# Day Ahead Solar Pv Power Forecasting Based On A Combination Of Statistical And Physical Modelling Utilizing Nwp Data For Solar Parks In India

Arindam Roy
NIWE

Aravindakshan Ramanan
GIZ GmbH

Barun Kumar
NIWE

Naveen Kumaar
NIWE

Chris Alice Abraham
NIWE

Anette Hammer
Univ of Oldenburg

Elena Barykina
Overspeed GmbH

Detlev Heinemann
Univ of Oldenburg

Hans Peter Waldl
Overspeed GmbH

Indradip Mitra
GIZ GmbH

Prasun Kumar Das
NIWE

Karthik Ramanathan
NIWE

Balaraman Kannan
NIWE

*Abstract*— **This paper attempts to develop appropriate models for solar power prediction for large MW scale Solar PV park in India. Focus is laid on Solar power prediction of large MW scale PV plant in India with hybrid approach of physical and statistical models . The objective of the exercise is to develop day ahead forecasts of 15 minutes temporal resolution and meet the current regulatory requirements in India mandated by Central Electricity Regulatory Commission. The hybrid method considers physical model supported by NWP data sets and plant specific measurement of irradiation along with Artificial Neural Network and other statistical techniques for bias correction of irradiance and AC power. As input data is critical in physical model, an appropriate filtering techniques on data quality has been carried out. As the Numerical Weather Prediction from NWP is available for every 3 hours/ 1 hours, necessary post processing to interpolate it to 15 minutes time resolution has been carried out. Ground based irradiation measurements was utilized for site specific bias correction. Two different bias correction approaches were cross-compared. The present approach considers Chandrasekhar Diffuse fraction model coupled with Klucher model to convert from irradiation on horizontal plane to panel tilt. For estimation of DC power output from solar PV panel, Huld and Beyer model are considered. Inverter characteristics are modelled using Schmidt Sauer model. These models are tested for various NWP data sets and their statistical combination. The results are validated against actual power output from plant and analysed in the form of NRMSE, NMAE for each 15 minutes time block. It is observed that accuracy of the model is directly dependant on availability of detailed information and data from the plant. It also seen that forecasting in general is within the deviation limits prescribed by regulatory authorities.**

***Keywords- Solar forecasting, Day ahead, physical, statistical, India***

## I. Introduction

The Government has set an ambitious target of 100 GW of grid-connected solar power by the year 2021-22 under the National Solar Mission. The installed capacity of grid connected wind power plants is 34.98 GW and for solar power plants is 24.31 GW [33].

The large scale integration of these renewable sources poses to be a challenge to the grid operators for balancing the demand and generation. In order to maintain grid security, essential measures need to be taken to manage the variability of renewable generation. Forecasting is very essential information for the grid operator which will help in maintain the stability of grid.

In order to align and cater to the needs of current situation, indigenous forecasting model chain was developed. Few large scale solar power plants co-located was considered as test case for solar power forecasting model development. Model was developed considering the system as a single large size solar power plant. The model chain was mixture of physical and statistical approaches. Scientific approach was followed to develop location specific coefficients for solar forecasting model chain. Indigenous data sets namely measured data from Solar power plant, Numerical Weather Prediction models from National Centre for Medium Range Weather Forecasting, etc was employed.

## II. LITERATURE REVIEW

The forecasting model chain consist of converting the irradiance prediction from NWP into tilted radiation and finally combining it with the model of solar PV power plant .Methodology for solar power forecasting is represented in Figure 2. Detailed explanation of the methodology followed in this work is provided in the subsequent section. In this section, different alternative techniques found in the literature for various blocks in solar forecasting model chain is explained.

### *A. Bias Correction of Numerical Weather Prediction Models*

The numerical weather prediction models have a coarse resolution spanning a large area. Therefore, bias following a specific pattern could be identified when it is compared with site specific measurements. There are some methods in the literature through which this bias could be modelled and reduced from forecasts. Some of the methods are detailed in as following:

*1) Bias as a standard bi-variate 4th order polynomial function of the cosine of Solar Zenith angle and Clear Sky Index*

The bias could be modelled as a standard bi-variate 4th order polynomial function of the cosine of Solar Zenith angle and Clear Sky Index. Based on NWP forecasts for a particular location and corresponding measured GHI data, the values of coefficients could be estimated by curve fitting equation.

*2) Bias as a binned function of the cosine of Solar Zenith Angle and Clear Sky Index*

Bias is defined as a binned function of the cosine of Solar Zenith Angle and Clear Sky Index. A user defined 2D bin is created with one edge of sequentially increasing values of $k_t^*$ from 0 till 1 with user defined interval size and other edge specified by increasing values of $(\cos\theta_z)$ [1,2]. The bins are represented in Figure 1. The average bias is calculated for each of the bins in the historical data set. Therefore, for the validation period, bias is assigned to the forecasts based on the bin it represents. The requirement for this method to work accurately is to have large training data set (minimum of one year of data).

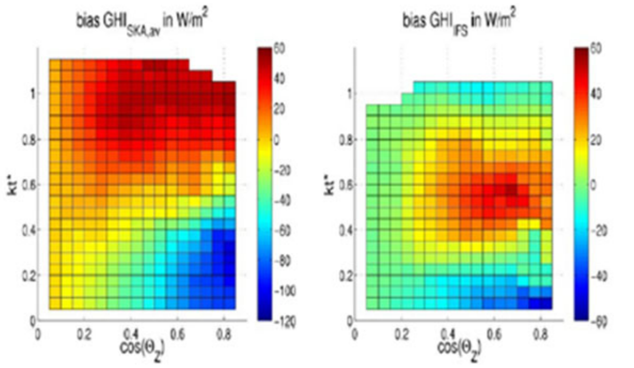

Figure 1:Representation of Binned bias as function of clear-sky index and cosine of zenith angle [2]

*3) Bias as the output of a Feed-forward Neural Network with hidden layers and two inputs – cosine of Solar Zenith angle and Clear Sky Index*

Bias could be modelled as output of a Feed-forward Neural Network with hidden layers and two inputs – cosine of Solar Zenith angle and Clear Sky Index [3]. On the process of training with historical dataset of Numerical Weather prediction models and ground measured data set, weights and offsets of neural network could be obtained.

### *B. Models for Conversion from GHI to GTI*

Diffuse fraction and Diffuse Sky models are empirical models that are employed for conversion of GHI to GTI in the solar forecasting model chain.

*1) Diffuse fraction models*

There are large groups of diffuse fraction models, where the diffuse fraction only depends on the clearness index.

*a) Orgill and Hollands' Model*

The correlation provided by Orgill and Hollands uses four years of data gathered in Toronto, Canada [9] based on the sky cover classified as below:

$d = (1 - 0.249k)$ For $0 < k < 0.35$ (2)
$d = (1.577 - 1.84k)$ For $0.35 \le k \le 0.75$ (3)
$d = 0.177$ For $0.75 < k \le 1$ (4)

*b) Erbs' Model*

Irradiance Data from five stations in the U.S. at different latitudes were used to develop correlations [10]. The model is represented as below:

$d = (1 - 0.09 * k)$ For $0 < k \le 0.22$ (5)

$d = (0.9511 - 0.1604 * k + 4.388 * k2 - 16.638 * k3 + 12.336 * k4)$ For $0.22 < k \le 0.8$ (6)

$d = 0.165$ For $0.8 < k \le 1$ (7)

d is the diffuse fraction
k is the clearness index

*c) Chandrasekar and Kumar Model*

Chandrasekar and Kumar diffuse fraction model is mentioned in the following set of equations [12].

$k_d = 1.0086 - 0.178k \quad for\ k \le 0.24$ (8)

$k_d = 0.9686 + 0.1325k + 1.4183k^2 - 10.1860k^3 + 8.3733k^4 \quad for\ 0.24 < k \le 0.8$ (9)

$k_d = 0.197 \quad for\ k > 0.8$ (10)

- Where $k_d$= diffuse fraction, k= clearness index

*2) Diffuse Sky Models*

The diffuse sky models available in literature are mainly divided into two categories isotropic and anisotropic.

*a) Liu and Jordan Model*

The simplest models are the isotropic model which assumes that diffuse radiation is distributed uniformly over the entire sky dome and it also assumes that reflection from the ground is isotropic. For a surface with tilt angle β, the total irradiance on the tilted surface can be obtained as follows [7] :

$$I_T = I_{hb}R_b + I_{hd}\frac{(1+\cos\beta)}{2} + I_h\,\rho\frac{(1-\cos\beta)}{2} \tag{11}$$

*b) Hay-Davies Model*

The model [8], assumes that the diffuse radiation is composed of isotropic and circumsolar components. The transmittance through atmosphere for beam radiation is defined as the anisotropy index which is used to quantify the amount of diffuse radiation coming from the circumsolar and isotropic components. It is defined as

$$A = \frac{I_{bn}}{I_{on}} \tag{12}$$

The equation for this model is described as :

$$I_T = (I_{hb} + I_{hd}A)R_b + I_{hd}(1-A) * \frac{(1+\cos\beta)}{2} * \sin^3\frac{\beta}{2} + I_h\,\rho\frac{(1-\cos\beta)}{2} \tag{13}$$

$I_{bn}$ direct-normal solar irradiance, W/m$^2$
$I_h$ global horizontal solar irradiance, W/m$^2$
$I_{h,b}$ direct-normal component of solar irradiance on the horizontal surface, W/m$^2$
$I_{h,d}$ global diffuse horizontal solar irradiance, W/m$^2$
$I_{on}$ direct extra terrestrial normal irradiance, W/m$^2$
$I_T$ solar irradiance on the tilted surface, W/m$^2$
$R_b$ variable geometric factor which is a ratio of tilted and horizontal solar beam irradiance
ρ hemispherical-hemispherical ground reflectance
β surface tilt angle from horizon

*c) Klucher Model*

Klucher model is described by the following set of equations [11]:

$$I_T = \frac{(I_H - I_D)*\cos(\psi)}{\sin(\alpha)} + I_D\left(\frac{1+\cos\varepsilon}{2}\right)\left(1 + F\sin^3\left(\frac{\varepsilon}{2}\right)\right)\left(1 + F\cos^2(\psi)\sin^3(90-\alpha)\right) \tag{14}$$

$$F = 1 - \left(\frac{I_D}{I_H}\right)^2 \tag{15}$$

$I_T$=insolation on surface tilted toward the equator at angle $\varepsilon$
$I_H$=total insolation received on horizontal surface
$I_D$=diffuse insolation received on horizontal surface
$(\alpha)$ = solar elevation angle
$(\psi)$= angle between sun direction and normal direction of tilted surface

*C. Models for Conversion of Global Tilted Irradiance into DC Power*

There are number of models which are used for conversion from GTI to DC power. Some of them are detailed below:

*1) Huld Model*

The power output in Huld Model is defined as a function of GTI and Module temperature as per the equation mentioned below [4,5]. Based on the historic measured data of DC Power, GTI and module Temperature, the equation could be trained to find out the coefficients k1 till k6.

$$P(G',T') = G'P_{STC}(1 + k_1\ln(G') + k_2(ln(G'))^2 + k_3T' + k_4T'\ln(G') + k_5T'(ln(G'))^2 + k_6T'^2 \tag{16}$$

$$G' = G/1000(Wm^{-2}) \tag{17}$$

$$T' = T_{mod} - 25°C \tag{18}$$

- $k_1,k_2,k_3,k_4,k_5,k_6$ are Huld coefficients
- G= global inclined irradiance
- $T_{mod}$ is the module temperature

*2) Beyer Model*

The Beyer model is represented by the following equations. On training with historical data sets of module temperature, GTI and Measured DC Power, the values of coefficients a1,a2 and a3 are estimated [6].

$$\eta_{MPP}(G, 25°C) = a1 + a2.G + a3.\ln(G) \tag{19}$$

$$\eta_{MPP}(G,T) = \eta_{MPP}(G,25°C).(1+\propto(T-25°C)) \tag{20}$$

$\eta\,(G,T)$ = efficiency at G, irradiance and T, module temperature
a1, a2, a3 are Beyer coefficients
∝= temperature coefficient

*3) DC Power model based on specifications from PV Module Data Sheet*

Forecasted DC power is calculated as a function of estimated tilted irradiance and module temperature as per the following equation. The parameters in the GTI to DC conversion model are chosen from the datasheet of PV Module and the design of the solar PV power plant.

$$DCPower = GTI * Number\ of\ Panels * Panel\ Area * Efficiency * (1 + \text{alpha} * (\text{temp\_mod} - 25)) \tag{21}$$

alpha= temperature coefficient of PV module.

*D. DC Power to AC Power conversion Model*

The conversion of DC power to AC power is based on the characteristic curve of the PV inverter. If the datasheet with efficiency values at discrete DC input levels are mentioned, the same could be used directly used in the model chain. In the absence of information from inverter datasheet, the "Schmidt and Sauer" could be used.

*1) Schmidt and Sauer Model*

Schmidt and Sauer model was used in the inverter modelling as per the following equations. The parameters of the Schmidt and Sauer model (Pself, Vloss, rloss) will be found out by fitting the equation with input of DC and AC power measurements.

$$Ploss = Pself + Vloss.Pout + rloss.P^2out \tag{22}$$

- Pself, Vloss, rloss are inherent, magnetic, resistive loss coefficients respectively for inverter model
- Equation (23) is valid when rloss is positive. Equation (24) is applicable when rloss is negative.

$$\eta * = -\frac{1+v*_{loss}}{2.r_{loss}.p_{in}} + \sqrt{\frac{(1+v*_{loss})^2}{(2.r*_{loss}.p_{in})^2} + \left(\frac{p_{in}-p*_{self}}{r*_{loss}.p_{in}^2}\right)} \quad (23)$$

$$\eta * = -\frac{1+v*_{loss}}{2.r_{loss}.p_{in}} - \sqrt{\frac{(1+v*_{loss})^2}{(2.r*_{loss}.p_{in})^2} + \left(\frac{p_{in}-p*_{self}}{r*_{loss}.p_{in}^2}\right)}$$
(24)

## III. Data Sets Used

### *A. NWP Data*

Two Numerical Weather Prediction models provided by Europe and India were used in the current work. The time and spatial resolutions of NWP models are given in Table 1.

*Table 1: Description of spatial and temporal resolution of NWP model*

| NWP Model Name | Spatial Resolution | Original Temporal resolution | Interpolated Temporal resolution (not given by the provider) |
|---|---|---|---|
| ECMWF | 0.25 degree x 0.25 degree (Approx. 25 km * 25 km) | 3 hours | 15 minutes |
| NCMRWF | 0.25 degree x 0.25 degree (Approx. 25 km * 25 km) | 1 hour | 15 minutes |

### *B. Power Plant Data*

Static and dynamic data were received from the reference power plants were utilized for the forecasting model development. GHI, GTI, AC Power, Module Temperature measurements were received from the power plants. The different measurements were passed through rigorous quality control algorithm (stuck values, physical range check, night time check, etc.) and used in the model development.

## IV Methodology, Results and Observations

The flowchart of entire process of forecasting model chain is shown in Figure 2. Forecasting model is developed for Day ahead forecasting of solar power generation for a group of large scale power plants geographically located together. The original NWP's are interpolated to 15 minutes time resolution which is the scheduling resolution. The interpolation is performed by clear-sky method. Different approaches are tried in various parts of the model chain. Accuracy measures of different techniques are cross compared with each other.

Two bias correction methodologies namely "Polynomial Method" and "ANN method" were cross compared for its effectiveness. The concept of both the methods are explained in the previous section. The total validation period considered for the analysis is between 17th December 2018 till 16th June 2019. Sliding window approach was adopted to test the methods for different set of training days. Since, the validation data period must be the same when different training days of sliding window are cross-compared, corresponding period before 17th December, 2018 is chosen for training the first set of coefficients. The coefficients/weights are calculated by fitting the polynomial equation, Neural network system with ground measured GHI data. The bias correction results for both the NWPs and for the two different methods is shown from Figure 4 till Figure 7. The results are expressed in relative terms. It can be seen "Polynomial method" is showing better results than ANN based approach. Training period of 25 and 30 days are showing the best results.

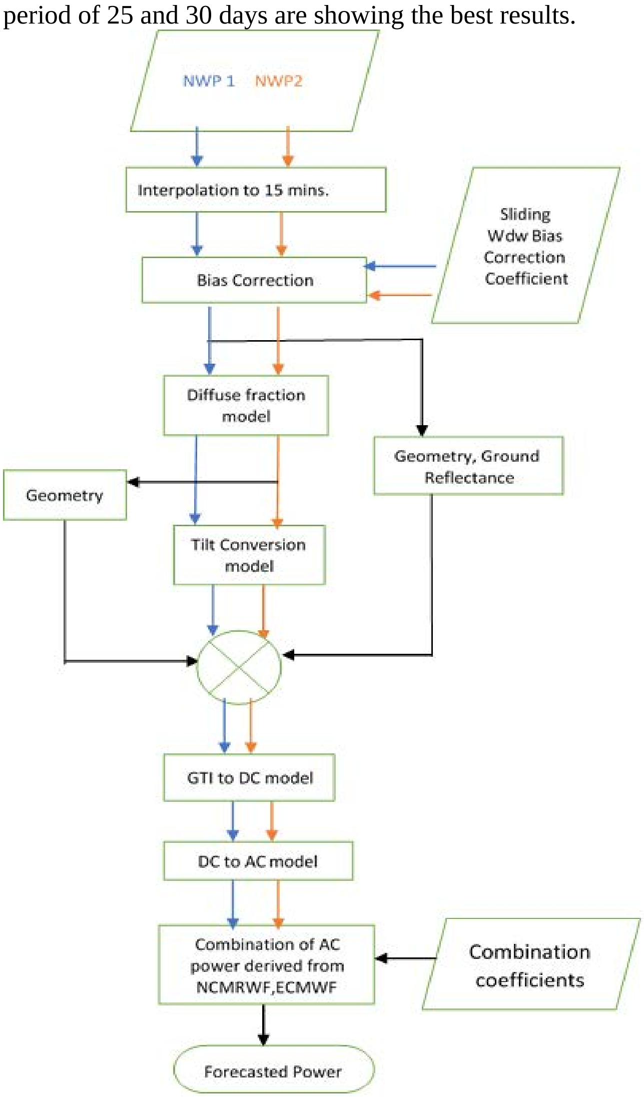


*Figure 2:Flow chart of the methodology followed for the forecasting model chain*

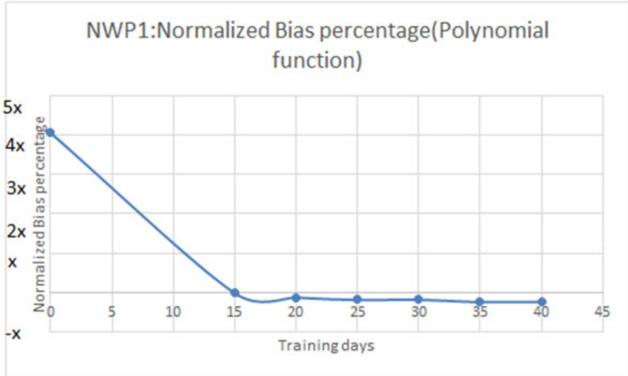


*Figure 3:NMBE (in relative terms) vs training days for NWP1 of Polynomial Function bias correction method*

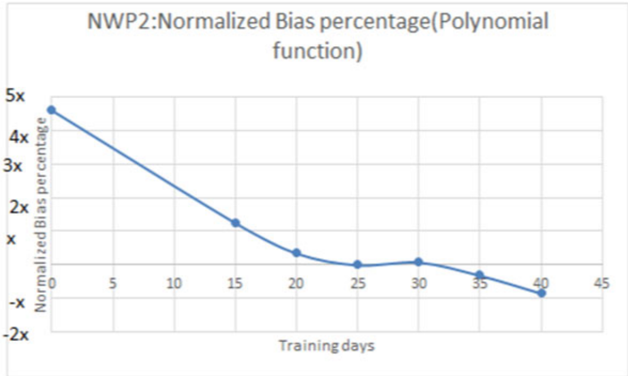


*Figure 4:NMBE (in relative terms) vs training days for NWP2 of Polynomial Function bias correction method*

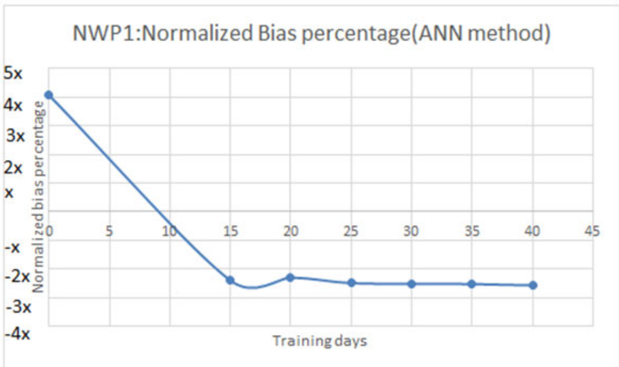


*Figure 5: NMBE (in relative terms) vs training days for NWP1 of ANN bias correction method*

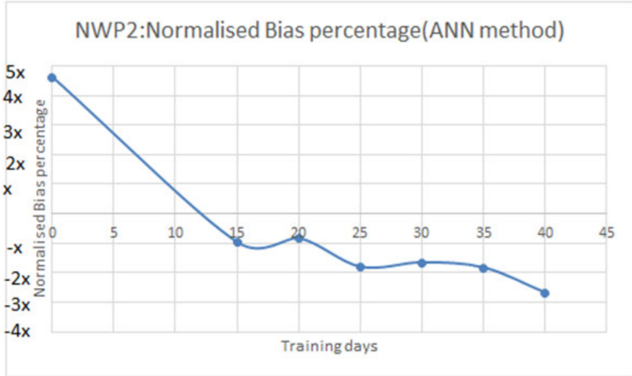


*Figure 6: NMBE (in relative terms) vs training days for NWP2 of ANN bias correction method*

Diffuse fraction model of "Chandrasekar and Kumar" and Tilt Conversion model of "Klucher" was employed in the model chain. The effectiveness of two GTI to DC models (Huld and Beyer) were cross compared. The equations of the models were shown in the previous section. Since DC power information was not available, it was estimated by reverse-calculating from Schmidt-Sauer model. Corresponding coefficients for Huld and Beyer model was calculated by fitting the equations with Module temperature, GTI measurements and estimated DC power. The Schmidt and Sauer coefficients, were calculated for different power plant where DC and AC power measurements were available. The same set of coefficients were used in this specific forecasting model chain. Combination of AC power were also performed. It was performed by the following formula:

$$ACPowerFcst_{Combi} = a1 * ACPower\ Fcst_{NCMRWF} + a2.* ACPowerFcst_{ECMWF}$$

The coefficient of the combination equation was trained on 15 days sliding window approach. The validation period for AC power validation is between 1$^{st}$ May 2019 till 16$^{th}$ June 2019. The results for the validation of AC power forecast with both the NWPs and the combination is shown in Table 2. The performance of NWP1 is better compared to NWP2 based model chain. The accuracy of combination of NWP's is better compared to individual NWP performance.

The weightage of different NWP's are different for the validation days. Relative performance of the NWP's are varying in the validation time period. The Scatter plots of measured and forecast (combination) value is represented in Figures 7 and 8.

*Table 2: The results for validation of AC power forecasts with individual NWP's and Combination*

| MODEL | NCMRWF AC FORECAST | | | ECMWF AC FORECAST | | COMBINATION OF NCMRWF & ECMWF AT POWER END | | | |
|---|---|---|---|---|---|---|---|---|---|
| | nRMSE (%) | nMAE (%) | nMBE (%) | nRMSE (%) | nMAE (%) | nMBE (%) | nRMSE (%) | nMAE (%) | nMBE (%) |
| BEYER | X1 | Y1 | Z1 | 0.908 X1 | 1.088 Y1 | - 0.314Z1 | 0.861 X1 | 0.955 Y1 | 1.04 Z1 |
| HULD | 1.014 X1 | 1.13 Y1 | - 1.064Z1 | 0.895 X1 | 1.025 Y1 | 0.232Z1 | 0.866 X1 | 0.955 Y1 | 1.17 Z1 |

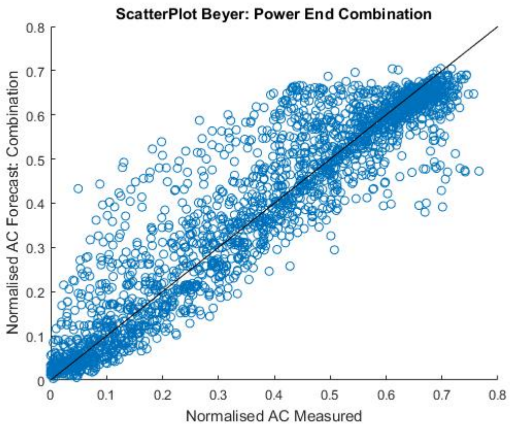


*Figure 7:The Scatter plot of Combined AC power forecast and Measured AC power (GTI to DC model: Beyer)*

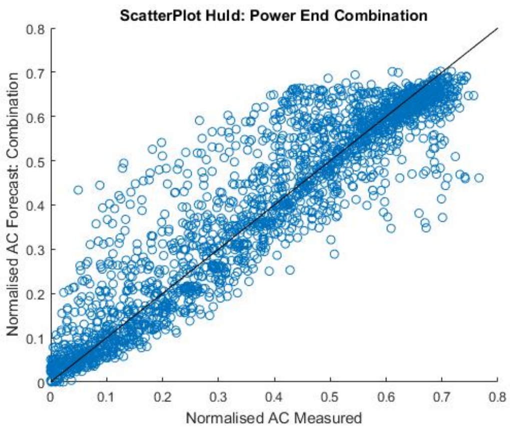


*Figure 8: The Scatter plot of Combined AC power forecast and Measured AC power (GTI to DC model: Huld)*

## V Discussion and Conclusion

Indigenous forecasting model was setup to forecast the power generation of group of solar power plants. Maximum datasets from Indian sources has been employed in the exercise. The forecasting system was evaluated to have good accuracy. Step by step scientific approach was followed in development of the model.

Some of the steps which would aid in improvement of model performance is listed in the following:

- Addition of aerosol information in NWP model by the meteorological agencies.
- Bias correction coefficients calculated on seasonal basis.
- Testng with advanced diffuse fraction models.
- Analyzing the NWP model performance based on weather classes defined by meteorological agencies and development of corresponding combination coefficients.

Its very important to receive all the possible data with highest quality for optimum performance of the model. In the current work, DC power data (dynamic), the exact period in which tilt angle is changing (manual seasonal tilt), were not received from the plant operator. The data from Availability Based Tariff meters is more accurate. However, this data is being received only in the interval of 15 days. The data from SCADA requires a lot of cleaning. It was also observed that pyranometers and other sensors are not being maintained well. Even the highest quality instruments not properly maintained, will result in poor data which will affect the model performance. Diffuse Horizontal radiation are not measured by Power plant operators. Availability of the above data will help is development of site-specific diffuse fraction models.

It is expected that as more quality data and implementation of advanced modelling approaches, the developed indigenous forecasting model chain will attain international standards.

## VI References

1. Elke Lorenz, Johannes Hurka, Detlev Heinemann and Hans Georg Beyer, “Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems”, IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, March 2009.
2. Elena Barykina,” Seminar on solar power forecasting”, January 2018.
3. Philippe Laureta, Maïmouna Diagnea, Mathieu David, “A Neural Network Post-processing Approach To Improving NWP Solar Radiation Forecasts”, Energy Procedia, 2014.
4. Thomas Huld, Gabi Friesen , Artur Skoczek, Robert Kenny, Tony Sample, Michael Field, Ewan D. Dunlop “A power-rating model for crystalline silicon PV modules”, August 2011.
5. Dr. Annette Hammer, “Seminar on power estimation in solar PV power plants”, November 2017.
6. Hans Georg Beyer, Jethro Bethke, Anja

Drews, Detlev Heinemann, Elke Lorenz, Gerd Heilscher, Stefan Bofinger, “Identification Of A General Model For The Mpp Performance Of Pv-Modules For The Application In A Procedure For The Performance Check Of Grid Connected Systems”, January 2000.

7. Liu, B. and Jordan, R. (1963). The long-term average performance of flat-plate solar-energy collectors. *Solar Energy*, 7(2), pp.53-74.

8. Hay, J.E., Davies, J.A., (1980). Calculations of the solar radiation incident on an inclined surface. In: Proceedings of First Canadian Solar Radiation Data Workshop. Canadian Atmospheric Environment Service, Canada, pp. 59–72.

9. Orgill, J.; Hollands, K. Correlation Equation For Hourly Diffuse Radiation On A Horizontal Surface. *Solar Energy* 1977, *19*, 357-359.

10. Erbs, D.; Klein, S.; Duffie, J. Estimation Of The Diffuse Radiation Fraction For Hourly, Daily And Monthly-Average Global Radiation. *Solar Energy* 1982, *28*, 293-302.

11. T. Klucher, “Evaluation of models to predict insolation on tilted surfaces”, Solar Energy, vol. 23, no. 2, pp. 111-114, 1979.

12. J. Chandrasekaran and S. Kumar, “Hourly diffuse fraction correlation at a tropical location”, Solar Energy, vol. 53, no. 6, pp. 505-510, 1994.

13. H.Schmidt, D.U.Sauer, (1996). Wechselrichter-Wirkungsgrade. Sonnenenergie 4, 43-47, 1996