# A PARTICLE FILTER APPROACH FOR SOLAR RADIATION ESTIMATE USING SATELLITE IMAGE AND IN SITU DATA

*Laurent Linguet[1], Jamal Atif[2]*

1. University of Antilles Guyane, UMR Espace-Dev - IRD, Cayenne, France; llinguet@guyane.univ-ag.fr
2. University of Paris Sud, LRI,TAO,INRIA/CNRS, Paris, France, jamal.atif@lri.fr

## ABSTRACT

The estimation of global solar radiation incident on the earth's surface is an important issue for several solar-based applications. From a signal processing point of view, it falls within non-stationary, non-linear/non-Gaussian dynamical inverse problems. In this paper, we propose a sequential Monte Carlo state space approach combining satellite images and in situ data. We propose original observation and transition functions taking advantages of the characteristics of both the involved type of data. A simulation study is carried along with a comparison with the state of the art established method, Heliosat.

## INTRODUCTION

The knowledge of the global solar radiation incident on the earth's surface and its geographical distribution is of prime importance for numerous solar-based applications (Climate change assessment, solar renewable energy systems). Hourly and daily values of solar radiation measurements with high spatial resolution that are always necessary for these applications imply unacceptable cost if provided by a high density ground based radiometric network. Besides interpolation techniques applied to radiation measurements become ineffective when the distance between the meteorological stations is greater than 34-50 km.

Satellite sensors can provide an alternative to the sparse coverage of radiometric networks since they can produce database over large regions on a high spatial resolution grid (1 by 1 km in visible range). However the computation of solar radiation by means of satellite images is unfortunately not straightforward. The satellite image is a top-of the-atmosphere (TOA) observation. The pixel value represents the flux density of the upward solar radiation emerging from the atmosphere, and the solar radiation absorbed by the ground is the fraction of the flux density of the downward solar radiation incident on the atmosphere. Determination of models capable of deriving global solar radiation at ground level from satellite images at high spatio-temporal resolution is an open issue in environmental research and solar applications that we are proposing to tackle in this paper.

Several mathematical models were studied, in order to estimate solar radiation from satellite images. Two different approaches to this subject were developed. Statistical models from the one hand, physical models from the other hand (1, 2). Statistical models (3) have evolved toward complex hybrid models by incorporating additional observation data and both empirical and physical information ((4, 5) among others). Direct information are given by meteorological satellites and indirect information such as transmittance are obtained by radiative transfert equation. This led to better spatial distribution of the models response. However, despite their increasing complexity and an improved usage of numerous available data, recurring obstacles (e.g., the difference in spatiotemporal scale between the model and measurements; measurement errors; or the simplification of physical processes) still introduce a significant amount of uncertainty into the model predictions.

In this paper we consider a Bayesian filtering approach to the dynamical estimation of the global solar radiation at ground level from satellite images. We defend the idea that an inverse approach based on sequential Monte Carlo filtering (6) helps to relax several assumptions and constraints while keeping estimations results in accordance with those of existing methods. Among these constraints, one can note the physical model and its parameter estimation. In fact, using a stochastic model allows, if the amount of data samples is sufficient, to associate (in a statistical sense) the satellite data to their corresponding in situ samples. This conducts to infer the radiation measure in a continuous way of a geographic map with a precision comparable to the well established methods.

The paper is organized as follows: after presenting the stochastic model along with the observation and transition laws in Section 2, we explain the sequential Monte Carlo sampling in Section 3. Section 4 presents our experiments and results and compares them to those obtained with a traditional model. Section 5 concludes and presents future directions.

## METHODS

Stochastic models are commonly used to describe the behavior of many processes. Model variables can be divided into hidden variables (that are not measured) as solar radiation estimates at surface, and observed variables as satellite image pixels. A combination of hidden and measured variables can be used to represent the dynamic behavior of the nonlinear process as described before.

### The data and notations

A common assumption underlying solar irradiance signal is that it can't be regarded as a stationary process due to the diurnal and annual variation related to the sun's changing angle. To remove these effects and obtain a weekly stationary stochastic process solar irradiance is often normalized by dividing solar radiation at the earth surface by the extraterrestrial solar irradiance. The result is defined as the clearness index. The horizontal irradiance outside the atmosphere is determined using:

$$G_0(i,j)=I_{sc}E_0\cos\theta_s(i,j) \tag{1}$$

where $I_{sc}$= 1367 W/m² is the solar constant, the extraterrestrial irradiance normal to the solar beam; $E_0$ is the excentricity correction factor and $\theta_s(i,j)$ is the sun zenithal angle at pixel (i,j). $E_0$ and $\theta_s$ depend on astronomical relationships and can analytically be determined for each instant k. Thus the knowledge of the clearness index allows the calculation of solar radiation at the earth's surface and inversely. Let $x_k$ denote the clearness index at time k :

$$x_k = G_k(i,j) / G_{0k}(i,j) \tag{2}$$

where $G_k(i,j)$, is the horizontal global irradiance at ground level for the time k and the pixel (i, j) and $G_{0k}(i,j)$ is the horizontal irradiance outside the atmosphere for the time k and the pixel (i, j). They are expressed in $W.m^{-2}$. Observations of our model refer to the apparent albedo $\rho(i,j)$ observed by the satellite sensor for the pixel (i, j) (containing the ground location). $\rho(i,j)$ has no unit and is equal to the bidimensional reflectance.

$$\rho(i,j) = \pi L(i,j) / (I_{sc}E_0\cos\theta_s(i,j)) \tag{3}$$

L(i,j) is the is the observed radiance. $x_k \in R^n$ is a state vector evolving according to the following equation: $x_{k+1} = f_k(x_k, v_k)$ where $v_k$ is i.i.d. random noise with unknown probability distribution function (pdf). At discrete times, observations $z_k \in \mathbb{R}^p$ become available and are related to the state vector via the observation equation: $z_k= h_k(x_k,w_k)$ The filtering problem can be formulated as:

$$x_k=f_k(x_{k-1})+v_{k-1} \tag{4 }$$

$$z_k=h_k(x_k)+w_k \tag{5}$$

v and w are the process noise and the observation noise. The state transition density is fully specified by $f_k$ and the process noise distribution and the observation likelihood are fully specified by $h_k$ and the observation noise distribution.

### The transition law (process law)

The first part of the stochastic model (eq. 4) is the transition law. In this work we use a transition law based on the ARMA (Auto-Regressive Moving Average Model) process called TAG (Time-dependant Autoregressive, Gaussian model) developped by Aguiar and Collares-Pereira (7) and designed to be independent of location and time of the year. This model generates synthetic daily sequences of the hourly clearness index $x_k$ as a Markov chain. To takes into account seasonal phenomena the variable $x_k$ is normalized (centrered and reduced):

$$y_k = \frac{x_k - x_{km}}{\sigma} \qquad (6)$$

where $x_{km}$ and $\sigma$ are the average hourly value and standard deviation of x and are calculated from the unique input of the model : $K_t$. $K_t$ is the monthly average of the daily clearness index. The wide availability of this type of monthly average data enables this model to be used almost anywhere. The sequential properties of the y variable have significative dependence on its previous value. The proposed ARMA(1,0) model is:

$$y_k = \phi_1 y_{k-1} + r \qquad (7)$$

where r is a random Gaussian variable with null average, and $\phi_1 = 0.38 + 0.06 \cos(7.4K_t - 2.5)$.

### The observation law

Apparent albedo $\rho^k(i,j)$ extracted from the digital satellite image over the time-interval are used to estimate what should be the hidden state $x_k$, clearness index at time k, by the knowledge of the observation law $h_k$.

The nonlinear function $h_k$ may be obtained using physical laws such such as radiative transfert functions. However, due to the complexity of physical processes, it is difficult to develop accurate and reliable nonlinear function, in particular in a tropical area (our area of study). The task is to obtain an estimate of the unknown $x_k$ where only the value of $z_k$ is known. One approach for estimating $x_k$ is modeling of the joint distribution p(x,z) with a learning dataset of clearness index data and apparent albedo data.

We consider a learning set consisting of M available paired data (X, Z). The $x_i$, (i = 1, 2,…,M), is the value of the clearness index obtained from a ground radiation measurement at time i associated with the apparent albedo value $z_i$ over the ground location. The inference task has three folds:

1. obtain the joint density p(x,z);
2. estimate the conditional distribution $p(x|z_k)$, for apparent albedo $z_k$ and
3. obtain an estimate $\hat{x}_k$ from such distribution.

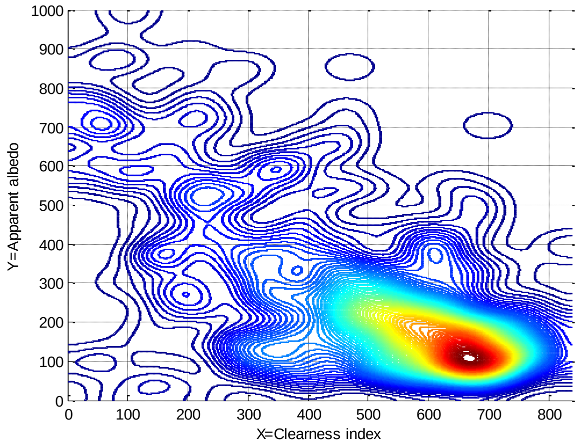


*Figure 1: Two dimensional estimated joint distribution p(x,z) between clearness index values and apparent albedo data* in

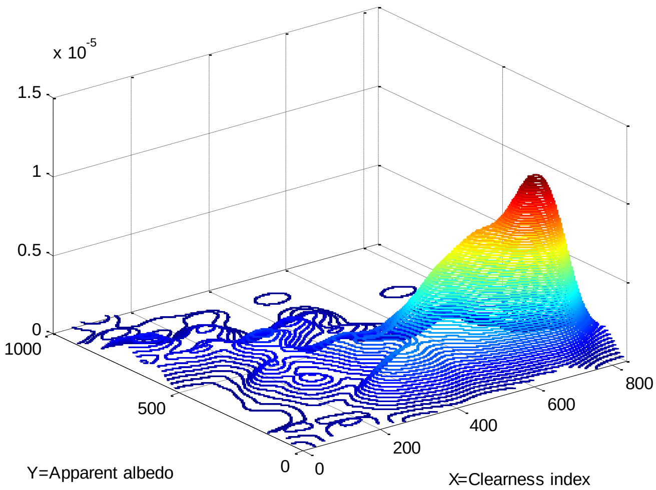


*Figure 2: Three dimensional estimated joint distribution p(x,z) between clearness index values and apparent albedo data*

We rely on a Monte Carlo approach to construct the joint density of p(x, z). The regular steps of a particle filter can generate an approximation of the joint pdf p(x, z) as the superposition of (equally weighted) local kernel densities centred about each sample ($x_i$, $z_i$), drawn from the learning set (8).

$$p(x,z) = \frac{1}{M}\sum_{i=1}^{M} K\left((x,z)|(x_i,z_i\right) \quad (8)$$

A common choice of Kernel density is the Gaussian Kernel. Each kernel can be propagated by using a local linearization (Fig. 1.) yielding a continuous output distribution p(x|z). Identifying the distribution of the clearness index state variable conditioned on the apparent albedo variable, p(x|$z_k$), reduces to identifying a marginal of this joint distribution. The conditional distribution p(x|$z_k$), is the section of p(x) at Z = $z_k$.

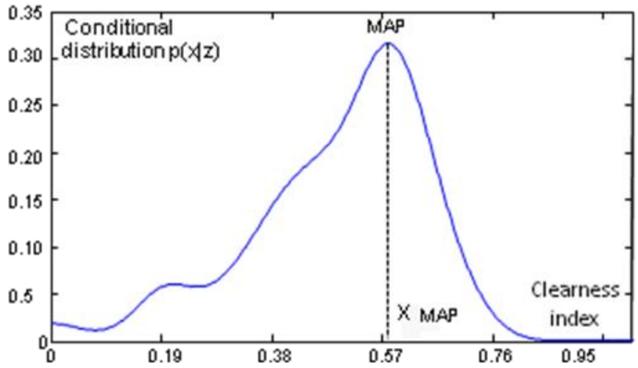


*Figure 3: Conditional distribution $p(x|z_k)$ extracted from* p(x|z)

Given the distribution of the clearness index x conditioned on the the apparent albedo $z_k$, the user has the freedom to choose any estimates of X. We choose the maximum a posteriori (MAP):

$$\tilde{x}_{MAP} = \text{argmax } p(x|z_k) \quad (9)$$

because it selects the maximum of the conditional distribution (note that other choices are possible).

## PARTICULE FILTERING

Solar radiation estimate through satellite images is a problem of causally estimating a hidden state sequence from a sequence of observations that satisfy the Hidden Markov Model (HMM) assumption. The problem is to recursively compute the "posterior" at time k using the posterior at time (k-1) and the current observation (probability density function of the current state conditioned on all observations until the current time). In others words, the problem is to find an update formula from $p(x_{k-1}|z_{1:k-1})$ to $p(x_k|z_{1:k})$ where $z_{1:k}$ denotes all observation $\{z_1, \ldots, z_k\}$.

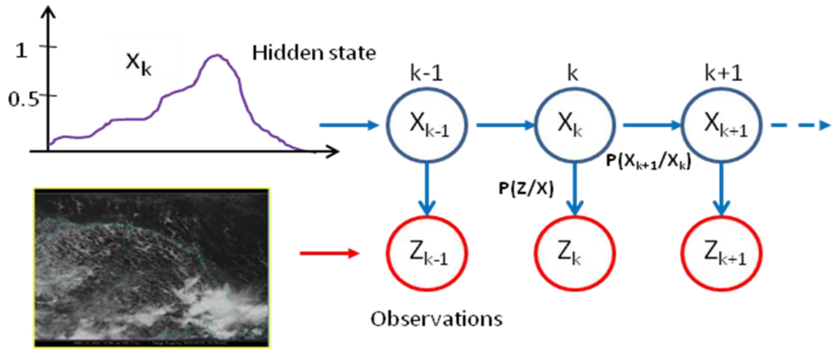


*Figure 4: Bayesian Network of the Hidden Markov Model. Clearness index, $x_k$, is the hidden state and apparent albedo extracted from the digital satellite image, $z_k$, is the observation.*

For most nonlinear or non-Gaussian state space models, the posterior cannot be computed analytically. However, it can be efficiently approximated using a particle filter (PF) based sampling (9) which is a Sequential Monte Carlo technique. A PF is a recursive algorithm which produces at each time k, a cloud of N "particles" (Monte Carlo samples), along with their corresponding weights, whose empirical measure closely approximates the true posterior for large N.

Time evolution is achieved with an importance sampling distribution via sequences of sampling and importance weighting. For simplicity reasons, we choose the importance density as the prior $p(x_k|x_{k-1})$. In order to overcome the major problem of PF techniques, the particles degeneracy, we introduce a resampling strategy. This algorithm is designed as the SIR particle filter [8]. Detailed sequences of our algorithm are given below :

▷ **(1) Initialization**:
set m = number of iterations;   set n = number of samples
set k = 1 and select an appropriate initialization density p($x_1$| $z_0$) [a]
for k = 1 to m
do
▷ **(2) Prediction**:
if k = 1
then
    Draw samples {$x_k^{(i)}$}from $p(x_1|\ z_0)$
else
  Draw samples {$x_k^{(i)}$}from $p(x_k|\ x_{k-1})$ [b]
▷▷Compute the predicted state: $E(x_k|z_{1:k-1}) = \sum_{i=1}^{n} \omega_{k-1}^{(i)} . x_k^{(i)}$

▷ **(3) Estimate the MAP** $\tilde{x}_{MAPk}^{(i)}$ from $p(x|z_k^{(i)})$ for each $z_k^{(i)}$ [c]
▷ **(4) Weights the samples** according to the likelihood:
Evaluate the weights: $\omega_k^{(i)} = p(\tilde{x}_{MAPk}^{(i)}|x_k^{(i)})$
Normalize the weights: $\omega_k^{(i)} = \frac{\omega_k^{(i)}}{\sum_{i=1}^{n} \omega_k^{(i)}}$
▷ **(5) Compute the state estimate**: $E(x_k|\ z_{1:k}) = \sum_{i=1}^{n} \omega_k^{(i)} . x_k^{(i)}$
▷ **(6) Re-sampling**
▷▷ Draw (with replacement) n samples from {$x_k^{(i)}$}
▷▷ so that $x_k^{(i)}$ is selected with probability $\omega_k^{(i)}$
▷▷ The new set is denoted as {$x_k^{(i)}$}

[a] SIR particle filter simulations were conducted according to the observation dataset. Since the dataset provided by meteorological stations consists of daily whole sequences, initial measurement corresponds in fact to sunrise time. Initial distribution is then chosen to be a white noise distribution with a zero mean. An initial set of particles {$x_0^{(i)}|_{i=1, 2,...., n}$} is formed with uniform weights: $\omega_0 = \frac{1}{n}$

[b] according to the transition law (2.2)

[c] according to the joint density $p(x,z)$ of the learning dataset. $\tilde{x}_{MAPk}^{(i)}$ refers to equation (9)

"Systematic resampling" scheme (10) is used in the last sequence. It is achieved by setting :

$U_i$ = (i-1)/n + U, where U is a single random drawn from the U([a,b]) denotes the uniform distribution on the interval [a; b]. Its performance is generally found to be close to that of "residual" and "stratified" resampling. This scheme is often preferred due to its computational simplicity.

## RESULTS

Hourly and daily global solar irradiance estimates derived from particle filter model are compared with hourly and daily global irradiance measurements performed at a single ground station by the French National Meteorological Service of the French Guiana. Comparisons with the Heliosat2 method statistics were also made (11). Heliosat2 is an existing satellite estimation method produced by Mines ParisTech. The Heliosat2 method (4) converts images acquired by meteorological geostationary satellites into data and maps of solar radiation received at ground level all over Europe, Africa, and the Atlantic Ocean.

In order to develop and validate the particle filter model a set of 4454 high resolution satellite images (GOES EAST) from the visible channel (0.4µm-1.1µm) covering a 207 days period from the year 2010 has been selected. This selection allows for various sky coverages. The apparent albedo observed by the satellite sensor is determined for cell of 0.2°x0.2° in size by averaging several pixels. The joint probability distribution function (pdf) between clearness index and apparent albedo is obtained using a learning dataset including randomly chosen satellite images and measurement data from 4 ground meteorological stations spread over the 84000 km² of the French Guiana territory.

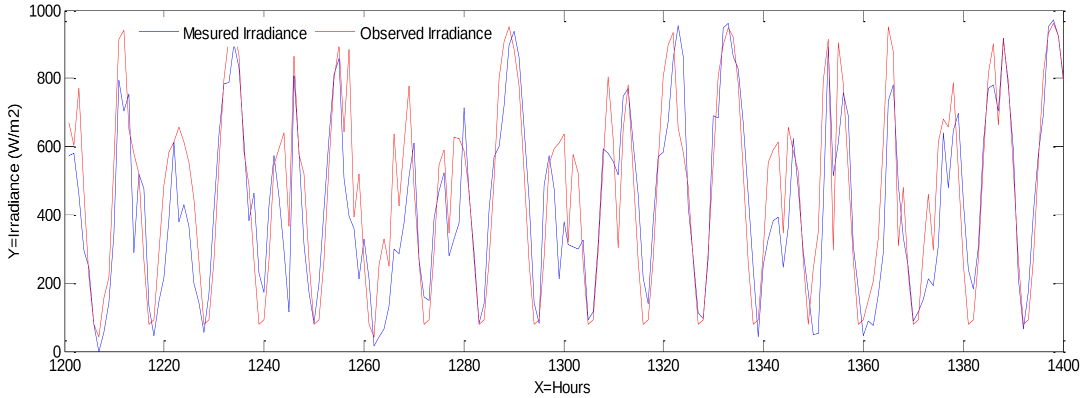


*Figure 5: Comparison between hourly measured and hourly estimated irradiance*

Hourly and daily solar irradiance for a single station were estimated using a test dataset (with data not used in the learning dataset) processed by the SIR particle filter with 400 samples. In Figure 5 the particle filter based estimates are compared with solar radiation measurements from the ground stations on a hourly basis. The results indicate that the models overestimated the radiation for lower irradiances. Figure 6. shows comparisons between daily measured and daily estimated irradiance.

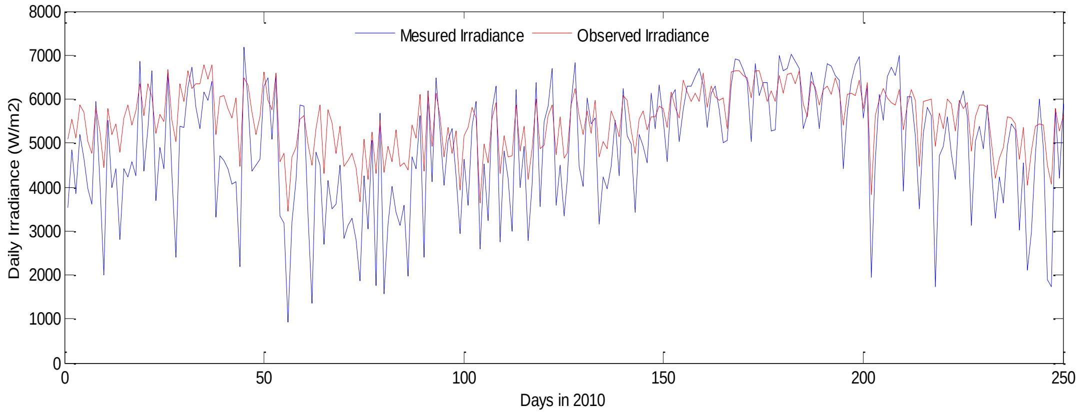


*Figure 6: Comparison between daily measured and daily estimated irradiance*

The performance of the particle filter based model was estimated using root mean squared error (RMSE) and mean bias error (BIAS) on a daily basis and are presented in Table 1. Relative BIAS and RMSE are also given as a percentage of the daily averaged measured irradiance. Relative RMSE obtained for Particle Filter based model is similar to average RMSE obtained with Heliosat model (11).

$$RMSE = (\Sigma^{n}(R_m - R_{est})^2/n)^{1/2}$$

$$BIAS = \Sigma^{n}(R_m - R_{est}))/n$$

Where $R_m$ is the measured ground solar radiation value and $R_{est}$ is the estimated solar radiation value. They are similar.

*Table 1: RMSE and BIAS between daily solar irradiance measurements and daily solar irradiance estimates (values in W/m² and percentage)*

| | RMSE | Bias (MBE) |
|---|---|---|
| W/m² | 566 | 220 |
| Relative (%) | 10% | 4% |
| Heliosat2 | ~10% | ~5% |

## CONCLUSIONS

The developed method connects global and local dynamics of solar irradiance in a Bayesian framework by using the existing relation between clearness index data and satellite apparent albedo (3). A particle filter approach has been developed to estimate solar radiation at surface using satellite images. The proposed method incorporates statistical model for observation process. The joint distribution of state variable and observation variable is not restricted by any prior assumption and gives a probabilistic perspective based on conditional distribution estimates. The observation model takes advantage of the statistical relationship between the clearness index data and apparent albedo of satellite image to avoid introduction of complex radiative transfert equations while keeping estimation results in accordance with those of existing methods.

We demonstrate the use of a SIR particle filter for deriving solar radiation estimates using remote sensing. However the method need to be improved for low daily solar irradiances. In this work we have focused on MAP estimate, however it may be supposed that a substantial reduction of the daily RMSE can be gained by optimizing the joint pdf and state estimates choice and further investigations will be made in this direction. Future works need to be pursued on a global scale and on various ground covers.

## ACKNOWLEDGEMENTS

The authors are grateful to Météo-France for supplying ground measurement data. The authors thank the Feder and the Conseil Régional of French Guïana.

## REFERENCES

[1] Noia M, C F Ratto and R Festa, 1993a. “Solar irradiance estimation from geostationary satellite data: I. Statistical models,” Solar Energy, vol 51: 449-456.

[2] Noia M, CF Ratto and R Festa, 1993b. “Solar irradiance estimation from geostationary satellite data: II. Physical models,” Solar Energy 51 : 457-465.

[3] Cano D, Monget, J-M, Albuisson M, Guillard H, Regas N, and Wald L, 1986. A method for the determination of the global solar radiation from meteorological satellite data. Solar Energy 37: 31–39.

[4] Rigollier C, M Lefevre, L Wald, 2004. “The method Heliosat2 for deriving shortwave solar radiation from satellite image,” Solar Energy, Vol 77: 159-169.

[5] Martins F R, E B Pereira, S L Abreu, 2007. “Satellite-derived solar resource maps for Brazil under SWERA project,” Solar Energy, Vol 81: 517-528.

[6] Doucet A, J F G de Freitas, N J Gordon Eds., 2001. Sequential Monte Carlo methods in practice, Springer-Verlag, NewYork.

[7] Aguiar R, M Collares-Pereira, 1992. “TAG: a Time-dependent, Autoregressive Gaussian model for generating synthetic hourly radiation,” Solar Energy Vol 49: 167-174.

[8]. Lehn-Schioler T, D Erdogmus, and J C Principe, 2004. “Parzen particle filters,” in Acoustics, Speech, and Signal Processing, 2004. Proceedings. IEEE, vol. 5.

[9] Arulampalam, S Maskell, N Gordon, T Clapp, 2002. “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking,” IEEE Transactions on Signal Processing, Vol.50.

[10] Kitagawa G, 1996. “Monte Carlo filter and smoother for non-Gaussian nonlinear state space models,” Journal of computational and graphical statistics, vol. 5, no. 1 : 1–25.

[11] Blanc P, Gschwind B, Lefèvre M, and Wald L, 2011. The HelioClim Project: Surface solar irradiance data for climate applications. Remote Sensing, 3 : 343-361.