Results and Discussion

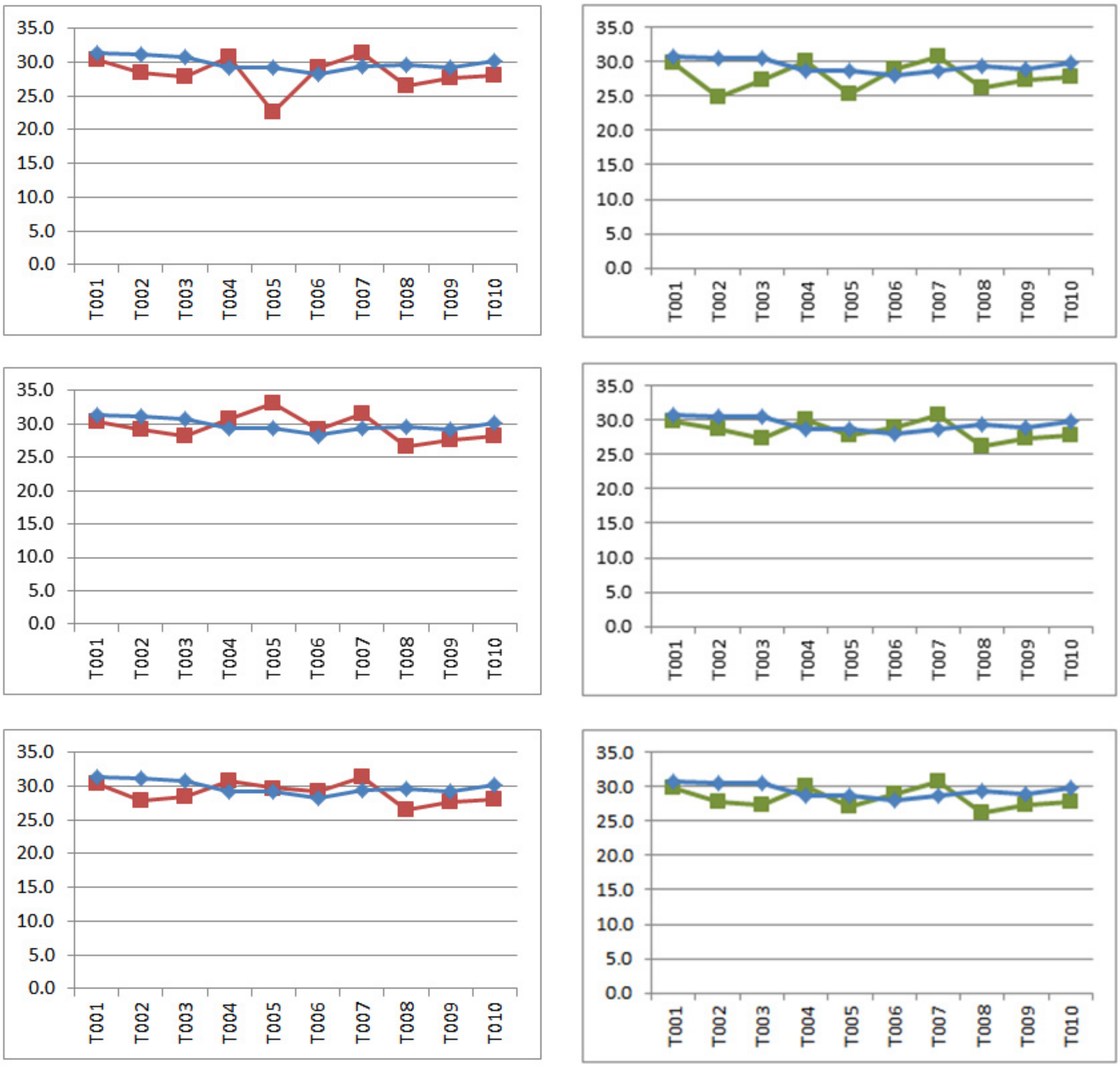


Fig. 5. Left: Actual (red) versus WAsP (blue) annual generation for centum MA and GA obtained using coarse extrapolation (top), EFM (middle) and MCM (bottom) | Right: Actual (green) versus WAsP (blue) annual generation for centum GA obtained using coarse extrapolation(top), TSPC (middle) and Reference Power Curve (PC) method.

A study of the results indicates that for most cases, the gap in predicted and extrapolated data sets was reduced to below 10 percent. The results obtained using the aforementioned methods are validated through statistical correlation. In particular, the following observations and conclusions are important:

- Since data for T003 was incomplete for almost a month, the error in its prediction exists for all methods.

Table 2. Comparison of WAsP Results and Actual Observed Power Generation

| Extraction Factor Method | | | | |
|---|---|---|---|---|
| WTG | MA, GA Losses (lakh units) | WAsP Losses (per cent) | Deviation (lakh units) | R |
| T002 | 4.6287 | 31.2 | -7.3% | 0.9450 |
| T003 | 2.0049 | 30.8 | -9.1% | 0.9222 |
| T005 | 10.3230 | 29.3 | 11.3% | 0.9702 |

| Turbine Specific Power Curve Method | | | | |
|---|---|---|---|---|
| WTG | MA, GA Losses (lakh units) | WasP Losses (per cent) | Deviation (lakh units) | R |
| T002 | 4.2830 | 30.7 | -6.6% | 0.9541 |
| T003 | 1.0847 | 30.5 | -11.6% | 0.9527 |
| T005 | 4.9515 | 28.7 | -3.5% | 0.9470 |

| Monthly Correlation Method | | | | |
|---|---|---|---|---|
| WTG | MA, GA Losses (lakh units) | WAsP Losses (per cent) | Deviation (lakh units) | R |
| T002 | 3.4918 | 31.2 | -11.8% | 0.8047 |
| T003 | 2.3537 | 30.8 | -7.9% | 0.8309 |
| T005 | 6.9633 | 29.3 | 1.5% | 0.8038 |

| Rated Power Curve Method | | | | |
|---|---|---|---|---|
| WTG | MA, GA Losses (lakh units) | WasP Losses (per cent) | Deviation (lakh units) | R |
| T002 | 3.3258 | 30.7 | -10.0% | 0.9513 |
| T003 | 1.2777 | 30.5 | -11.5% | 0.9586 |
| T005 | 4.3775 | 28.7 | -5.8% | 0.9530 |

- The Extraction Factor method, although effective for a distributed lack of MA, underperforms for T005 which has a case of acute lack of MA in the HW season. Hence, over-prediction for HW seasons is a drawback of the EFM.
- The MCM provides very good results for temporally concentrated lack of MA, as is shown in the case of T005. Hence, it can be a great tool for analysis of turbines that have high lack of MA in specific time intervals.
- Out of all methods analyzed, the TSPCM gives the best and the most consistent results over the different scenarios provided by T002, T03 and T005. Except for T003, which has HW data missing, the TSPCM manages to bring down the error in calculation of actual generation losses well below 10%. The polynomial curve constructed shows high value of $R^2$, indicating a goodness of fit. Also, the TSPCM calculates generation losses due to lack of MA, which is of more utility to a WTG company.

**Acknowledgements**

The authors would like to express their gratitude towards Suzlon Group for the opportunity to carry out this project and for providing turbine and mast data for calculations.

**References**

[1] Shika, S., Bhatti, T.S., Kothari, D.P., Wind energy in India: Shifting paradigms and challenges ahead. J Energy Engg (2004); 67-74.

[2] Proceedings of Conference on Developments in Wind Power. Centre for Wind Energy Technology, India. (2011)

[3] Supervisory Control And Data Acquisition - Data Banks. Suzlon Energy Limited, One Earth, Pune, India.

[4] Help and Reference Files. Wind Atlas Analysis and Application Program (WAsP 10.0.0.1) © DTU Wind Energy, Technical University of Denmark, Netherlands

[5] Albers, A., Gerdes, G., Micrositing of Small-scale Wind Turbines is Economically Viable for Rural Electrification in Sri Lanka. DEWI Magazine (1999); 14: p. 24-35.

[6] Holman, J.P., Experimental Methods for Engineers. McGraw Hill Book Company; (1971) p. 37-52.