Forecasting Global Solar Insolation Using the Ensemble Kalman Filter Based Clearness Index Model

Pravat Kumar Ray, *Senior Member*, *IEEE*, Bidyadhar Subudhi, *Senior Member*, *IEEE*, Ghanim Putrus, Mousa Marzband, *Senior Member*, *IEEE*, and Zunaib Ali, *Student Member*, *IEEE*

***Abstract*—This paper describes a novel approach in developing a model for forecasting of global insolation on a horizontal plane. In the proposed forecasting model, constraints, such as latitude and whole precipitable water content in vertical column of that location, are used. These parameters can be easily measurable with a global positioning system (GPS). The earlier model was developed by using the above datasets generated from different locations in India. The model has been verified by calculating theoretical global insolation for different sites covering east, west, north, south and the central region with the measured values from the same locations. The model has also been validated on a region, from which data was not used during the development of the model. In the model, clearness index coefficients (K_{T}) are updated using the ensemble Kalman filter (EnKF) algorithm. The forecasting efficacies using the K_{T} model and EnKF algorithm have also been verified by comparing two popular algorithms, namely the recursive least square (RLS) and Kalman filter (KF) algorithms. The minimum mean absolute percentage error (MAPE), mean square error (MSE) and correlation coefficient (R) value obtained in global solar insolation estimations using EnKF in one of the locations are 2.4%, 0.0285 and 0.9866 respectively.**

***Index Terms*—Clearness index, ensemble Kalman filter, extra-terrestrial irradiance, forecasting, global solar insolation.**

I. Introduction

INTERMITTENT and variable nature of renewable generation poses new challenges to power system operations and controls. With high penetration of solar irradiance in a photovoltaic (PV) system, more acute problems, such as voltage and frequency fluctuations, occur resulting in additional requirements of ancillary generation and challenges in the electricity markets. Therefore, devising accurate forecasting methods has attracted more attention to researchers, for resolving the earlier issues with intermittency in PV power systems. Since the majority of generation units are associated with the output of the day ahead market, day ahead forecasting is important. A new model proposed [1] is to estimate solar insolation, through comparing it with numerical simulation and climatology measured data. Results show accurate evaluation of solar radiation for any day of a year with inputs, such as altitude, longitude and latitude. However, a main concern is that in order to collect the greatest solar energy, the orientation and inclination of the receiver should be varied, as per the variation of the solar declination angle. The diffuse ratio (k) vs. clearness index (K_{T}) was reviewed [2] for hourly, daily, monthly and yearly frequency regression models. From the regressor equation, the averaged diffuse irradiation values have been estimated from the averaged global irradiation values. A method [3] has been developed for calculation of total solar radiation from the evaluated direct and scattered solar radiation. In this method, the bias, which is the difference between the estimated and calculated values of diffuse and direct solar irradiance, act as an offset in the estimation of global solar insolation. It was also determined that temperature and relative humidity are the considered factors influencing the bias for direct and diffuse solar radiation. In addition, the assessment of global solar insolation is limited using this model. A rigorous review [4] has been made for the development and analysis of different models for estimating diffuse horizontal solar radiation during the day for different baroscopic places in China. The System Advisor Model (SAM) [5] is used to obtain the past solar generation information, and provide the input data from Solar anywhere. Using MATLAB software and the System Identification Toolbox, the model has been validated. The main concern here is that when the forecast horizon increases, both the persistence and Auto Regressive Moving Average (ARMA) models provide erroneous results. In [5] because of geographical differences, for prediction of irradiance, each location requires one model and the construction of the model requires two phases, i.e., for finding out orders and coefficients of the ARMA. For a country covering a vast geographical location, a single model cannot accurately forecast the solar power. A simple theoretical model has been developed in [6] for assessment of total solar insolation on a horizontal plane. The model is developed considering the latitude and the quantity of the entire precipitable water content in the vertical column of the desired site as input

Manuscript received August 21, 2021; revised January 2, 2022; accepted February 25, 2022. Date of online publication May 6, 2022; date of current version June 17, 2022. This work was supported in part by the DST, Govt. of India and British Council, UK vide no. DST/INT/UK/P-178/2017.

P. K. Ray is with Department of Electrical Engineering, National Institute of Technology Rourkela, Rourkela -769008, India.

B. Subudhi (corresponding author, e-mail: bidyadhar@iitgoa.ac.in; ORCID: https://orcid.org/0000-0003-4383-6783) is with School of Electrical Sciences, Indian Institute of Technology Goa, Goa 403401, India.

G. Putrus, M. Marzband and Z. Ali are with Department of Maths, Physics and Electrical Engineering, Northumbria University, UK.

DOI: 10.17775/CSEEJPES.2021.06230

constraints. However, the main lacunae in [6] is that the perfect estimation of the clearness index K_{T}, is not achievable, which plays a major role in the forecasting of global solar irradiance. A two-stage technique was introduced [7], where, at first, a statistical regulation of the lunar power has been carried out using a clear sky model. Then an adaptive linear time series model is used for forecasting the regulated solar power. The main disadvantage is that the model relies on the dependency of two types of inputs, such as for forecasts up to 2 h ahead solar power, while for longer horizons, Numerical Weather Predictions (NWP) are the most important inputs. A hybrid model based on the mesoscale meteorological Weather Research and Forecasting (WRF) model [8] and the clearness index-based Kalman filter were developed for day-ahead solar radiation prediction and were validated for two sites in Japan and China. This model has a drawback that the forecasting accuracy is much better in clear skies than in overcast conditions. Solar irradiance forecasting by employing artificial neural network (ANN) is suggested in [9]. In the ANN model, the Multilayer Perceptron MLP-model is found to be reasonably good to estimate the 24 hours based solar irradiance by applying the daily temperature of air and mean solar irradiance. The main issue lies in the complexity of the MLP-forecasting which requires huge computing time for attaining a decent performance. The authors propose the use of power output forecasting [10] that is based on 24 hours leading insolation forecasting for a PV system by utilizing weather testified statistics, fuzzy theory, and ANN. In this model, also more complexity is involved in training NN by output power information created on fuzzy theory and weather reported information. A comparative study [11] has been carried out among empirical models and ANN models and it was found that empirical models perform better compared to ordinary ANN models. However, when ordinary ANN models are coupled with the Genetic Algorithm, their performance improved. However, taking into account all the factors, such as skill, processing time and equipment, the empirical model is a better choice for evaluation of day-to-day total solar insolation in the climatic conditions of Iran. A feedforward backpropagation model [12] and its application in predicting the daily global solar radiation was presented. The proposed neural network runs with ten neurons and the log-sigmoid transfer function of the hidden. Fifteen numbers of different Geographical and meteorological parameters were used as input variables and daily global solar radiation as output variables. After accessing the neural network-based method [13], [14], experiments have been performed on machine learning based methods [15], such as random forest, gradient boosting, regression tree and many others, for the prediction of solar irradiance. In machine learning [15], the construction and study of systems can be learned from data sets, giving computers the ability to learn without being explicitly programmed. Machine learning models find relationships between input and output even without any possible representation. Before using machine learning models for forecasting problems, classification and data mining are of prime importance because one has to work with big datasets and the task of pre-processing, data protection and transmission can be taken care of by machine learning models, so a forecasting model using machine learning requires more skill and computational time as compared to other conventional methods. It was difficult to rank those methods as per their performance because of the data set diversification, forecasting horizon, time step and performance indicators. For improving prediction performance, hybrid models have been suggested. Another solar irradiance forecasting method [16] built on the Markov Switching Model has been established. The above technique has been applied to only remote locations where cloud based and other numerical prediction techniques may not be used.

A review of previousstudies in the field of forecasting in solar irradiance exhibit shortcomings (Sh) and weaknesses can be divided into the following categories:

1) Sh-1: Problem lies in the variation of the receiver as per the variation of the solar declination angle and modeling for the accuracy of the clearness index K_{T}, which plays a vital role in forecasting of global solar insolation ([1], [2], [6], [7]).
2) Sh-2: Temperature and humidity, affect the bias for direct and diffuse solar irradiance and finally affect the forecasting performance of global solar insolation ([3], [4]).
3) Sh-3: Model developed for solar irradiance forecasting is not uniform for any location in the world, i.e., it is location specific and also when the forecast horizon increases, accuracy in forecasting decreases ([5], [16]).
4) Sh-4: Main concern in neural network (NN) and hybrid model-based forecasting lies in training of NN requiring a huge amount of weather data, involvement of more skill, computation time and equipment in forecasting of solar irradiance ([8]–[12]).

Solar insolation forecasting is currently more important because of more inclusion of PV in conventional power systems. Statistical methods of forecasting provide good results for forecast horizons of up to 6 hours but for greater than a 6 hour forecast horizon, the numerical weather prediction (NWP) method is preferred. However, in this paper, a clearness index model based Ensemble Kalman Filtering is used for accurate forecasting of global solar insolation on a horizontal plane. The performance of the proposed method has also been compared with the recursive least square (RLS) and Kalman filtering (KF) methods.

The model developed in this study for forecasting direct insolation depends on two dominant measured constraints, such as latitude and precipitable water in the vertical column of the desired site. By using a global positioning system GPS receiver or a geographical map, the latitude of the desired site can be found. The source of insolation data for different locations within India is the Indian Meteorological Department, Pune i.e., https://imdpune.gov.in/ site. From the data related to daily relative humidity provided by the climatological department, precipitable water at these places can be found, or can be measured using radiosonde or GPS receivers. In the developed model for calculating clearness index (i.e K_{T}), Fourier coefficients are updated using an ensemble Kalman filter (EnKF) algorithm. So, a model, based on few input parameters will aid in fast and easy estimations of insolation for any place for any specified day.

The major contributions of this paper are summarized as follows:

1) Solar irradiance forecasting was performed based on two important parameters (latitude and precipitable water in the vertical column of the desired location). Environmental conditions and solar angles are also taken into account during the development of the model (tackling Sh1, Sh2).
2) In the developed model K_{T}, Fourier coefficients are updated using an ensemble Kalman filter (EnKF) algorithm, which provides more accuracy in estimating for K_{T} (tackling Sh1).
3) In the proposed model, a few input parameters are used for fast and easy estimation of insolation for any place for any specified day (tackling Sh3, Sh4).

The remainder of this paper is organized as follows: The development of the Ensemble Kalman Filter based clearness index model for global solar insolation forecasting is explained in Section II. The test and validation results are discussed in Section III. Section IV concludes this paper.

II. Ensemble Kalman Filter (EnKF) Based Clearness Index Model

The schematic for forecasting of global solar insolation is shown in Fig. 1. First K_{T} modeling is carried out and then the final estimate of K_{T} is performed using the EnKF/RLS/KF algorithm, Finally, global solar insolation value is forecasted.

The average monthly data for daily global solar insolation is usually accessible for various places in a given region. The data must be collected in a way that covers a large area of latitudes. This data ise then reduced to an index of average monthly daily clearance (i.e., K_{T}) by taking the global solar radiation to the estimated extra-terrestrial horizontal insolation, at a specified site. Extra-terrestrial horizontal insolation per day is an insolation on the horizontal surface without atmospheric influences. Extra-terrestrial horizontal insolation is expressed in terms of latitude and day of the year. It can be estimated for any site and for any day as described below.

The extra- terrestrial horizontal insolation, H_{o} is given by:

$$H_0 = \frac{24 I_0}{\pi}[\cos\phi\cdot\cos\delta\cdot\sin\omega_{\mathrm{sr}} + \omega_{\mathrm{sr}}\cdot\sin\phi\cdot\sin\delta]\ \mathrm{kWh/m^2} \tag{1}$$

where H_0 = Extra – terrestrial horizontal insolation in kWh/m^2

$$I_0 = I_{\mathrm{sc}}\left[1 + 0.033\left(\frac{360N}{365}\right)\right] \tag{2}$$

$I_0 =$ Extra-terrestrial insolation in kW/m^2
$I_{\mathrm{sc}} =$ Solar constant $= 1.367$ kW/m^2
$N =$ Day of the particular year
$\phi =$ Latitude of the location in degrees

$$\delta = 23.45\sin\left[\frac{2\pi(N-80)}{365}\right] \tag{3}$$

$\delta =$ Declination angle in degrees [13]
$\omega_{\mathrm{sr}} =$ Hour angle during Sunrise in radians

$$\omega_{\mathrm{sr}} = \cos^{-1}(-\tan\phi\cdot\tan\delta) \tag{4}$$

From the above equations, it can be observed that extra-terrestrial horizontal insolation can be determined for each location and day of the year as it depends on day of the year and latitude only. But in the above calculations, atmospheric effects had not been taken into account.

The effect of the atmosphere on insolation is determined by the clearness index, K_{T}. But, K_{T} is a stochastic parameter, that depends on time of year, climatic state, season and geographical position. So, in order to take into account, the effects of the atmosphere on the insolation of a place, a model of the clearness index is needed. For the modeling of K_{T}, the data on the insolation on a horizontal surface over a time interval covering each of the seasons and atmospheric situations, for a few locations, are to be measured. Utilizing Eq. (1), the H_0 can be estimated for desired sites, which provides measured global insolation. In the calculated value, atmospheric effect is not taken into account. After the calculation of H_0 for specific locations and the measured global horizontal insolation for the same locations, K_{T} for these locations is calculated. Then a graph is plotted by taking K_{T} vs month of the year as shown in Fig. 2. It is found from the graph that the variation of K_{T} is a periodic function having a periodicity of one-year. Thus, for the modeling of K_{T}, the Fourier series is regarded as a suitable curve fitting method.

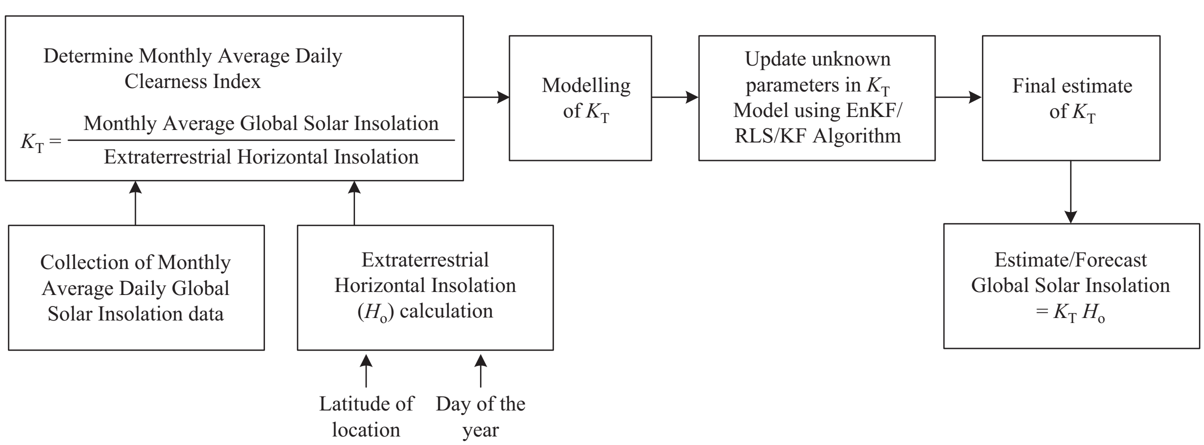

Fig. 1. Schematic for forecasting of global solar insolation.

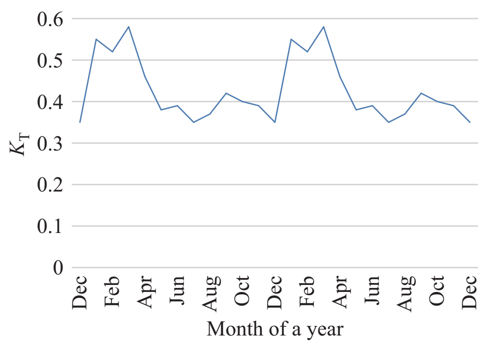

Fig. 2. Variation of K_{T} throughout year.

K_{T} can be represented by the Fourier series:

$$K_{\mathrm{T}} = f(x, w, t) + e \tag{5}$$

$$\begin{aligned} f(x, w, t) = {} & A_1 + A_2 \sin t + A_3 \sin 2t + A_4 \sin 3t \\ & + A_5 \cos t + A_6 \cos 2t + A_7 \cos 3t \end{aligned} \tag{6}$$

Trigonometric terms (t) are functions of the day of the particular year (N)

$$x = \phi - 35 \tag{7}$$

$\phi =$ Latitude in degrees
$w =$ Total precipitable water vapor in gm/cm^2

$$t = (2\pi/365)(N - 80) \tag{8}$$

$e =$ error

$$e = K_{\mathrm{T}} - f(x, w, t) \tag{9}$$

x has been updated based on the best fit of the available data. Since all the data used have been collected only from India, we have found that an offset value of 35-degree latitude gives the best fit for this sub-continent. In this way, the function of x given in (7) has been determined.

$A_1, A_2, \cdots, A_7$: Functions of (ϕ) and (w)

Fourier coefficients are evaluated from the succeeding equation:

$$A_i = a_{i1} + a_{i2}x + a_{i3}x^2 + a_{i4}w + a_{i5}w^2 \tag{10}$$

Now replacing A_i in (6) for $f(x, w, t)$

$$\begin{aligned} f(x, w, t) = {} & (a_{11} + a_{12}x + a_{13}x^2 + a_{14}w + a_{15}w^2) \\ & + (a_{21} + a_{22}x + a_{23}x^2 + a_{24}w + a_{25}w^2) \sin t \\ & + (a_{31} + a_{32}x + a_{33}x^2 + a_{34}w + a_{35}w^2) \sin 2t \\ & + (a_{41} + a_{42}x + a_{43}x^2 + a_{44}w + a_{45}w^2) \sin 3t \\ & + (a_{51} + a_{52}x + a_{53}x^2 + a_{54}w + a_{55}w^2) \cos t \\ & + (a_{61} + a_{62}x + a_{63}x^2 + a_{64}w + a_{65}w^2) \cos 2t \\ & + (a_{71} + a_{72}x + a_{73}x^2 + a_{74}w + a_{75}w^2) \cos 3t \end{aligned}$$

$$f(x, w, t) = H(x, w, t)\theta \tag{11}$$

$H(x, w, t)$: system structure matrix
θ: vector of unknown parameter

Here the unknown parameters are estimated using the Ensemble Kalman Filter (EnKF) [18] algorithm. The EnKF is a Monte Carlo approximation of the conventional Kalman Filter (KF). Rather than developing the covariance matrix of probability density function of the state vector, it uses the distribution characterized by a sample of the state vector x, called an ensemble.

θ is updated using Ensemble Kalman Filtering (EnKF) as given below:

Take the ensemble matrix θ and data matrix as f.

Ensemble mean and covariance are:

$$E(\theta) = \frac{1}{Q} \sum_{n=1}^{Q} \theta_n \tag{12}$$

Q: No. of Ensembles

$$C = \frac{GG^{\mathrm{T}}}{Q - 1} \tag{13}$$

where

$$G = \theta - E(\theta) \tag{14}$$

$$\hat{\theta}(k) = \theta(k-1) + CH^{\mathrm{T}}(HCH^{\mathrm{T}} + R)^{-1}(f - H\hat{\theta}(k-1)) \tag{15}$$

After finding the updated values of θ, i.e., $a_{11}, a_{12}, \cdots, a_{75}$ These coefficients are utilized to find the Fourier coefficients

$$A_1, A_2, \cdots, A_7$$

which leads to the following model for K_{T}

$$\begin{aligned} K_{\mathrm{T}} = {} & A_1 + A_2 \sin t + A_3 \sin 2t + A_4 \sin 3t + A_5 \cos t \\ & + A_6 \cos 2t + A_7 \cos 3t \end{aligned} \tag{16}$$

Now for a particular location and for a particular day of the year, K_{T} can be estimated by applying the above Equation.

$$H_{tf} = K_{\mathrm{T}} \cdot H_0 \tag{17}$$

$H_{tf} =$ Estimated/forecasted value of Global solar insolation on a horizontal surface at a particular location for a particular day.

The EnKF algorithm, for estimation of daily global solar insolation is given as the flow chart in Fig. 3.

III. Results and Discussions

The model validation has been carried out in a two-step method. In the first stage, measured daily median values of global insolation data for 12 different locations covering a large range of latitude across India was collected. The model was developed using measured data from 12 different locations in India covering the whole area of the country over a period of 5 years. Based on this data, a model for K_{T} was developed. From the model, theoretical global insolation was calculated and then compared with the measured value for validation.

In the second stage, the conceptual global insolation was estimated for another location in India, whose latitude lies inside the range of latitudes considered for development of the model. The estimated and measured insolation curves are studied for validation of the proposed model. Table I shows the measured and forecasted values of global solar insolation at five different locations covering north, south, east, west and the central parts in India. In the table, Ins stands

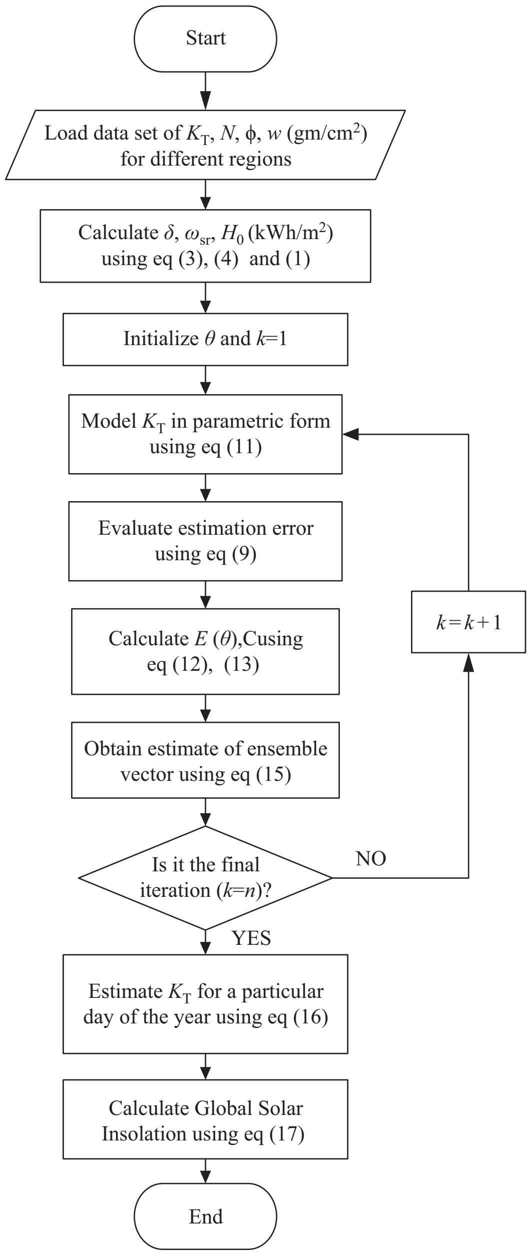

Fig. 3. Flow chart for estimation of global solar insolation using the EnKF algorithm.

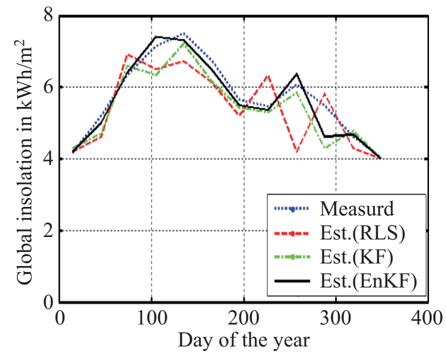

Fig. 4. Measured and estimated global solar irradiance for region A (North).

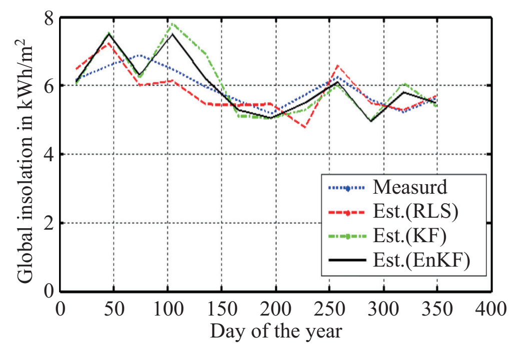

Fig. 5. Measured and estimated global solar irradiance for region B (South).

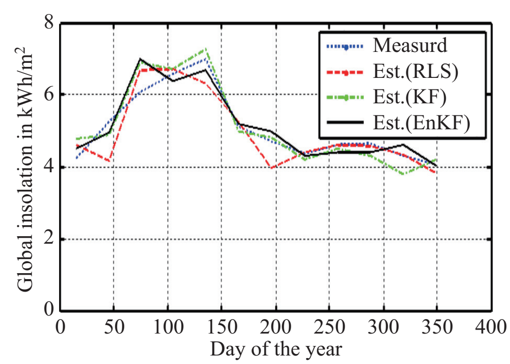

Fig. 6. Measured and estimated global solar irradiance for region C (East).

for Insolation, H_{tm} is the measured value of global solar insolation in kWh/m^2 and H_{tf} is the forecasted value of global solar insolation in kWh/m^2.

Figure 4 shows the measured and calculated global solar irradiance at region A, which is located in the northern part of India. It is found that the estimation of global solar irradiance, using RLS is comparatively better than using KF. But estimating using EnKF outperforms RLS and KF in estimation of global solar insolation. Fig. 5 shows the comparison of estimation accuracy using all the three discussed algorithms at region B, located in southern part of India and it was found that also in this case, estimation accuracy is greater in the case of EnKF estimation as compared to the RLS and KF algorithms. Fig. 6 shows the comparative estimation of global solar irradiance using the three algorithms in region C, which is in the eastern part of India. In this region the measured and estimated irradiances almost match with each other, with more accuracy in the estimation using the EnKF algorithm.

For verification of the proposed EnKF based estimation of global solar irradiance, for the rest of the two regions, Figs. 7 and 8 show the comparative estimation of global solar irradiance in regions D and E, located in the western and central parts of India. In both the Figs., the estimation performance of global solar irradiance using EnKF is comparatively better as contrasted to RLS and KF.

In order to measure the accuracy of the estimation, the errors between the estimated values and measure data are examined here. In this paper, mean absolute percentage error (MAPE), mean square error (MSE) and correlation coefficient ® as defined in Eqs. (18), (19) and (20) respectively, are used as the error indices to verify the forecasting technique.

TABLE I
MEASURED AND FORECASTED VALUES OF GLOBAL SOLAR INSOLATION

Loc	Ins	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
A	H_{tm}	4.15	5.19	6.34	7.13	7.51	6.76	5.66	5.45	6.07	5.5	4.6	4.02
	H_{tc} (RLS)	4.2	4.61	6.92	6.48	6.72	6.15	5.2	6.33	6.33	5.25	4.3	4.01
	H_{tc} (KF)	4.3	4.7	6.6	6.32	7.2	6.16	5.42	5.31	5.86	4.3	4.77	4.00
	H_{tc} (EnKF)	4.2	5.0	6.4	7.4	7.3	6.5	5.5	5.35	6.36	5.63	4.68	4.00
B	H_{tm}	6.17	6.58	6.89	6.48	5.58	5.58	5.2	5.75	6.23	5.59	5.24	5.62
	H_{tc} (RLS)	6.48	7.23	6	6.15	5.46	5.44	5.47	4.8	6.6	5.49	5.3	5.69
	H_{tc} (KF)	6.05	7.54	6.2	7.8	6.92	5.12	5.04	5.3	6.0	4.98	6.04	5.4
	H_{tc} (EnKF)	6.1	7.5	6.3	6.85	6.2	5.3	5.06	5.5	6.1	5.53	5.8	5.5
C	H_{tm}	4.24	5.26	6.09	6.59	7.01	5.14	4.71	4.36	4.62	4.64	4.3	4.07
	H_{tc} (RLS)	4.6	4.18	6.7	6.74	6.3	5.27	3.96	4.4	4.62	4.59	4.35	3.84
	H_{tc} (KF)	4.77	4.89	6.91	6.73	7.26	5.0	4.83	4.2	4.5	4.3	3.8	4.2
	H_{tc} (EnKF)	4.5	4.94	7.0	6.4	6.7	5.2	5.0	4.3	4.4	4.4	4.6	4.02
D	H_{tm}	5.08	5.77	6.58	7.13	7.42	5.8	4.17	4.09	5.2	5.66	5.23	4.83
	H_{tc} (RLS)	4.9	5.5	5.83	6.16	6.63	5.62	4.32	4.88	4.81	4.92	5.0	4.75
	H_{tc} (KF)	4.78	5.86	5.15	6.4	6.5	5.27	4.54	4.9	5.4	5.63	5.32	4.35
	H_{tc} (EnKF)	5.2	5.6	6.78	7.13	7.48	5.56	4.86	4.8	5.5	5.33	5.1	4.5
E	H_{tm}	5.03	5.78	6.43	6.97	7.12	6.01	4.41	4.22	5.38	5.87	5.3	4.85
	H_{tc} (RLS)	4.32	5.84	6.07	7.20	6.74	6.64	4.51	3.53	5.28	5.87	4.78	5.02
	H_{tc} (KF)	5.24	6.2	6.8	7.5	6.86	6.4	4.43	5.21	4.32	5.6	5.32	5.01
	H_{tc} (EnKF)	5.1	6.4	6.24	7.3	7.1	5.85	4.5	4.3	5.0	5.3	5.4	5.00

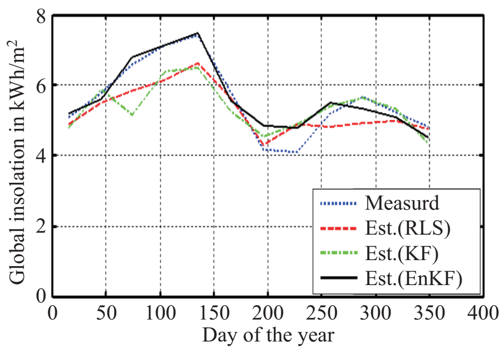

Fig. 7. Measured and estimated global solar irradiance for region D (West).

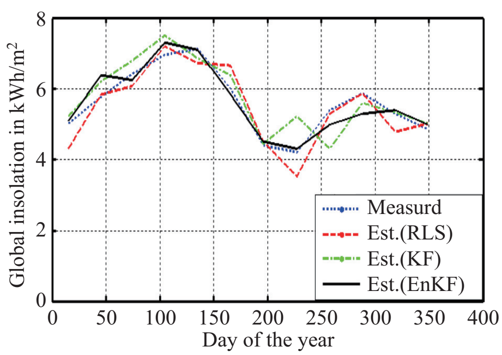

Fig. 8. Measured and estimated global solar irradiance for region E (Central).

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{H_{tm}(t) - H_{tf}(t)}{H_{tm}}\right| \tag{18}$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(H_{tm}(t) - H_{tf}(t))^2 \tag{19}$$

$$R = \frac{n\sum H_{tm}H_{tf} - \left(\sum H_{tm}\right)\left(\sum H_{tf}\right)}{\sqrt{\left[n\sum H_{tm}^2 - \left(\sum H_{tm}\right)^2\right]\left[n\sum H_{tf}^2 - \left(\sum H_{tf}\right)^2\right]}} \tag{20}$$

where n is the number of measurements.

From Table II, it is seen that for all the cities, the MAPE and MSE using EnKF method is minimal as compared to the other two methods. It can also be seen from Table II that the correlation co-efficie®(R) for most cases of forecasting lies between 0.7–1, which shows strong association between measured and forecasted values. In the majority of cases, for R values in the EnKF estimation, they are very close to 1, which shows that the EnKF estimation has the strongest association between measured and estimated values as compared to the other two algorithms

TABLE II
COMPARISION OF MAPE, MSE AND R OF DIFFERENT PLACES WITH DIFFERENT METHODS

City	Methods	MAPE (%)	MSE	R
A	RLS	7.5	0.2750	0.8868
	KF	6.4	0.2527	0.9228
	EnKF	2.4	0.0285	0.9866
B	RLS	6.6	0.2153	0.7082
	KF	10.3	0.5419	0.6118
	EnKF	5.6	0.1858	0.7661
C	RLS	6.5	0.2361	0.8841
	KF	5.9	0.1357	0.9528
	EnKF	5	0.1183	0.9361
D	RLS	8	0.3059	0.9359
	KF	8.8	0.4063	0.8317
	EnKF	5.6	0.0893	0.9418
E	RLS	6.1	0.168	0.9336
	KF	7.4	0.255	0.8634
	EnKF	3.7	0.0897	0.9462

The MAPE in the estimation of global solar insolation using different methods, is shown in Fig. 9, and it is seen that for all the cities, the MAPE error using the EnKF method is minimum as compared to other methods. Minimum MAPE using EnKF method is found as 2.4% for city A and maximum MAPE value using EnKF is 5.6% for both city B and D, which can be accepted for getting accuracy in forecasting. Fig. 10 shows the comparison of MSE in the assessment of global solar insolation applying various methods. It has been found that MSE value in case of EnKF estimation for all the five

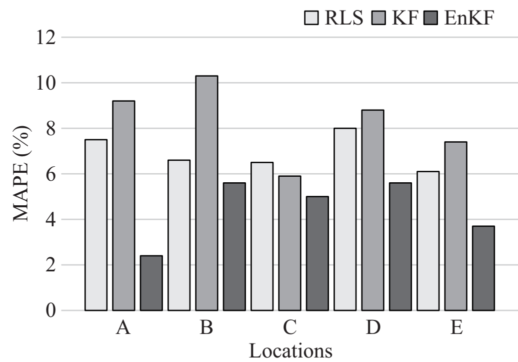

Fig. 9. Comparison of MAPE of different places with different methods.

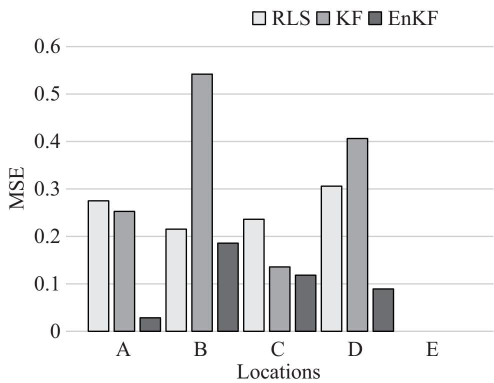

Fig. 10. Comparison of MSE of different places with different methods.

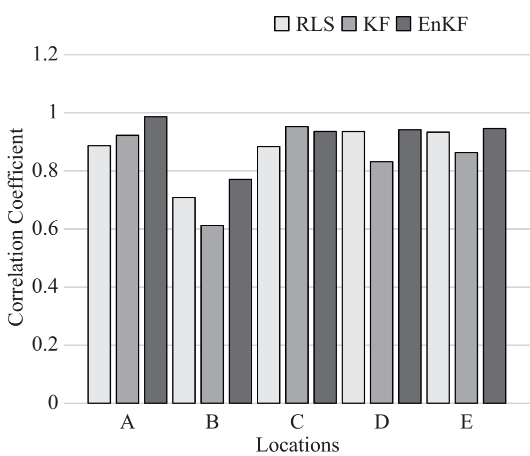

Fig. 11. Comparison of Correlation Coefficient ® of different places with different methods.

cities is minimal as compared to RLS and KF. The minimum value of MSE is 0.0285 in the EnKF estimation for city A. The maximum value of MSE is 0.1858 for city B, which can be accepted for estimation accuracy. Fig. 11 shows a comparison of correlation coefficient ® (R) values in the global solar insolation assessment using various methods and it is seen that there is the strongest association between measured and estimated value in each city using the EnKF algorithm as compared to the other two methods.

In each case of estimation, it is seen that the estimated and measured values of insolation are following very closely and the estimation using EnKF is giving more accuracy as compared to the others.

Table III shows a performance comparison of the proposed EnKF algorithm with the methods used in [13] and [14]. On comparing R^2 values using the EnKF algorithm with the methods employed in [13] for day ahead forecasting, it was found that the proposed EnKF forecasting outperforms the algorithms used in [13] in most of the cases of forecasting with a maximum R^2 value of 0.9733 in city A. RMSE performance of forecasting using the proposed EnKF algorithm has also been compared with [13] and [14] and it has been found that EnKF outperforms the methods used in the above two papers, with a minimized RMSE value of 0.168 in city A. Comparison of MAPE using EnKF with [14] shows that it has a very lesser MAPE error in all cases of forecasting as compared to [14] except in Jodhpur. The overall comparison from Table II shows that forecasting using the proposed EnKF method outperforms the methods used in [13] and [14].

TABLE III
PERFORMANCE COMPARISON OF PROPOSED ENKF ALGORITHM WITH ALGORITHMS [13], [14]

Papers	Locations	Model/Algorithms	R^2	RMSE	MAPE
[13]	Algeria	AR	0.6706	1.29	–
		NAR	0.6651	1.28	–
		SVR	0.6753	1.26	–
		RF	0.6759	1.25	–
		PER	0.5217	1.54	–
[13]	Ghardia	AR	0.6989	1.08	–
		NAR	0.6645	1.10	–
		SVR	0.6849	1.08	–
		RF	0.7001	1.07	–
		PER	0.5738	1.26	–
[14]	Jaisalmar	Multi step	–	0.445	6.193
	Barmer		–	0.365	7.977
	Bikaner		–	0.333	6.968
	Jodhpur		–	0.367	0.474
This paper	City A	EnKF	0.9733	0.168	2.4
	City B		0.5869	0.431	5.6
	City C		0.8762	0.344	5
	City D		0.8869	0.298	5.6
	City E		0.8952	0.299	3.7

However, since the model has been developed on using the data from the above five locations out of all other locations, the accuracy in the above results requires itto be validated again. For this, the measured values and estimated values of insolation for a location were chosen, whose data was not used during the model development. Place Z, is one such location whose data was not considered during model development and its latitude lies within the range of latitudes for the model. The measured and estimated insolation values of place Z for a year, using three estimation algorithms, are given in Table IV. Here also the insolation estimation performance of EnKF is better when compared to the other two methods. We determined from Fig. 12 that the measured values and estimated values are very close to each other in the case of the estimation using EnKF. Table V shows the MAPE, MSE and R values comparison using different methods for region Z and it is also found that, EnKF outperforms the other two methods. Fig. 13 shows the MAPE comparison using different methods for city Z and it is also found that, in the case of the EnKF estimation, MAPE is minimal, i.e., 4% as compared to the other two methods.

TABLE IV
CITY Z (WHOSE DATA HAS NOT BEEN INCLUDED IN MODEL)

Place	Inso	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Z	Htm	5.74	6.41	6.83	7.03	7.00	5.04	3.93	5.05	5.81	5.91	5.85	5.58
	Htc (RLS)	5.56	6.75	6.35	6.5	6.67	5.7	3.44	5.7	5.5	6.7	6.76	4.48
	Htc (KF)	6	6.64	6.2	6.8	6.4	5	3.85	5.72	4.5	6.5	5.34	4.43
	Htc(EnKF)	5.6	6.8	6.9	7.0	6.84	5	3.85	5.5	5.25	6.1	6.06	5.09

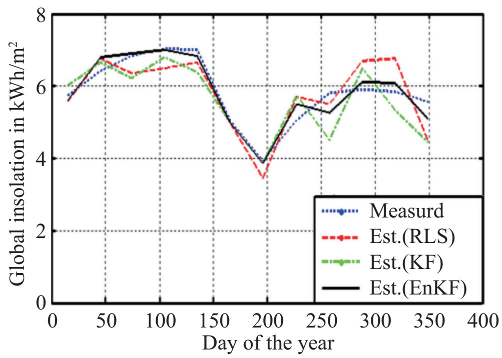

Fig. 12. Measured and estimated global solar irradiance for region Z (Data not included during model development).

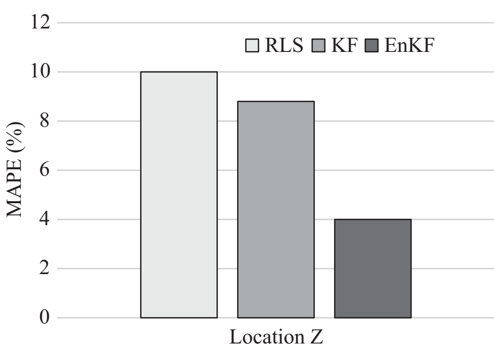

Fig. 13. Comparison of MAPE of city Z with different methods.

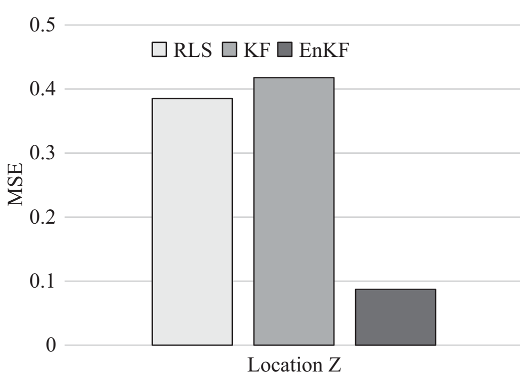

Fig. 14. Comparison of MSE of city Z with different methods.

Fig. 14 shows a comparison of MSE values in the global solar insolation estimation for city Z and it is seen that also in this case, MSE is minimal (0.0872) using the EnKF estimation. Fig. 15 shows a comparison of correlation coefficient ® (R) of city Z with different methods and it is found that the R value in the case of the EnKF estimation is 0.9486, which shows a stronger association of measured and estimated values

TABLE V
COMPARISION OF MAPE, MSE AND R OF CITY Z WITH DIFFERENT METHODS

City	Methods	MAPE (%)	MSE	R
Z	RLS	10	0.3853	0.7862
	KF	8.8	0.4178	0.7780
	EnKF	4	0.0872	0.9486

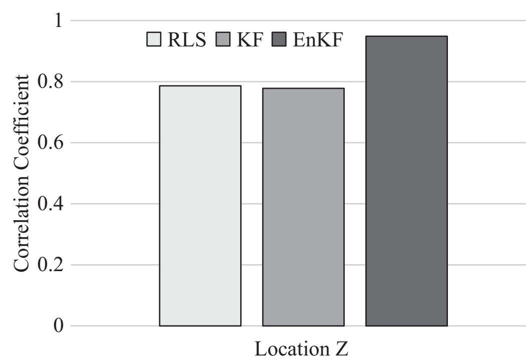

Fig. 15. Comparison of Correlation Coefficient ® of city Z with different methods.

as compared to the other two methods. The above results definitely validate not only the developed model but also the used methodologies.

IV. CONCLUSION

A method for the development of the solar insolation model for certain regions is explained. As the forecasting of global solar insolation primarily depends on the K_{T} value and using the EnKF algorithm, more accuracy in an updated K_{T} value was obtained, so the accuracy in forecasting is more in the estimation using this algorithm. From the obtained results, it is found that the forecasted value of insolation using the EnKF method is more closely matched with the measured values (minimum MAPE of 2.4 %, MSE of 0.0285 and R of 0.9866), and this method can be used for any region in the world for forecasting solar insolation. For getting more accuracy in forecasting, measured data should be taken covering a large range of latitudes for the selected region. Also data available should be as accurate as possible to obtain a low error model, in order to calculate global solar insolation for a specific day and for a specific location. Currently, with the help of satellites, data for certain regions can be obtained with good accuracy. With more accurate data, using the proposed EnKF algorithm, more accurate Fourier's Coefficients can be obtained to forecast the global solar insolation. Since accuracy on forecasting of global solar insolation depends on both K_{T} and H_0 values, proper modeling for obtaining K_{T} and an accurate calculation for the H_0 value of the desired location are very essential, so more accurate modeling for K_{T} and a less input parameters model for the H_0 calculation can be used in the future.

REFERENCES

[1] Y. El Mghouchi, A. El Bouardi, Z. Choulli, and T. Ajzoul, "New model to estimate and evaluate the solar radiation," *International Journal of Sustainable Built Environment*, vol. 3, no. 2, pp. 225–234, Dec. 2014.

[2] S. E. Berrizbeitia, E. J. Gago, and T. Muneer, "Empirical models for the estimation of solar sky-diffuse radiation. A review and experimental analysis," *Energies*, vol. 13, no. 3, pp. 1–23, Feb. 2020.

[3] M. Z. Xiao, Z. B. Yu, and Y. Z. Cui, "Evaluation and estimation of daily global solar radiation from the estimated direct and diffuse solar radiation," *Theoretical and Applied Climatology*, vol. 140, no. 3, pp. 983–992, Feb. 2020.

[4] J. L. Fan, L. F. Wu, F. C. Zhang, H. J. Cai, X. Ma, and H. Bai, "Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of China," *Renewable and Sustainable Energy Reviews*, vol. 105, pp. 168–186, May 2019.

[5] R. Huang, T. N. Huang, R. Gadh, and N. Li, "Solar generation prediction using the ARMA model in a laboratory-level micro-grid," in *Proceedings of the 2012 IEEE Third International Conference on Smart Grid Communication*, 2012, pp. 528–533.

[6] R. Kumar and L. Umanand, "Estimation of global radiation using clearness index model for sizing photovoltaic system," *Renewable Energy*, vol. 30, no. 15, pp. 2221–2233, Dec. 2005.

[7] P. Bacher, H. Madsen, and H. A. Nielsen, "Online short-term solar power forecasting," *Solar Energy*, vol. 83, no. 10, pp. 1772–1783, Oct. 2009.

[8] Y. Z. Che, L. W. Chen, J. F. Zheng, L. Yuan, and F. Xiao, " A novel hybrid model of WRF and clearness index-based Kalman filter for day-ahead solar radiation forecasting," *Applied Sciences*, vol. 9, no. 19, pp. 3967, Sep. 2019.

[9] A. Mellit and A. M. Pavan, "A 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at Trieste, Italy," *Solar Energy*, vol. 84, no. 5, pp. 807–821, May 2010.

[10] A. Yona, T. Senjyu, T. Funabashi, and C. H. Kim, "Determination method of insolation prediction with fuzzy and applying neural network for long-term ahead PV power output correction," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 2, pp. 527–533, Apr. 2013.

[11] B. Jahani and B. Mohammadi, "A comparison between the application of empirical and ANN methods for estimation of daily global solar radiation in Iran," *Theoretical and Applied Climatology*, vol. 137, pp. 1257–1269, Oct. 2019.

[12] B. Belmahdi, M. Louzazni, M. Akour, D. T. Cotfas, P. A. Cotfas, and A. El Bouardi, "Long-term global solar radiation prediction in 25 cities in morocco using the FFNN-BP method," *Frontiers in Energy Research*, vol. 9, no. 733842, Sep. 2021.

[13] A. Takilalte, S. Harrouni, and J. Mora, "Forecasting global solar irradiance for various resolutions using time series models - case study: Algeria," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 44, no. 1, pp. 1–20, Aug. 2019, doi: 10.1080/15567036.2019.1649756.

[14] D. Chandola, H. Gupta, V. A. Tikkiwala, and M. K. Bohra, "Multi-step ahead forecasting of global solar radiation for arid zones using deep learning," *Procedia Computer Science*, vol. 167, pp. 626–635, Jan. 2020.

[15] C. Voyant, G. Notton, S. Kalogirou, M. L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, "Machine learning methods for solar radiation forecasting: a review," *Renewable Energy*, vol. 105, pp. 569–582, May 2017.

[16] A. Shakya, S. Michael, C. Saunders, D. Armstrong, P. Pandey, S. Chalise, and R. Tonkoski, "Solar irradiance forecasting in remote microgrids using markov switching model," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 895–905, Jul. 2017.

[17] R. H. B. Exell, "A mathematical model for solar radiation in south-east Asia (Thailand)," *Solar Energy*, vol. 26, no. 2, pp. 161–168, Jan. 1981.

[18] P. K. Ray and B. Subudhi, "Ensemble Kalman filter based power system harmonics estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 12, pp. 3216–3224, Dec. 2012.

Pravat Kumar Ray (SM'18) received the B.E. degree from Indira Gandhi Institute of Technology, Sarang, Odisha, India, in 2000, the M.E. degree from Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India, in 2003, and the Ph.D. degree from National Institute of Technology (NIT) Rourkela, India, in 2011, all in Electrical Engineering. He is currently an Associate Professor with the Department of Electrical Engineering, NIT Rourkela. He was also a Postdoctoral Fellow at Nanyang Technological University, Singapore, during January 2016 to June 2017. He is having eighteen years of experience in teaching and research. He has supervised 10 Ph.D theses. Currently, he is handling 4 research projects at NIT Rourkela. He has published more than 100 papers in international journals and conferences. His research interests include signal processing and soft computing applications to power system, power quality, and grid integration of renewable energy systems.

Bidyadhar Subudhi (SM'08) is currently working as a Professor, School of Electrical Sciences and Dean (Research & Development) in Indian Institute of Technology (IIT) Goa, India. Prior to this, he was working as a Professor in the Dept. of Electrical Engg. in NIT Rourkela, India. He also served as a Visiting Professor at the Univ. of Saskatchewan, Canada and in Asian Institute of Technology, Bangkok in 2011 and 2013 respectively. He was a recipient of the Prestigious Samanta Chandra Sekhar Award of the Odisha Bigyan Academy, Govt. of Odisha for his contribution to science and technology in the year 2013. He is a Fellow of the Indian National Academy of Engineering and Fellow of IET (UK). His research interests include system & control, control of PV and microgrid systems, and wide area control of power system.

Ghanim Putrus is Professor of Electrical Power Engineering at the Department of Mathematics, Physics and Electrical Engineering, Northumbria University, Newcastle upon Tyne, UK. He has over 30 years of research experience with over 200 publications, including one patent and has given several invited talks at national and international events. He has led several research projects and has often provided consultancy for industry. He has led several research projects on the integration of renewable energy generation and electric transport into the electricity grid. His main research interests are application of power electronics in power systems and integration of electric vehicles and renewable energy sources into power distribution networks (smart grids).

Mousa Marzband (SM'17) received the Ph.D. degree in Electrical Engineering from the Department of Electrical Engineering, Polytechnic University of Catalonia, Barcelona, Spain, in 2014. He then joined the University of Manchester, UK as a Post-Doctoral Research Fellow and the University College Cork, Cork, Ireland, as a Senior Researcher. He is currently a Senior Lecture (Associate professor) with the Department of Maths, Physics and Electrical Engineering, Northumbria University, Newcastle upon Tyne, UK. His research findings have been published in two books and more than 80 top journals and over 40 proceedings of international conferences. His research interests include operation and control strategies of DGs, mathematical modelling and control of optimal energy management system within multi-energy carrier systems, and cooperative and non-cooperative game theory applications in energy market.

Zunaib Ali (S'15) received the B.Sc. degree in Electronic Engineering from the University of Engineering and Technology, Peshawar, Pakistan, in 2013, the M.Sc. degree in Electrical Engineering from COMSATS University Islamabad, Abbottabad, Pakistan, in 2015, and the Ph.D. degree in Electrical Engineering from Frederick University, Nicosia, Cyprus, in 2018. From 2018 to 2019, he was a Research Assistant with Real Time Power Laboratory, Queen Mary University of London, London, U.K., for Innovative U.K. project. He is currently a Research Fellow with Northumbria University, Newcastle upon Tyne, U.K. His research interests include grid synchronization and control methods for distributed renewable energy systems and how such systems can be strategically utilized in applications related to power electronic converters and consequently in smart grids and microgrids.