

VALIDATION OF GHI AND DNI PREDICTIONS FROM GFS AND MACC MODEL IN THE MIDDLE EAST

Luis Martin-Pomares¹, Jesus Polo², Daniel Perez-Astudillo¹ and Dunia A. Bachour¹

¹ Qatar Environment and Energy Research Institute (QEERI), HBKU, Qatar Foundation, Doha, Qatar

² Energy Department, CIEMAT, Madrid, 28040, Spain

Abstract

Production of electricity from solar energy is gaining a tremendous significance. The integration of all solar energy power to the electricity grid challenges new horizons. Mainly, the prediction of short-term power generation to optimise its management, avoid situations of load reduction and anticipate supplying problems.

This paper presents a methodology to forecast hourly global horizontal solar irradiance (GHI) and direct normal irradiance (DNI) using global forecast system (GFS) model from NOAA and MACC model from ECMWF. Three clear sky models are tested to increase temporal resolution of GFS model from 3 hours to 1 hour. The forecasting horizon of the predictions is six days. The model has been validated using a ground radiometric station in Qatar with data from 2014 to 2015. The errors of the best model tested are 19% in terms of relative RMSD and -1.68% regarding relative bias for GHI. In the case of DNI, relative RMSD is 48.43%, and relative bias is -3.40%.

Keywords: *solar radiation forecasting, AOD forecasting, DNI prediction, Atmospheric aerosols*

1. Introduction

Solar energy is gaining an enormous significance due to the unsustainable current energetic model. In the case of solar power, technology to produce electricity, the integration of all these power to the electricity grid challenges new horizons. The most important are the estimation of short range power generation to optimise its management, avoid situations of load reduction and anticipate supplying problems (Pelland, Remund, Kleissl, Oozeki, & De Brabandere, 2013).

In this study, we present a methodology to forecast hourly global horizontal solar irradiance (GHI) and direct normal irradiance (DNI) using cloudiness from global forecast system (GFS) model from NOAA and AOD and water vapour predictions from MACC model from ECMWF. The forecasting horizon of the predictions is six days. The model has been validated using a ground radiometric station in Qatar with data from 2013 to 2015.

2. Experimental Solar Radiation Dataset

The radiometric data employed in this study for the validation is located in the Middle East. The instruments in the station are Kipp & Zonen high precision and measurements are available at the end of November 2012. The station is equipped with a Solys2 sun tracker with sun sensor kit for improved tracking accuracy and shading ball assembly for diffuse measurements. Mounted on the sun tracker are one CHP1 pyr heliometer for measuring DNI and two CMP11 pyranometers (one of them shaded) for GHI and DHI measurements; both pyranometers are fitted with CVF 3 ventilation units. Data from the monitoring station are collected as minute

by minute averages in W/m². Maintenance and checks are conducted on a daily basis and consist of cleaning all sensors and tests on the tracking and alignment of the sensors and shadow ball as well as visual checks on the collected data (Bachour & Perez-Astudillo, 2014; Perez-Astudillo & Bachour, 2014). Two type of quality filters has been applied to the data. The first group of filters is based on physical limits of solar radiation. We have removed from the comparison values lower than zero and higher than extraterrestrial solar radiation. The second group of filters is more restrictive and is based on BSRN quality filters (Long & Dutton, 2002).

3. Methodology

The model developed uses 3-hourly solar radiation predictions from GFS model to obtain hourly values using a statistical downscaling model and a clear sky model.

The Global Forecast System (GFS) is a global numerical weather prediction model from NOAA. The GFS model covers the entire globe at a base horizontal resolution of 28 kilometres between grid points, which is used by the operational forecasters who predict weather out to 16 days in the future. In this work, we obtain hourly predictions for six days ahead.

The predictions of hourly GHI are calculated using directly the output variable DSWRF at a surface level from GFS model. The nearest grid output point of GFS model to each station is used as the basic predictions. Afterwards, 3-hourly values from GFS model are calculated as hourly values with the interpolation of k_c which is the quotient between measured global solar radiation and clear sky solar radiation. Clear sky solar radiation is calculated with three different models as they behave better than others depending on the climatic conditions of the sites. ESRA (Rigollier, Bauer, & Wald, 2000), SOLIS (Ineichen, 2008) and REST2 (Gueymard, 2008) clear sky models have been tested. Hourly atmospheric parameters of AOD at 550nm and water vapour column are obtained from MACC-II model (Monitoring Atmospheric Composition & Climate from ECMWF) for clear sky models. The spatial resolution of MACC-II model is 1.125°x1.125°. MACC-II uses a comprehensive global monitoring and forecasting system that estimates the state of the atmosphere on a daily basis, combining information from models and observations, and it provides a daily 5-day forecast.

Once GHI hourly predictions are calculated, DNI hourly predictions are obtained using DirIndex model (Perez, P.Ineichen, E.Maxwell, R.Seals, & A.Zelenka, 1992).

An algorithm to detect clear sky days has been applied (Zarzalejo, Polo, Martín, Ramírez, & Espinar, 2009) to calculate to clear sky conditions. In such a case, the prediction for the next day for GHI and DNI components are made using the clear sky model. Fig 1. shows the schema of the methodology employed to obtain hourly predictions of GHI and DNI for six days ahead.

4. Results

The parameters to calculate the errors of the models are mean bias deviation (MBD), root mean squared deviation (RMSD) and its relative value as rMBD and rRMSD. The relative expressions are obtained using the following equations:

$$\text{rMBD} = \frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i) \quad (\text{eq. 1})$$

$$\text{rRMSD} = \frac{\sqrt{\sum_{i=1}^n (x_i - \hat{x}_i)^2 / n}}{\bar{x}} \quad (\text{eq. 2})$$

where N is the population size, x is the observed value, \bar{x} is the mean of observed values and \hat{x} is the predicted value.

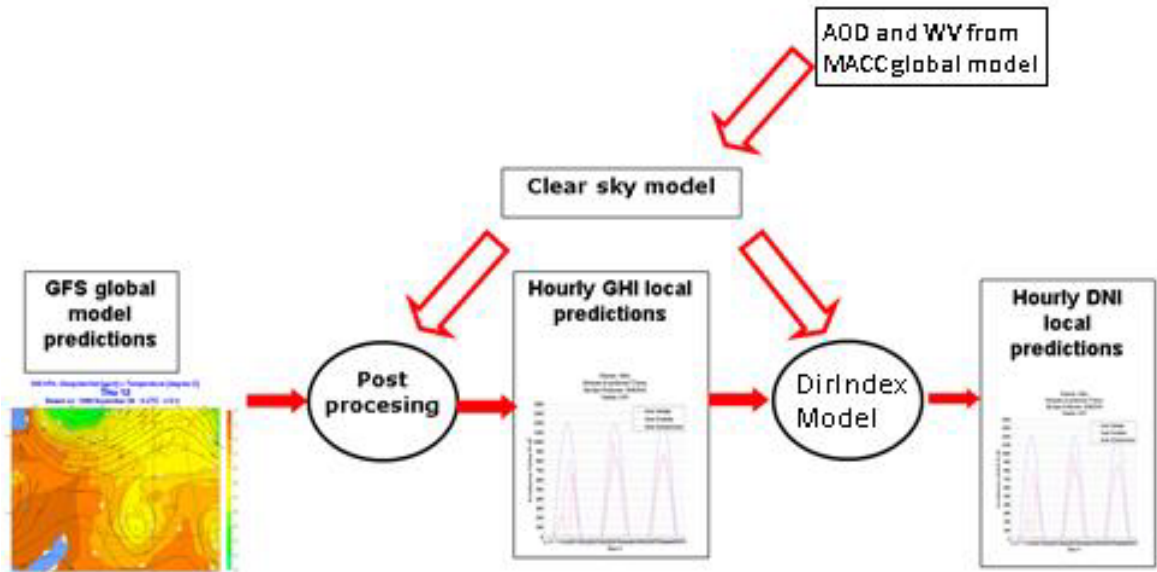


Fig. 1. Schema of the methodology used to obtain hourly predictions of GHI and DNI

Last parameters are often insufficient to establish a complete, coherent comparison for benchmarking. Due to this, additional measures for model quality based on the analysis of cumulative distribution function (CDF) are used. A comprehensive approach to analyse the deviations of measured and predicted data is applied. It is based on the Kolmogorov-Smirnov (KS) test and defines new parameters to quantify the similarity of the two CDFs. Although there are several statistical tests and ways of evaluating the goodness of a model, the KS test has the advantage of making no assumption about the data distribution and is thus a non-parametric, distribution-free test. The KS test determines if two datasets differ significantly. The test consists of comparing the distribution of a dataset to a reference distribution. It can be done by converting the list of data points to an unbiased estimator $S(x_i)$ of the CDF, $i = 1..N$, N is the population size. The KS statistic D is defined as the maximum value of the absolute difference between the two CDFs:

$$D = \max |S(x_i) - R(x_i)| \quad (\text{eq. 3})$$

where $R(x_i)$ is the CDF of the reference dataset. Thus, if the D statistic is lower than the threshold value V_c , the null hypothesis that the two datasets come from the same distribution cannot be rejected. The critical value depends on N and is calculated for a 99% level of confidence as:

$$V_c = \frac{1.63}{\sqrt{N}}, \quad N \geq 35 \quad (\text{eq. 4})$$

This test detects smaller deviations in cumulative distributions than the χ^2 test does. However, instead of using the original one, an extended KS test is used, in which the distances between the CDFs are calculated over the whole range of the variable x , i.e. the solar radiation. A discretization in $n = 1..m$ levels is applied here. In the following $m = 100$ intervals as a reasonable choice are used. Greater order of magnitude for m is not recommended since it implies more computational cost for no improvement in the accuracy of the result. The interval distance p is thus defined as:

$$p = \frac{x_{\max} - x_{\min}}{m}, \quad m = 100 \quad (\text{eq. 5})$$

where x_{\max} and x_{\min} are the extreme values of the independent variable. Then, the distances between the CDFs are defined, for each interval, as:

$$D_n = |S(x_j) - R(x_j)|, \quad x_j \in [x_{\min} + (n-1)p, x_{\min} + np] \quad (\text{eq. 6})$$

The representation of the values D_n , along with the critical value, gives the complete testing behaviour of the CDF on the reference over the whole range. However, although application of the KS test contributes valuable information, it only materialises in the acceptance or rejection of the null hypothesis. In the next paragraphs, new parameters are proposed, which, based on the estimation of the distance between the two CDFs for the sets compared, define quantitative measures that can be used to rank models (Espinar et al., 2009).

The KSI parameter (Kolmogorov-Smirnov test Integral) is defined as the integral of the area between the CDFs for the two sets. The unit of this index is the same for the corresponding magnitude, the value of which depends on it. The KSI is defined as the integral:

$$KSI = \int_{x_{\min}}^{x_{\max}} D_n dx \quad (\text{eq. 6})$$

As D_n is a discrete variable and the number of integration intervals is identical in all cases, trapezoidal integration is possible over the whole range of the independent variable x . A percentage of KSI is obtained by normalising the critical area, $a_{critical}$:

$$KSI\% = \frac{\int_{x_{\min}}^{x_{\max}} D_n dx}{a_{critical}} * 100 \quad (\text{eq. 6})$$

where $a_{critical}$ is calculated as:

$$a_{critical} = V_c * (x_{\max} - x_{\min})$$

where V_c is the critical value for the level of confidence selected and (x_{\max}, x_{\min}) are the extreme values of the independent variable. Normalisation to the key area enables the comparison of different KSI values from various tests. The minimum value of the KSI Index is zero, in which case, it can be said that the CDFs of the two sets compared are the same.

Over, Over %, KSE and RIO parameters are also used to calculate the error of the models. More information about them can be found in (Espinar et al., 2009).

The assessment of the models proposed is performed with hourly ground radiometric data (GHI and DNI) for the period from 1st of January 2014 to 28th of February 2015. In the validation, only values for solar zenithal angle higher than 85 degrees are considered.

4.1 Results for ESRA clear sky model

In this section, the results for ESRA clear sky model is presented. The next graphics show the scatter plot for predicted and measured hourly GHI and DNI.

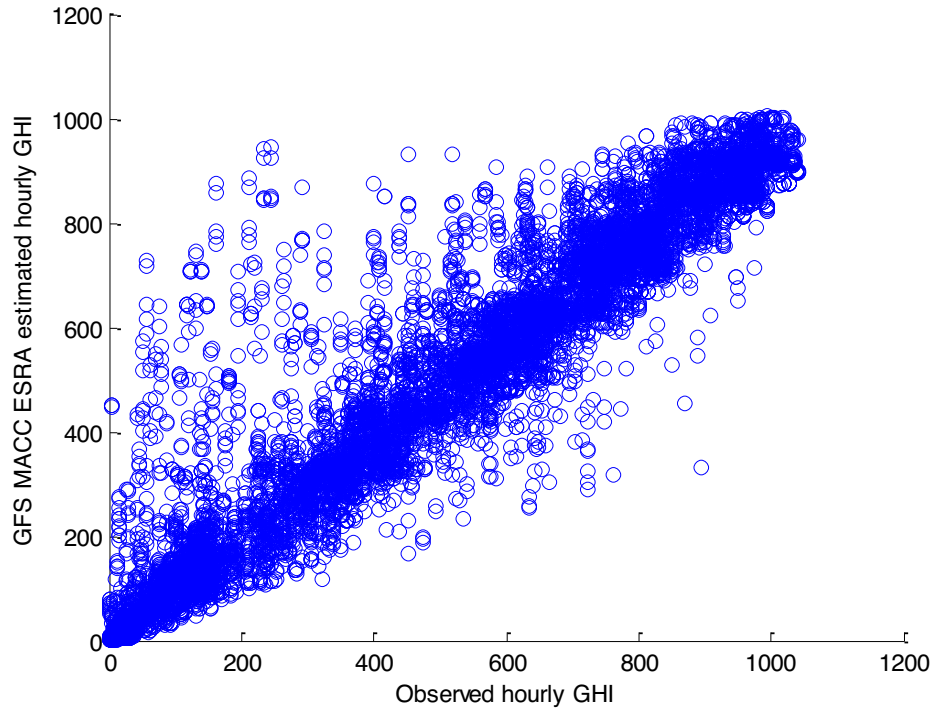


Fig. 2: Scattering plot of GFS ESRA clear sky model predictions and measurements of GHI for the next six days.

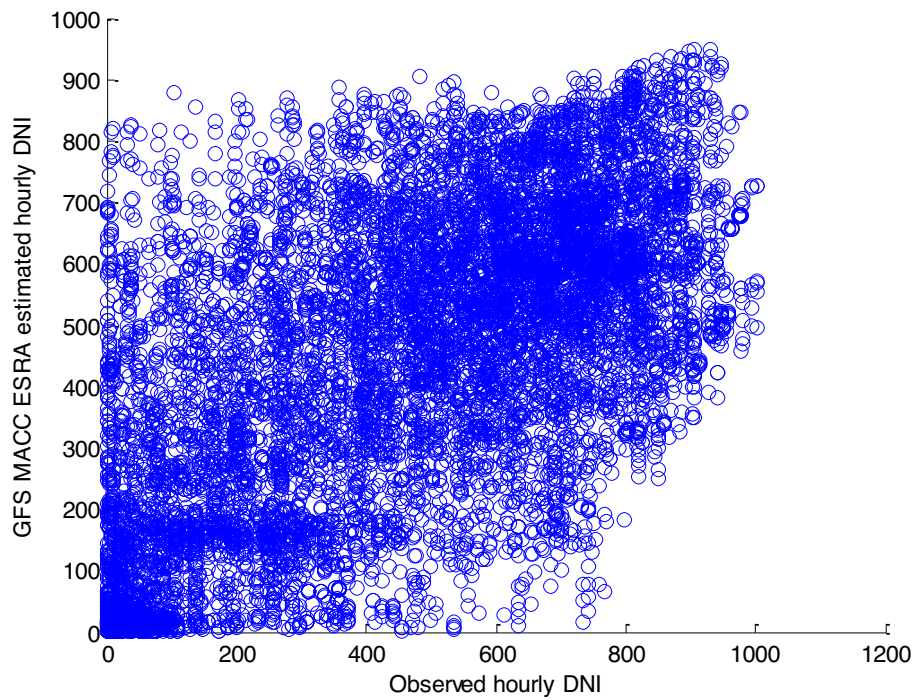


Fig. 3: Scattering plot of GFS ESRA clear sky model predictions and measurements of DNI for the next six days.

As can be seen in the last scatter plots, predictions of GHI are much better correlated than DNI predictions.

Next figures present forecasts (in blue colour) and measurements (in green colour) of hourly GHI and DNI and extraterrestrial solar irradiance (in red colour) in true solar time temporal reference for date 01/01/2014 for six days of prediction ahead as an example of the results of the model.

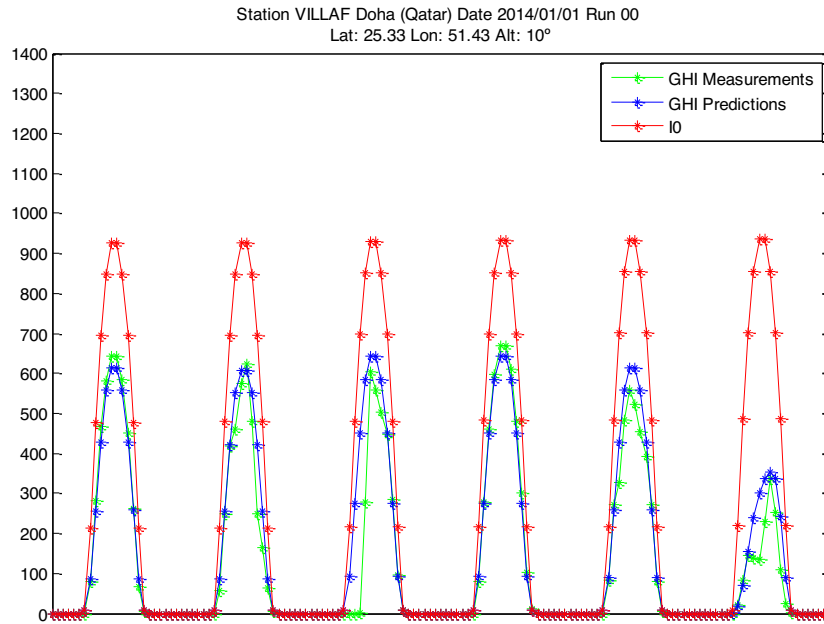


Fig. 4: GFS ESRA clear sky model predictions and measurements of hourly GHI for the next six days.

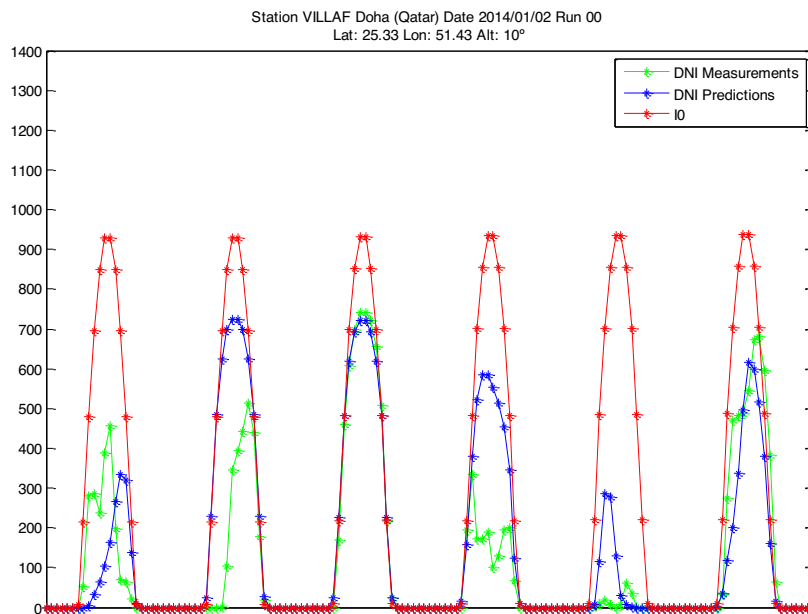


Fig. 5: GFS ESRA clear sky model predictions and measurements of hourly DNI for the next six days.

The next tables show MBD, RMSD, rMBD, rRMSD and R for GHI and DNI, for each day ahead of prediction from 1 to 6 and total errors of the models.

Tab. 1: Statistical errors of hourly GHI predictions validated with ground measurements. ERSA clear sky model

Day ahead	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)	R
1	-11.39	87.54	-2.32	17.82	0.96
2	-8.36	94.34	-1.70	19.23	0.95
3	-8.61	94.88	-1.75	19.27	0.95
4	5.37	98.17	-1.08	19.80	0.95
5	-9.86	89.15	-1.98	17.87	0.96
6	-6.37	98.90	-1.28	19.91	0.95
TOTAL	-8.32	93.93	-1.68	19.00	0.95

Tab. 2: Statistical errors of hourly DNI predictions validated with ground measurements. ERSA clear sky model

Day ahead	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)	R
1	-19.17	215.17	-4.17	46.78	0.67
2	-14.23	226.91	-3.10	49.41	0.62
3	-15.72	225.30	-3.42	48.96	0.63
4	-12.96	229.82	-2.78	49.33	0.61
5	-16.53	222.18	-3.56	47.77	0.63
6	-15.67	222.72	-3.39	48.25	0.64
TOTAL	-15.71	223.73	-3.40	48.43	0.63

GHI predictions are almost unbiased, and rRMSD is reasonably small. The value of rMBD decreases as the temporal horizon of prediction increases. On the contrary, the value of rRMSD increases as the horizon of prediction increases. The predictions are well correlated as the total mean value of R is 0.95.

DNI predictions have a small bias of around 3%. rRMSD is big in the case of DNI. The value of rMBD decreases as the temporal horizon of prediction increases. On the contrary, the value of rRMSD increases as the horizon of prediction increases. The predictions are relatively well correlated as the total mean value of R is 0.63.

Next, we present tables with MBD, RMSD, rMBD and rRMSD for GHI and DNI, for each hour ahead forecasted in true solar time (TST) within the first day of prediction.

In the case of GHI, rRMSD is lower in the central hours of the day and considerably high during sunrise and sunset. The predictions are biased during the first hours of the day, most probably due to some walls near the station which hide the horizon when the sun is appearing in the sky. During the rest of the day, the model presents unbiased results with a considerable value in the last two hours of the day.

In the case of DNI, we have almost the same analysis. rRMSD is lower in the central hours of the day and considerably high during sunrise and sunset. In this case, the predictions of DNI are not biased during the first

hours of the day because there is no wall between the pyrheliometer and the sun disk. During the rest of the day, the model presents unbiased results with a considerable value in the last two hours of the day.

Tab. 3: Statistical errors for each hour of the first day of prediction. Hourly GHI predictions validated with ground measurements. ERSA clear sky model

Hour ahead in TST	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)
7	9.69	22.71	9.29	21.78
8	5.83	59.50	2.56	26.09
9	-7.46	77.55	-1.69	17.61
10	-5.20	104.58	-0.85	17.01
11	-8.01	112.95	-1.07	15.06
12	-18.97	116.95	-2.31	14.25
13	-23.85	113.39	-2.89	13.73
14	-20.70	106.64	-2.71	13.94
15	-15.88	98.49	-2.52	15.61
16	-15.41	80.41	-3.41	17.81
17	-24.90	61.58	-9.99	24.70
18	-9.17	27.92	-12.77	38.90

Tab. 4: Statistical errors of hourly DNI predictions validated with ground measurements. ERSA clear sky model

Hour ahead in TST	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)
7	0.58	127.90	0.37	80.52
8	5.65	203.90	1.88	67.98
9	-21.56	231.66	-4.50	48.35
10	-13.18	239.33	-2.34	42.39
11	-9.16	231.93	-1.51	38.22
12	-12.71	226.03	-2.01	35.67
13	-16.73	232.43	-2.62	36.35
14	-18.56	233.91	-2.98	37.57
15	-15.83	232.01	-2.77	40.65
16	-26.62	232.13	-5.56	48.47
17	-67.71	222.38	-20.16	66.22
18	-34.74	129.01	-28.99	107.69

The next graphics show the Kolmogorov-Smirnov test (D_n) for hourly GHI and DNI. The black dotted line represents the critical limit, V_c , calculated for the number of values available.

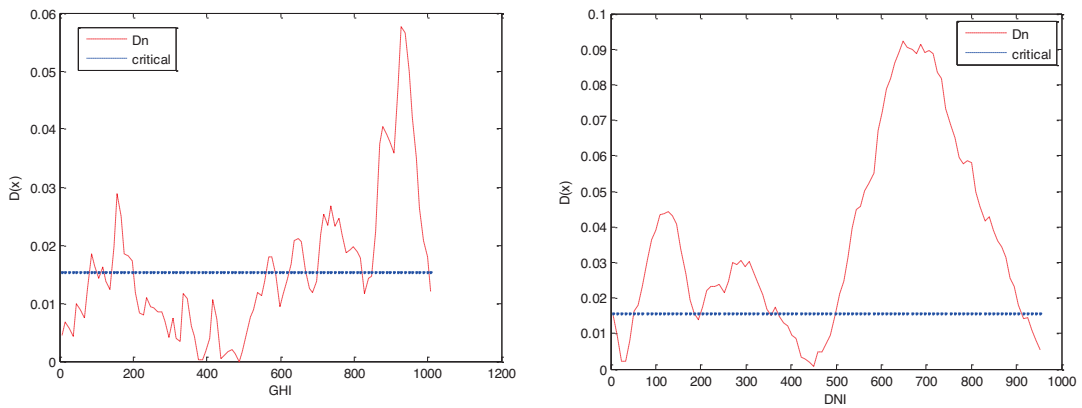


Fig. 6: Kolmogorov-Smirnov test for hourly GFS predictions and ground measurements of GHI and DNI (Wh/m^2) for Doha.

As we can see from the last graphics, there is a notorious difference in GHI predictions and measurements for high values. In the case of DNI, the difference is notorious for low and high values.

Next graphics show the distribution of hourly measurements (red line) and predictions (blue line) for GHI and DNI. The frequency of the range of values from 0 to 1000 Wh/m^2 for intervals of 100 Wh/m^2 is presented.

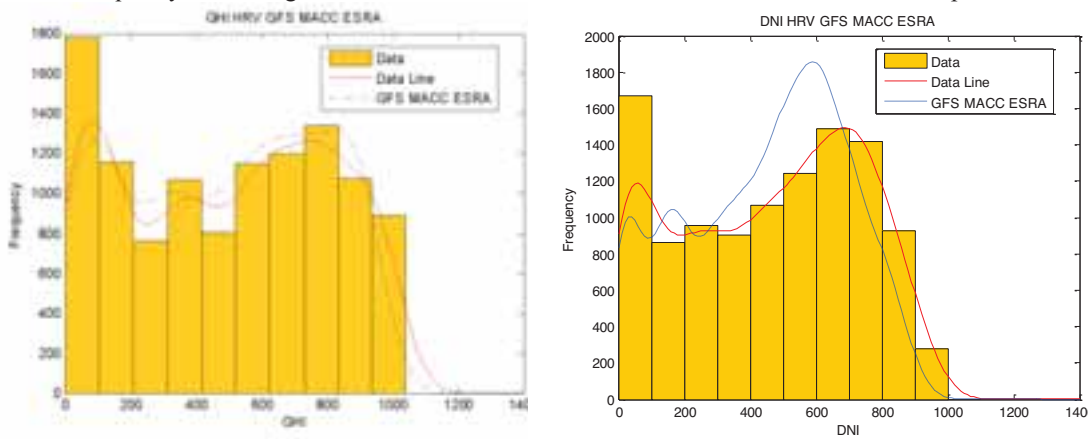


Fig. 7: Cumulative distribution function of hourly GHI and DNI predictions in DOHA.

Next graphics show the Cumulative Distribution Function of hourly GHI and DNI:

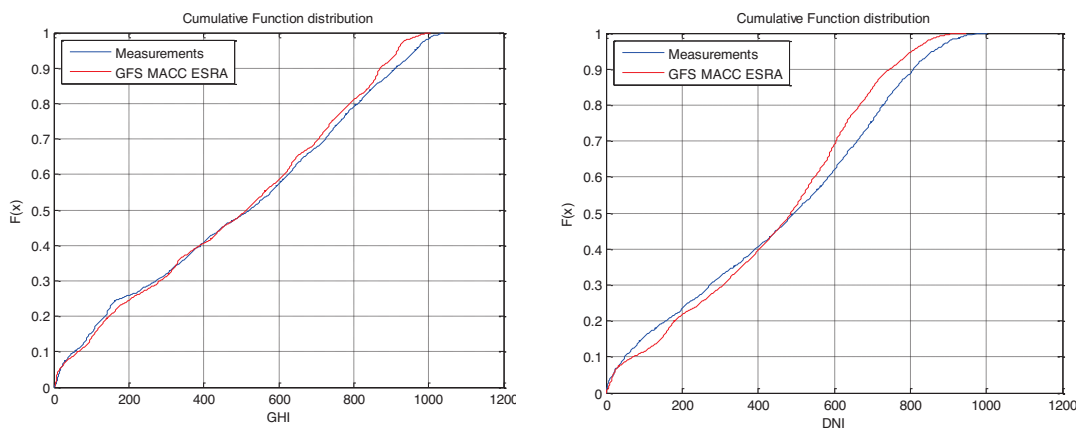


Fig. 8: Cumulative distribution function of hourly GHI and DNI GFS predictions and measurements in DOHA.

Finally, IKS, IKS %, Over, Over %, KSE and RIO for hourly GHI and DNI is presented:

Tab. 5: Errors for hourly GHI and DNI predictions in DOHA. ESRA clear sky model

Variable	IKS	IKS %	Over	Over %	KSE	RIO
GHI	15.98	102.95	4.77	30.73	133.69	76.34
DNI	34.68	232.24	21.48	143.86	376.10	212.26

As main conclusions for ESRA clear sky model, the results for GHI presents acceptable magnitude of the error regarding rRMSD. However, DNI is considerable high. One of the reasons could be that the model used to transform from GHI to DNI has been developed in the climatology of the USA and is not valid for the Gulf climate.

4.1 Results for SOLIS and REST2 clear sky models

In this section, we present the results for SOLIS and REST2 clear sky models. The next tables show MBD, RMSD, rMBD, rRMSD and R for GHI and DNI, for each day ahead of prediction from 1 to 6 and total errors of the models.

Tab. 6: Statistical errors of hourly GHI predictions validated with ground measurements. SOLIS clear sky model

Day ahead	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)	R
1	-71.39	131.61	-16.62	30.64	0.92
2	-67.78	133.61	-15.70	30.95	0.91
3	-67.57	135.86	-15.56	31.27	0.91
4	-63.39	139.55	-14.52	31.97	0.89
5	-62.91	137.31	-14.51	31.67	0.90
6	-63.63	139.75	-14.56	31.97	0.90
TOTAL	-66.10	136.32	-15.24	31.42	0.90

Tab. 7: Statistical errors of hourly DNI predictions validated with ground measurements. SOLIS clear sky model

Day ahead	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)	R
1	-62.43	243.11	-13.60	52.95	0.62
2	-51.26	249.78	-11.10	54.10	0.58
3	-49.12	246.56	-10.61	53.28	0.59
4	-37.23	250.16	-7.99	53.73	0.56
5	-42.23	242.86	-9.20	52.94	0.60
6	-42.47	249.98	-9.10	53.55	0.57
TOTAL	-47.45	247.10	-10.26	53.43	0.59

Tab. 8: Statistical errors of hourly GHI predictions validated with ground measurements. REST2 clear sky model

Day ahead	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)	R
1	-17.82	91.73	-4.15	21.36	0.94
2	-12.31	91.10	-2.85	21.10	0.94
3	-12.22	97.02	-2.81	22.33	0.94
4	-4.09	97.69	-0.93	22.35	0.94
5	-3.98	105.38	-0.92	24.25	0.92
6	-2.71	106.47	-0.62	24.33	0.92
TOTAL	-8.84	98.43	-2.04	22.67	0.93

Tab. 9: Statistical errors of hourly DNI predictions validated with ground measurements. REST2 clear sky model

Day ahead	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)	R
1	-71.92	240.23	-15.79	52.74	0.64
2	-61.99	247.43	-13.61	54.33	0.60
3	-61.74	242.45	-13.53	53.13	0.62
4	-47.69	245.26	-10.39	53.45	0.60
5	-52.53	239.70	-11.61	52.96	0.62
6	-52.22	246.62	-11.34	53.55	0.60
TOTAL	-57.99	243.64	-12.70	53.36	0.61

Conclusions for these clear sky models are similar as for ESRA model. Errors for SOLIS clear sky model are considerable high in term of rRMSD. The GFS-SOLIS models also present a substantial underestimation.

In the case of REST2 model, the errors regarding rRMSD are not so high for GHI. In the case of DNI, rRMSD is great and the value underestimated.

3.3 Data filtered with BSRN standards

In this section, we present the errors for ESRA clear sky model compared with ground measurements filtered with more restrictive BSRN quality filters. The next tables show MBD, RMSD, rMBD, rRMSD and R for GHI and DNI, for each day ahead of prediction from 1 to 6 and total errors of the models.

Tab. 10: Statistical errors of hourly GHI predictions validated with filtered ground measurements. ESRA clear sky model

Day ahead	MBD (Wh/m2)	RMSD(Wh/m2)	rMBD (%)	rRMSD (%)	R
1	-23.44	82.29	-3.62	12.71	0.94
2	-22.42	86.03	-3.46	13.27	0.94
3	-22.78	83.76	-3.51	12.91	0.94
4	-20.55	79.14	-3.15	12.15	0.95
5	-22.06	81.00	-3.39	12.43	0.94
6	-23.02	80.78	-3.52	12.37	0.94
TOTAL	-22.37	82.19	-3.44	12.64	0.94

Tab. 11: Statistical errors of hourly DNI predictions validated with ground measurements. ESRA clear sky model

Day ahead	MBD (Wh/m ²)	RMSD(Wh/m ²)	rMBD (%)	rRMSD (%)	R
1	-34.77	225.37	-6.08	39.44	0.41
2	-31.80	233.09	-5.57	40.82	0.36
3	-31.36	233.02	-5.49	40.76	0.37
4	-29.12	231.41	-5.05	40.13	0.34
5	-32.07	227.08	-5.59	39.60	0.37
6	-35.26	223.64	-6.15	39.00	0.40
TOTAL	-32.39	228.95	-5.65	39.96	0.38

The conclusions are similar as for ESRA model presented in section 4.1. There is a considerable decrease of the magnitude of rRMSD for GHI and DNI. It is because there are some reflections due to walls near to the station. This effect affects primarily to GHI. We have also detected a temporal deviation in the clock of around 2 to 5 minutes. When the deviation is corrected, rRMSD improves around 1% in the case of GHI.

5. Conclusion

The results show that the combination of the predictions of cloudiness from GFS and AOD and Water vapour from MACC and ESRA clear sky model provide good results for hourly GHI. In the case of DNI, the results are not so good. Bias can be removed using MOS, NN or correction with k_t and solar zenith angle matrix. We expect this model to have poor results regarding performance predicting variability of solar radiation due to coarse resolution of GFS model which is 3 hours. This last result has been presented recently in Task 46 of IEA. In the future, we are going to step up a mesoscale model customised for Qatar. The results in this paper will be used as a reference to improve prediction of solar radiation.

6. References

- Bachour, D., & Perez-Astudillo, D. (2014). Ground measurements of Global Horizontal Irradiation in Doha, Qatar. *Renewable energy*, 71(0), 32-36. doi: <http://dx.doi.org/10.1016/j.renene.2014.05.005>
- Espinar, B., Ramírez, L., Drews, A., Beyer, H. G., Zarzalejo, L. F., Polo, J., & Martín, L. (2009). Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations. *Solar Energy*, 83(1), 118-125. doi: DOI: 10.1016/j.solener.2008.07.009
- Gueymard, C. A. (2008). REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation $\Gamma\check{\circ}$ Validation with a benchmark dataset. *Solar Energy*, 82(3), 272-285.
- Ineichen, P. (2008). A broadband simplified version of the Solis clear sky model. *Solar Energy*, 82(8), 758-762.
- Long, C. N., & Dutton, E. G. (2002). BSRN Global Network recommended QC tests, V2.0 BSRN Technical Report.
- Pelland, S., Remund, J., Kleissl, J., Oozeki, T., & De Brabandere, K. (2013). Photovoltaic and Solar Forecasting: State of the Art.
- Perez-Astudillo, D., & Bachour, D. (2014). DNI, GHI and DHI Ground Measurements in Doha, Qatar. *Energy Procedia*, 49(0), 2398-2404. doi: <http://dx.doi.org/10.1016/j.egypro.2014.03.254>
- Perez, R., P.Ineichen, E.Maxwell, R.Seals, & A.Zelenka. (1992). Dynamic Global-to-Direct Irradiance Conversion Models. *ASHRAE Transactions-Research Series*, 354-369.
- Rigollier, C., Bauer, O., & Wald, L. (2000). On the clear sky model of the ESRA -- European Solar Radiation Atlas -- on the heliosat method. *Solar Energy*, 68(1), 33-48.
- Zarzalejo, L. F., Polo, J., Martín, L., Ramírez, L., & Espinar, B. (2009). A new statistical approach for deriving global solar radiation from satellite images. *Solar Energy*, 83(4), 480-484. doi: DOI: 10.1016/j.solener.2008.09.006