Grey-box Identification for Photovoltaic Power Systems via Particle-Swarm Algorithm

Naji Al-Messabi, Cindy Goh, Yun Li

School of Engineering, University of Glasgow, Glasgow G12 8 QQ, U.K.,

(n.al-messabi.1@research.gla.ac.uk, Cindy.Goh@glasgow.ac.uk, Yun.Li@glasgow.ac.uk)

Abstract — **Amongst renewable generators, photovoltaics (PV) are becoming more popular as the appropriate low cost solution to meet increasing energy demands. However, the integration of renewable energy sources to the electricity grid possesses many challenges. The intermittency of these non-conventional sources often requires accurate forecast, planning and optimal management. Many attempts have been made to tackle these challenges; nonetheless, existing methods fail to accurately capture the underlying characteristics of the system. There exists scope to improve present PV yield forecasting models and methods. This paper explores the use of apriori knowledge of PV systems to build clear box models and identify uncertain parameters via heuristic algorithms. The model is further enhanced by incorporating black box models to account for unmodeled uncertainties in a novel grey-box forecasting and modeling of PV systems.**

I. INTRODUCTION

Photovoltaic (PV) energy is now positioned amongst the top three new power generation means installed in Europe and is expected to remain so [1]. Power from PV sources provides a number of benefits over other renewable energy sources (RES). It can be supplied locally to loads, reducing the cost of transmission lines and associated power losses. Furthermore, advances in technology and large scale manufacturing have led to the decline in PV cost at a steady rate [2]. Despite a high capital setup cost, the operation and maintenance costs of PV are almost zero [3].

Nonetheless, like other RES, PV sources pose a number of integration challenges such as the impact on voltage profile [4],[7], impact on operational costs of the grid [5], regulation and load-following requirements [6]. Advance knowledge of the expected yield from PV sources will help tackle these challenges, to allow for proper planning of available generation sources and provide insights into the impact of PVs on the power network. However, the forecasting task requires non-primitive techniques, as power yield from PVs is intermittent in nature. The intermittent and non-linear characteristics of PV data is due to an interplay of various factors such as the variability in sunrise and the amount of sunshine, sudden changes in atmospheric conditions, cloud movements and dust [8]. The PV power data can thus be viewed as consisting of two parts: the deterministic and the stochastic parts. The former represents the mathematical equations of irradiance that depend on location, sun's position, and equations of PV cells, whilst the latter represents the

sudden atmospheric changes such as dust, clouds, and wind blow.

Various mathematical models that capture physics of PVs or clear-box models are possible but are inaccurate or impractical for large systems [9]. However, clear-box models possess various strengths such as that their structures are of physical meaning and usually have fewer parameters to estimate [10].

On the other hand, data-driven or black-box models based on statistics or artificial-intelligence are popular methods as they are simple and easy to use. Dynamic Neural Networks (DNNs) such as the 'Focused Time-Delay Neural Networks' (FTDNN) and the 'Distributed Time-Delay Neural Networks' (DTDNN) [11] have been studied for PV forecasting. These methods can handle nonlinear time-series data that are dynamic in nature. However, black-box models require good data for proper modeling – both quality and quantity. It is also difficult to design due to large number of parameters and lack of a systematic way to arrive at an optimal structure.

This paper will therefore focus on the identification of proper grey box photovoltaic models. Section II provides an overview of related works in clear-box models for solar PV. In Section III, uncertain parameters in the clear-box model are identified and optimised using Particle Swarm Optimisation (PSO). The model is then extended into a grey-box model to account for unknown effects during forecast. Two grey-box models are developed and the results and observations are discussed in Section IV. Section V concludes the paper.

II. SOLAR PV POWER MODELS

A. General overview

There exist various forecasting models proposed for PV systems [9]-[30]. The simplest ones are naïve or persistence models where the next power value is assumed to be same as the previous step. Such models are usually taken as reference models in forecasting studies [13].

The other approach to forecast PV power is to model solar data using statistical methods. Regression models can be used where power value is expressed as a regression of previous power values, irradiance, and temperature [14]. Statistical approaches adopt classical time-series forecasting methods that assume the data to be stationary. Auto-regressive (AR), AR with exogenous input (ARX), and AR with integrated moving average (ARIMA) [13] are some of the famous

statistical models used in solar PV forecasting. As the parameters in these models usually do not represent a physical phenomenon or quantity, such models are often referred to as 'black-box' models or functional approximates. The artificial neural network (NN) is another example of these models and is gaining popularity in PV forecasting owing to their modularity in handling non-linear models. There are various structures of NN, but they can be categorized into two: static [15] and dynamic [16]. However, artificial intelligence models can suffer from generalization problems. Also, there is no systematic way in arriving at the structure of the model. An alternative to this is the 'clear-box' model based on physical principles. The benefits of these models were outlined in the introduction. These equations are based on the physics of PV modules and are detailed in the preceding section.

B. PV Module Equations

There are different physical models proposed for PV modules - double diode models [17], simplified single diode model (SSDM), and further SSDM are some of them in a descending order of complexity. The higher complexity can provide better accuracy on the expense of increased computational burden which not suitable for real-time online applications.

The best model that gives a good compromise between simplicity and accuracy [18] is the simplified single diode model shown in Fig. 1.
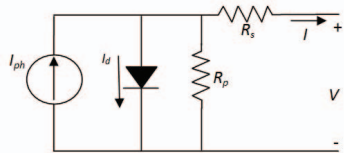

Figure 1. PV cell/array Single Diode Simplified model

The following equations describe the relation between the current and voltage output of the PV cell/array:

$$I = I_{ph} - I_d - \frac{V + R_s I}{R_p} \tag{1}$$

$$I_d = I_o (e^{\left(\frac{V + R_s I}{a V_t}\right)} - 1) \tag{2}$$

Where I is the output current of the cell in Amperes, V is the solar cell voltage in Volts, I_{ph} is the photocurrent in Amperes, I_d is the Shockley diode equation, I_o is the reverse saturation or leakage current of the diode, $V_t = kT/q$ is the thermal voltage of the array, q is the electron charge ($1.60217646 \times 10^{-19}$ C), k is the Boltzman constant (1.3806503×10-23 J/K), T is the temperature of the cell in Kelvin, and a is the ideality factor constant. More details of these equations can be found in [18]. To calculate power yield, values for I and V are usually computed using numerical methods [17],[18]. The mathematical approach is usually tedious especially when applied to large or widely spread PV systems [9].

C. Simplified PV Equations

Furthermore, the aforementioned equations of PV modules require numerical solution and thus are sometimes replaced with simplified equations that relate the power output with the efficiency of the system and variation in radiation and temperature [19],[20]. These equations are basically a translation of performance measurement from standard test measurements (STC; Air Mass 1.5 spectrum with global irradiance (G=1000W/m^2 and module temperature = 25°C). One famous simple method is that of Osterwarld [19] which can be described as follows:

$$P_m = P_{mo} \cdot \frac{G}{G_o} [1 - \gamma \cdot (T - 25)] \tag{3}$$

Where P_m is the cell/module maximum power (W), P_{mo} is the cell/module maximum power in STC (W), γ is the cell maximum power coefficient (^{o}C^{-1}) which ranges from -0.005 to -0.003 ^{o}C^{-1} in crystalline silicon and can be assumed to be -0.0035^{o}C^{-1} with good accuracy.

Another version of equation (7) is given below [20]:

$$P_{MPP} = G_t \cdot A_a \cdot \eta \cdot [1 + K_T (T_C - T_{ao})] \tag{4}$$

$$\eta = \eta_m \eta_{dust} \eta_{mis} \eta_{DCloss} \eta_{MPPT}$$

G_t is the global irradiance on the titled surface in W/m^2, K_T is thermal derating coefficient of the PV module in %/^{o}C, A_a area of the PV array in m^2, η_m is the module efficiency, η_{dust} is 1-the fractional power loss due to dust on the PV array, η_{mis} is 1-the fractional power loss due module mismatch, η_{DCloss} is 1- the fractional power loss in the dc side, η_{MPPT} is 1-fractional power loss due to the MPPT algorithm, T_C is the cell temperature in ^{o}C, T_{ao} is the ambient temperature at STC conditions in ^{o}C. The ac power of the PV system is then estimated by using manufacturer's efficiency curve of three phase inverter. The simplified PV equation adopted for this work is given below [21]:

$$P_{pv} = G_t \cdot A \cdot \eta_{PV} \cdot \eta_{loss} \cdot \eta_{inv} [1 - \gamma \cdot (T_m - 25)] \tag{5}$$

In this equation, miscellaneous losses including dust were lumped together in η_{loss}; PV cell efficiency η_{PV} and MPPT or inverter efficiency η_{inv} are kept separate. T_m is the module temperature.

The aforementioned equations require detailed modeling of the global irradiance falling on a tilted surface G_t as outlined in the next section.

D. Irradiance Falling on a Tilted Surface: Hottel's equations

There exist various models for calculating irradiance on a tilted panel. However, some of these models rely on other meteorological data such as total irradiance on horizontal surface, diffuse irradiance on horizontal surface, beam normal irradiance. Models of this type include those of Perez [22],[23] and Klucher [22],[24]. Others are not accurate in cloudy conditions, Temps and Coulson [25], or in clear skies, Liu and Jordan [26]. Simple models that require no additional solar measurements were proposed by Hottel [9],[27],[28] and are

adopted in this work. Description of this model is outlined below:

To explain irradiance equations, it is important first to present equations of solar angles as they are a pre-requisite to calculate solar equations.

The derivation of irradiance on tilted surfaces requires the calculation of different solar angles. These equations are mainly based on [29] and [30]. Solar angles that define the position of the sun with respect to a PV plane are illustrated in Fig. 2.
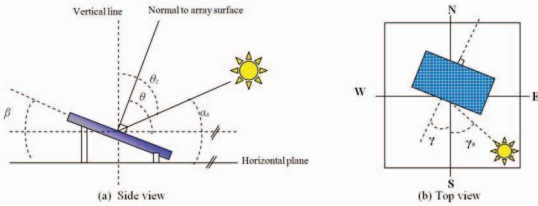

Figure 2. Solar angles of a PV plane

β = Tilt angle of array.

A_s = Solar elevation (altitude): the angle between the horizontal and line to the sun.

θ = Angle of incidence: the angle between normal to array surface and direct irradiance on a tilted surface (or line to the sun).

θ_z = Zenith angle: the angle between vertical line to earth and line to the sun.

γ_s = Solar azimuth angle: the angular displacement from south of the projection of beam radiation on the horizontal plane. Displacements east of south are negative and west of south are positive.

γ = Surface azimuth angle: the deviation of the projection on a horizontal plane of the normal to the surface from the local meridian, with zero due to south, east negative, and west positive; $-180° \leq \gamma \leq 180°$.

The zenith angle θ_z can be written as follows:

$$cos\theta_z = cos\varphi.cos\delta.cos\omega + sin\varphi.sin\delta \qquad (6)$$

Where

δ is the declination angle given by

$$\delta = 23.45.sin\left(360.\frac{284+n}{365}\right) \qquad (7)$$

φ is the latitude in degrees is the angular location north or south of the equator, north positive; $-90° \leq \varphi \leq 90°$.

ω is the hour angle which is the angular displacement of the sun east or west of the local meridian due to rotation of the earth on its axis at $15°$ per hour; morning negative, afternoon positive. The hour angle can be calculated by first calculating the solar time given by:

Solar time = standard time + 4. (L_{st} - L_{loc}) + E (8)

Where L_{st} is the standard meridian for the local time zone, L_{loc} is the longitude of the location in question, and longitudes are in degrees west. The parameter E is the equation of time in minutes and is given by:

$$E = 22.92(0.000075 + 0.001868 \cos B \qquad (9)$$
$$- 0.032077 \sin B$$
$$- 0.014615 \cos 2B$$
$$- 0.04089 \sin 2\ B)$$

Where B is calculated as follows:

$$B = (n-1)\frac{360}{365} \qquad (10)$$

The hour angle ω can then be written as:

$$\omega = (Solar\ time - 12).15 \qquad (11)$$

Furthermore, the incidence angle θ can be calculated using the following formula:

$$\cos\theta = sin\delta\ sin\varphi\ cos\beta - sin\delta\ cos\varphi\ sin\beta\ cos\gamma \qquad (12)$$
$$+ cos\delta\ cos\varphi\ cos\beta\ cos\omega$$
$$+ cos\delta\ sin\varphi\ sin\beta\ cos\gamma\ cos\omega$$
$$+ cos\delta\ sin\beta\ sin\gamma\ sin\omega$$

The solar irradiance falling on a tilted surface, G_t (W/m^2) is composed of three parts: the direct irradiance G_{tb} (W/m^2), the diffuse irradiance G_{td} (W/m^2) and reflected irradiance G_{tr} (W/m^2), i.e.

$$G_t = G_{tb} + G_{td} + G_{tr} \qquad (13)$$

The three components of irradiance can be calculated as follows:

$$G_{tb} = G_{on}\ \tau_b\ \cos\theta \qquad (14)$$
$$G_{td} = G_{on}\ \cos\theta_z\ \tau_d.\frac{(1+\cos\beta)}{2} \qquad (15)$$
$$G_{tr} = \rho.G_{on}\ \cos\theta_z\ \tau_r.\frac{(1+\cos\beta)}{2} \qquad (16)$$

Where G_{on} is the extraterrestrial radiation (W/m^2), τ_b is the beam atmospheric transmittance, τ_d is the diffuse atmospheric transmittance, and τ_r is the reflected atmospheric transmittance. G_{on} can be calculated as follows:

$$G_{on} = G_{sc}.\left(1 + 0.033.cos\left(\frac{360.d}{365}\right)\right) \qquad (17)$$

Where G_{sc} is 1367±5 W/m2 and d is the day of the year.

The beam atmospheric transmittance τ_b can be calculated as follows:

$$\tau_b = a_0 + a_1.e^{-k/\cos\theta} \qquad (18)$$

Where a_0, a_1, and k are constants that can be calculated as follows:

$$a_0 = r_0.[0.4237 - 0.00821(6 - A)^2] \qquad (19)$$
$$a_1 = r_1.[0.5055 + 0.00595(6.5 - A)^2] \qquad (20)$$
$$k = r_k.[0.2711 + 0.01858(2.5 - A)^2] \qquad (21)$$

Where A is the altitude of the location in km, r_o, r_1, and r_k are correction factors for different types of climates and are given in Table I.

TABLE I. COFFIECIENTS VALUES DEPENDENT ON CLIMATE

Climate type	r_0	r_1	r_k
Tropical	0.95	0.98	1.02
Midlatitude Summer	0.97	0.99	1.02
Subarctic Summer	0.99	0.99	1.01
Midlatitude Winter	1.03	1.01	1.00

III.   PV SYSTEM DESCRIPTION AND IDENTIFICATION VIA PSO

The test-bed system, located in Masdar city close to Abu Dhabi airport, is a 220,000 m^2, 10MW PV plant [31],[32]. The plant consists of around 87,777 panels: 17,777 are

polycrystalline and 70,000 are thin-film from Suntech and First Solar respectively. The parameters for the model were taken from data sheets of panels [33],[34] and from Engineers in Masdar and are given in Table II. These are the parameters with best engineering values taken from data sheet and engineers of the PV system. The model with these values will be referred to as the *clear box* model.

TABLE II. PARAMETERS OF **CLEAR BOX PV MODEL**

Longitude	54.45°
Latitude	24.43°
Altitude	1 m
Area of SunTech panels, A_{Sun}	30,911 m^2
Area of First Solar Panels, A_{First}	72,500 m^2
Miscellaneous losses (Suntech group), $\eta_{loss\ Sun}$	5% (i.e. $\eta_{loss\ Sun} = 95\%$)
Miscellaneous losses (Firstsolar group), $\eta_{loss\ First}$	6% (i.e. $\eta_{loss\ First} = 94\%$)
Efficiency of PV panel (Suntech), $\eta_{PV\ Sun}$	11%
Efficiency of PV panel (Firstsolar), $\eta_{PV\ First}$	10%
Efficiency of Inverter (Suntech panels), $\eta_{inv\ Sun}$	95%
Efficiency of Inverter (Firstsolar panels), $\eta_{inv\ First}$	94%
Temperature coefficient (%/C°), γ	5%
r_o	0.27
r_1	0.29
r_2	0.32
Albedo, ρ	0.35

A. CAutoD for Uncertainities in Clear-box Model:

The clear box model assumes different parameters with the values outlined in data sheet or are of constant value throughout the system. However and in reality values *change* with variation in atmospheric conditions and with aging of the materials. For the PV system studied in this work, some parameters were of uncertain value and thus were candidates for exploration of *better* practical values. These are the different PV efficiencies, albedo of ground, temperature coefficient, and miscellaneous losses; eight parameters in total. The practical ranges of these values are given in Table III. The identification of the best or optimized parameter value was conducted through PSO as explained in the preceding paragraphs.

TABLE III. RANGE OF UNCERTAIN PARAMETERS FOR **CAutoD CLEAR BOX PV MODEL**

$\eta_{loss\ Sun}$	92 - 96%
$\eta_{loss\ First}$	92 - 96%
$\eta_{PV\ Sun}$	10 - 12 %
$\eta_{PV\ First}$	10 - 12 %
$\eta_{inv\ Sun}$	92 - 99 %
$\eta_{inv\ First}$	92 - 99%
γ	2% - 6%
r_o	0.2 - 1.0
r_1	0.2 - 1.0
r_2	0.2 - 1.0
ρ	0.3 - 0.5

B. Grey-box PV Model:

Enhancement of the model was explored by introducing black-box models to account for unknown effects. The parameters of efficiency of inverter are known to be adaptive and are function of their loading [35]. Therefore, the following two grey-box models are proposed to model the change in efficiency with loading:

$\eta_{inv\_Sun} = 0.45 + a_1 . P_o + a_2. P_o^2 + a_3. P_o^3 + a_4. P_o^4$ $+ a_5. P_o^5$ $\eta_{inv\_First} = 0.45 + a_6 . P_o + a_7. P_o^2 + a_8. P_o^3 + a_9. P_o^4$ $+ a_{10}. P_o^5$	**Grey-box 1**
$\eta_{inv\_Sun} = \dfrac{0.45 + a_1.P_o + a_2.P_o^2 + a_3.P_o^3}{1 + a_4.P_o + a_5.P_o^2 + a_6.P_o^3}$ $\eta_{inv\_First} = \dfrac{0.45 + a_7.P_o + a_8.P_o^2 + a_9.P_o^3}{1 + a_{10}.P_o + a_{11}.P_o^2 + a_{12}.P_o^3}$ $P_o = \dfrac{P_{DC}}{P_{DC\,o}}$	**Grey-box 2**

where P_o is the fractional loading calculated by dividing the P_{dc} output of PV by nominal DC rating of the inverter.

The first model, Grey-box 1, is of a polynomial form while the second, Grey-box 2, is based on Padé approximation. Both will be explored to find the most suitable model for this application. The coefficients of these models were identified using PSO.

PSO was chosen as an identification algorithm to optimize the parameters in the CAutoD clear model and to find best coefficients in the grey box models respectively. The objective function for the identification is to minimize the Root Mean Square Error (*RMSE*) between actual and identified model in terms of PV power output. The RMSE is calculated as follows:

$$RMSE = \sqrt[2]{\frac{\sum_{i=1}^{i=n}(P_a^{\ i} - P_p^{\ i})^2}{n}} \quad (22)$$

Where P_i^a is the ith actual output power, P_i^p is the ith predicted power by model, and n is number of data points.

Particle Swarm Optimization [36] is inspired by social behaviour of bird flocking or fish schooling. It can be applied as follows:

Step 1: Initialize a population (array) of particles with random positions and velocities v on d dimension in the problem space. The particles are generated by randomly selecting a value with uniform probability over the dth optimized search space [x_d^{min}, x_d^{max}].

Step 2: For each particle x, evaluate the desired optimization fitness function, J, in d variables.

Step 3: Compare particles fitness evaluation with x_{pbest}, which is the particle with best local fitness value. If the current value is better than that of x_{pbest}, then set x_{pbest} equal to the current value and x_{pbest} locations equal to the current locations in d-dimensional space.

Step 4: Compare fitness evaluation with population overall previous best. If current value is better than x_{gbest}, the global best fitness value then reset x_{gbest} to the current particle's array index and value.

Step 5: Update the velocity v as follows:

$v_{id}(k) = w(k)v_{id}(k-1) +$
$\varphi_1.rand_1(x_{idpbest}(k-1) - x_{id}(k-1)) +$ (23)
$\varphi_2.rand_2(x_{idgbest}(k-1) - x_{id}(k-1))$

where, k is the number of iteration, i is the number of the particles that goes from 1 to n, d is the dimension of the variables, and $rand_{1,2}$ is a uniformly distributed random

number in (0, 1), $\varphi_{1,2}$ are acceleration constants and are set, as recommended by investigators [36], equal to 2. The weight w is often decreased linearly from about 0.9 to 0.4 during the search process.

Step 6: Update position of the particles,

$$x_{id}(t) = v_{id}(t) + x_{id}(t-1) \qquad (24)$$

Step 7: Loop to Step 2, until a criterion is met, usually a good fitness value or a maximum number of iterations (generations) m is reached.

PSO identification will search for 8 parameters in the CAutoD clear box model, 10 parameters in Grey-box model 1, and 12 parameters in Grey-box model 2.

IV.   SIMULATION RESULTS AND DISCUSSION

The models described in the preceding sections are simulated and compared as outlined below. The clear irradiance model based on Hottel's equations (6)-(21) is simulated and illustrated in Fig. 3. Furthermore, the clear PV model is simulated with values given in Table II. For the CAutoD clear box model, PSO is used to search for the optimum values of the uncertain parameters in the clear-box model. The data available from the system was used in two phases: first to tune the models and second to test or validate the model. To train the model, data (hourly PV power and module temperature) of days 5-20 in July 2010 (summer) and days 5-20 in January 2011 (winter) were used. The fitness function of the identification is chosen as the average error of July and January as shown below:

$$J = RMSE_{train} = (RMSE_{July} + RMSE_{January}) / 2 \qquad (25)$$

Once models are identified, the five consecutive days in both July 2010 and January 2011 are used to test the models and to compute the average forecasting $RMSE$ ($RMSE_{test}$). For PSO, the number of particles is set to $n = 50$ with a maximum number of search iterations of 300.

The results of the four models: **clear-box, CAutoD clear-box, Grey-box 1,** and **Grey-box 2** are summarized below. The coefficients identified by PSO for the respective models are given in Table IV. In the same table, RMSEs for training and for testing are given for the four models. The progress for identification is given in Fig. 4 for three optimized models. PV power predicted (test data) by the four models is given in Figures 5 and 6. The difference between accuracy of the models is seen better by analyzing the errors of modeling shown in Figures 7 and 8.

The following observations can be deduced from the results:

1. In general, tuning the parameters of the clear-box model proved to enhance the forecasting capabilities of the model. This was clear in the CAutoD clear- box, Grey-box 1, and Grey-box 2. Parameters from manufacturers require further adaptation to the unique atmospheric and location conditions of a given PV system. The *RMSEs* of tuned models are generally better than that of the clear-box model.

2. On the whole, introducing grey-box model enhanced the modeling accuracy compared with clear-box and optimized CAutoD clear-box as evident from both the training and testing the models. This can also be seen error plots in Figures 7 and 8.

3. PSO exhibited stagnation in identifying clear-box and grey-box models but at different times of progress. CAutoD clear-box identification was first to go through stagnation followed by Grey-box 2, though grey-box 2 reached a better *RMSE*. Grey-box 1 was better at stagnation indicating a better suited model for tuning via PSO.

4. Although, Grey-box 2 model produced the best $RMSE_{train}$, it was slightly surpassed by Grey-box model 1 in the testing phase. It can be said that Grey-box 2 exhibited *generalization* issue in comparison with the less complex Grey-box model 1.

5. Increasing or decreasing the order of Grey-box 1 and 2 was found to deteriorate the models accuracy and therefore kept at the given values.

6. General observation: the simplified model, equation 5, was found more sensitive to module temperature in July than in January i.e. excluding the temperature coefficient part (in brackets) had higher impact (worsen accuracy) in July than in January. This is expected as higher temperatures in July will impact the performance of the PV panels.

The following areas of further improvements were identified:

1. Different identification methods can be explored to further enhance the modeling capabilities. For example, PSO can be further enhanced to overcome the stagnation issue.

2. Different structure for grey-box models can be explored. Also, clear-box model can be further explored to identify candidate parameters to be replaced by black-box models to account for uncertainties.
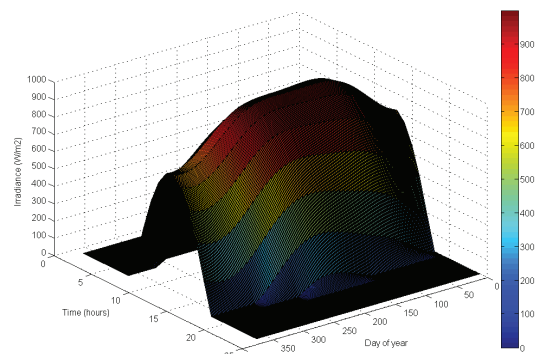

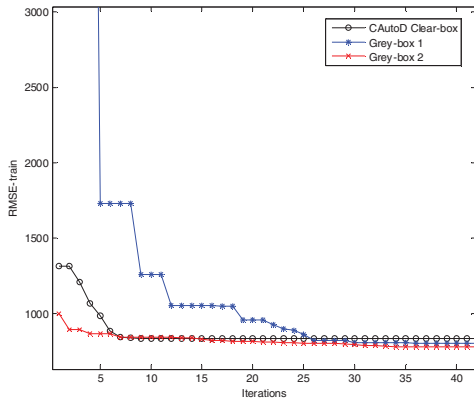Figure 3. Clear-day irradiance model for one year

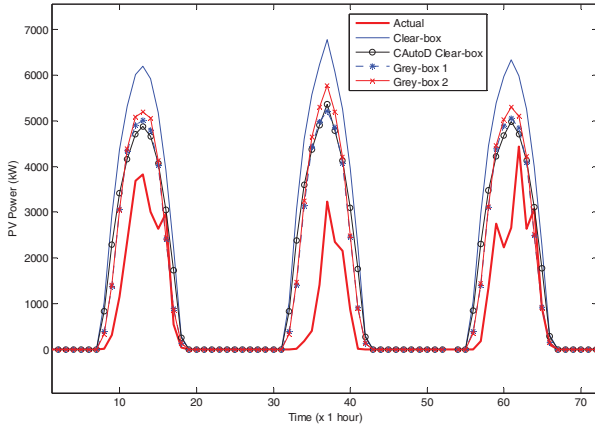Figure 4. Progress of PSO identification for different models

Figure 5. PV power predicted for different models: January 2011 (20-22 January). These were cloudy days.
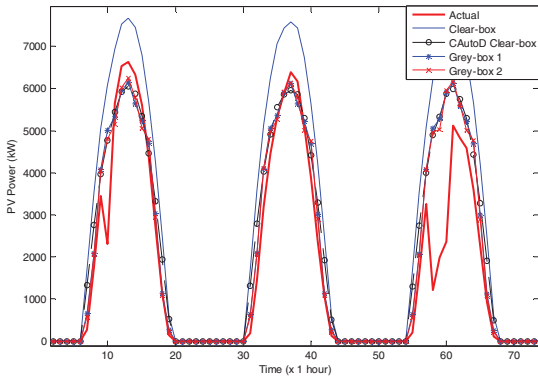

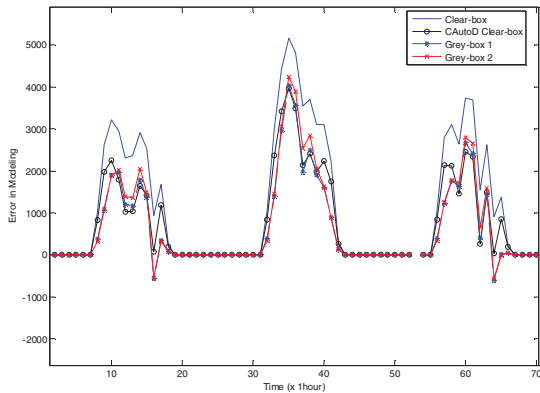Figure 6. PV power predicted by different models: July 2010 (20-22 July).

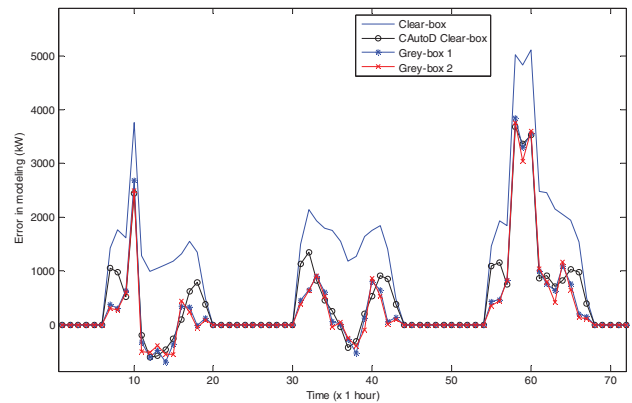Figure 7. Testing error in modeling for different models: January 2011 (20-22 January).

TABLE IV. IDENTIFIED MODELS AND RMSE

Parameters	Clear-box	CAutoD clear-box	Grey-box 1	Grey-box 2
η_{loss_Sun}	95%	92%	96%	96%
η_{loss_First}	94%	96%	96%	96%
η_{PV_Sun}	11%	10%	10%	10%
η_{PV_First}	10%	10%	10%	12%
η_{inv_Sun}	95%	92%	*Replaced by Grey box*	
η_{inv_First}	94%	92%		
γ (%/Co)	5%	5.3%	2.3%	6%
ro	0.27	0.2	1	0.2
$r1$	0.29	0.2	0.2	0.2
$r2$	0.32	1	1	0.2
ρ	0.35	0.3	0.3	0.5
a_1	*Not applicable*		-7.11	-2.42
a_2			0.55	5.21
a_3			6.36	-4.58
a_4			-1.46	-2.23
a_5			-5.24	-2.97
a_6			0.054	7.34
a_7			2.55	-1.32
a_8			3.73	4.16
a_9			-8.20	6.47
a_{10}			3.37	3.70
a_{11}			*Not applicable*	5.30
a_{12}				2.79
$RMSE_{train}$, kW	1322	835	775	**770**
$RMSE_{test}$, kW	1636	1067	**993**	1010

Figure 8. Testing error in modelling for different models: July 2010 (20-22 July).

V. Conclusion

The enhancement of modeling and forecasting of clear-box PV power models through introduction of black-box models was discussed in this paper. It was found that practical values of parameters can be tuned to improve the accuracy of the models. Further enhancement can be achieved through the introduction of grey-box models to account for uncertainties in the PV models. A free-derivative particle swarm engine was utilized in the identification process and was found particularly beneficial with simple grey-box models. The work presented is a novel step towards exploring the benefits grey-box models can add in the PV forecasting and hence supporting a better integration of these renewable energy sources in power networks.

References

[1] Global Market Outlook for photovoltaics 2013-2017, European Photovoltaic Industry Association (EPIA), May 2013.

[2] R. M. Swanson, "Photovoltaic Power Up", Science, Vol. 324, pp. 891-892, 15 May 2009.

[3] IRENA reference: "Renewable Energy Technologies: Cost Analysis Series, Photovoltaics", IRENA working paper, Volume 1: Power sector, Issue 4/5, June 2012.

[4] A. Woyte, V. V. Thong, R. Belmans, J. Nijs, "Voltage Fluctuations on Distribution Level Introduced by Photovoltaic Systems," IEEE Trans. On Energy Conversion, Vol. 21, No. 1, pp. 202- 209, March 2006.

[5] M. Sandoval, S. Grijalva, "An assessment on the impacts of photovoltaic systems on operational costs of the grid? The case of the state of Georgia," 2013 IEEE PES Conference On Innovative Smart Grid Technologies Latin America (ISGT LA), pp.1-6, 15-17 April 2013.

[6] J. Ma, S. Lu, R. P. Hafen, P. V. Etingov, Y. V. Makarov, V. Chadliev, "The impact of solar photovoltaic generation on Balancing Requirements in the Southern Nevada system," Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES, pp.1-9, 7-10 May 2012

[7] T. Stetz, F. Marten, M. Braun, "Improved Low Voltage Grid-Integration of Photovoltaic Systems in Germany," IEEE Transactions on Sustainable Energy, vol.4, no.2, pp.534-542, April 2013.

[8] L. Cibulka, M. Brown, L. Miller, A. V. Meier, "User Requirements and Research Needs for Renewable Generation Forecasting Tools that will Meet the Need of the CAISO and Utilities for 2020," A White Paper Report Prepared by CIEE, September 2012.

[9] C. Tao, D. Shanxu, C. Changsong, "Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement," , 2010 2nd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), pp.773-777, 16-18 June 2010.

[10] U. Forsell, P. Lindskog, " Combining semi-physical and neural network modelling: An example of its usefulness" Proceedings of the 11th IFAC symposium on system identification, Vol. 4, pp. 795–798, Kitakyushu, Japan, 1997.

[11] N. Al-Messabi, Yun Li, I. El-Amin, C. Goh, "Forecasting of photovoltaic power yield using dynamic neural networks," The 2012 International Joint Conference on Neural Networks (IJCNN), pp.1-5, 10-15 June 2012.

[12] Y. d. Valle, G. K. Venayagamoorthy, S. Mohagheghi, J.-C. Hernandez, R. G. Harley, "Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power Systems" IEEE Transactions on Evolutionary Computation, Vol. 12, No. 2, pp. 171-195, April 2008.

[13] P. Bacher, H. Madsen, H. A. Nielsen, "Online short-term solar power forecasting," Solar Energy, Vol. 83, pp. 1772-1783, 2009.

[14] B. Kroposki, K. Emery, D. Myers, and L. Mrig, "A Comparison of Photovoltaic Module Performance Evaluation Methodologies For Energy Ratings," 24th IEEE Photovoltaic Conference, pp. 858-862, 1994.

[15] A. Mellit, 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, pp. 807-821, 2010.

[16] L. A. Fernandez-Jimenez, A. Munoz-Jimenez, A. Flaces, M, Mendoz-Villena, E. Garcia-Garrido, P. Lara-Santillan, E. Zorzano-Alba, J. Zorzano-Santamaria, "Short-term Power Forecasting System for Photovoltaic plants," Renewable Energy, Vol. 44, pp. 311-317, 2012.

[17] J. A. Gow, C. D. Manning, "Development of a Photovoltaic Array Model for use in Power-Electronics Simulation Studies," IEE proceedings on Electric Power Applications, Vol. 146, no. 2, pp. 193-200, March 1999.

[18] M. G. Villalva, J. R. Gazoli, E. R. Filho, "Comprehensive Approach to Modeling and Simulation of Photovoltaic Arrays," IEEE Transactions on Power Electronics, Vol. 24, No. 5, May 2009.

[19] C. R. Osterwald, "Translation of Device Performance Measurements to Reference Conditions," Solar Cells, Vol. 18, pp. 269-279, 1986.

[20] W. Omran, M. Kazerani, M. M. A. Salama, "A Clustering-Based Method for Quantifying the Effects of Large On-Grid PV Systems," IEEE Transactions on Power Delivery, Vol .25, No. 4, October 2010.

[21] Md. Habibur Rahman and Susumu Yamashiro "Novel Distributed Power Generating System of PV-ECaSS Using Solar Energy Estimation," IEEE TRANSACTIONS ON ENERGY CONVERSION, VOL. 22, NO. 2, JUNE 2007

[22] R. Perez, "An anisotropic model for the diffuse radiation incident on slopes of different orientations," Proceedings of ASES, Minneapolis, pp. 883-888, 1983.

[23] R. Perez, "A new simplified version of the Perez diffuse irradiation model for tilted surfaces," Solar Energy, Vol. 39, No. 3, pp. 221-231, 1987.

[24] T. M. Klucher, "Evaluation of models to predict insolation on tilted surfaces," Solar Energy, Vol. 23, pp. 111-114, 1979.

[25] R. C. Temps, K. L. Coulson, "Solar radiation incident upon slopes of different orientations," Solar Energy, Vol. 19, pp. 179-814, 1977.

[26] B. Y. H. Liu and R. C. Jordan, "The long term average performance of flat-plate solar-energy collectors," Solar Energy, Vol. 7, No.2, pp. 53-74 1963.

[27] H. C. Hottel, "A simple model for estimating the transmittance of direct solar radiation through clear atmospheres," Solar Energy, vol. 18, no. 2, pp. 129 – 134, 1976.

[28] Charnon Chupong and Boonyang Plangklang, "Forecasting power output of PV grid connected system in Thailand without using solar radiation measurement," Energy Procedia, Vol. 9, pp. 230-237, 2011

[29] J. Duffie, W. Beckman, Solar Engineering of Thermal Process, Third Edition, John Wiley and Sons, 2006.

[30] M. Iqbal, "An introduction to solar radiation," Academic Press Canada, 1983.

[31] Website: Date 18 April 2015: http://www.masdar.ae/en/energy/detail/masdar-city-solar-pv-plant

[32] 10MW Solar Power Plant, Enviromena Power Systems Report, 2014 Website: http://enviromena.com/2015/wp-content/uploads/2014/01/10MW-Masdar_PD_2013-WEB.pdf.

[33] SunTech: SunTech panel data sheet: STP270S - 20/Wd+, STP265S - 20/Wd+ : www.suntech-power.com IEC-STP-WdS+-NO1.01-Rev 2014

[34] Thin film: Models: First Solar FS Series 2 PV Module data sheet: www.firstsolar.com, FS Series 2 PV Module PD-5-401-02 NA MAY 2011

[35] Aleksandar Pregelj, "Impact of Distributed Generation on Power Network Operation," PhD thesis, Georgia Institute of Technology, December 2003.

[36] R. Eberhat and Y. Shi, Particle Swarm Optimization: Developments, applications, and resources, Proceedings of the 2001 Congress on Evolutionary Computation, vol. 1, 2001, pp. 81-86.