

# Kalman filter model, as a tool for short-term forecasting of solar potential: case of the Dakar site

A. Mbaye<sup>1</sup>, J. Ndong<sup>2</sup>, M.L.NDiaye<sup>1</sup>, M. Sylla<sup>1</sup>, M.C. Aidara<sup>1</sup>, M. Diaw<sup>1,3</sup>, M.F.NDiaye<sup>1</sup>, Papa Alioune Ndiaye<sup>1</sup>, A.Ndiaye<sup>1</sup>

<sup>1</sup>Electrical Engineering Department of the Polytechnic High School of Dakar Senegal, Laboratory International Center for Solar Energy, BP 5085 Dakar-Fann, Senegal.

<sup>2</sup>Department of Mathematics and Computer Science, University Cheikh Anta Diop of Dakar Senegal, Laboratory LID.

<sup>3</sup>Department of physical, University Cheikh Anta Diop of Dakar Senegal, Laboratory of Hydraulic and fluid Mechanics.

**Abstract.** The prediction of solar potential is an important step toward the evaluation of PV plant production for the best energy planning. In this study, the discrete Kalman filter model was implemented for short-term solar resource forecasting on the Dakar site in Senegal. The model input parameters are constituted at a time  $t$  of the air temperature, the relative humidity and the global solar radiation. The expected output at time  $t+T$  is the global solar radiation. The model performance is evaluated with the square root of the normalized mean squared error (NRMSE), the absolute mean of the normalized error (NMAE), the average bias error (NMBE). The model Validation is carried out by means of the data measured within the Polytechnic Higher School of Dakar for one year. The simulation results following the 20 minute horizon show a good correlation between the prediction and the measurement with an NRMSE of 4.8%, an NMAE of 0.27% and an NMBE of 0.04%. This model could contribute to help photovoltaic based energy providers to better plan the production of solar photovoltaic plants in Sahelian environments.

**Keywords.** Kalman filter, solar potential, Prediction, Dakar, short-term.

## 1 Introduction

Senegal has significant potential for the solar resource development (PV and Thermal) with 3000 hours of sunshine per year and an average global irradiation of 5.8 kWh / m<sup>2</sup> / day. The solar photovoltaic market is growing in the world in general and in Africa in particular. Currently, we are witnessing a proliferation of solar photovoltaic plants in Senegal. The integration of these intermittent sources on the SENELEC grid is causing many problems in terms of planning for the maintenance of the balance between consumption and production. Planning is less easy because of the stochastic nature of the renewable energy sources production. A solution can be considered to overcome this limitation : forecasting. Potential prediction can make a substantial contribution in this context by evaluating the penetration rate of intermittent energies on the grid that would limit the risks. The forecasting of the production of intermittent energies makes it possible to anticipate the availability of the sources of production and thus to facilitate the management of the network. The prediction of the solar potential on a horizontal surface can be done in the long term [1], [2] and the short term [3], [4]. It is in this perspective that, this study is made to contribute to the search for reliable prediction methodologies of photovoltaic power. Four

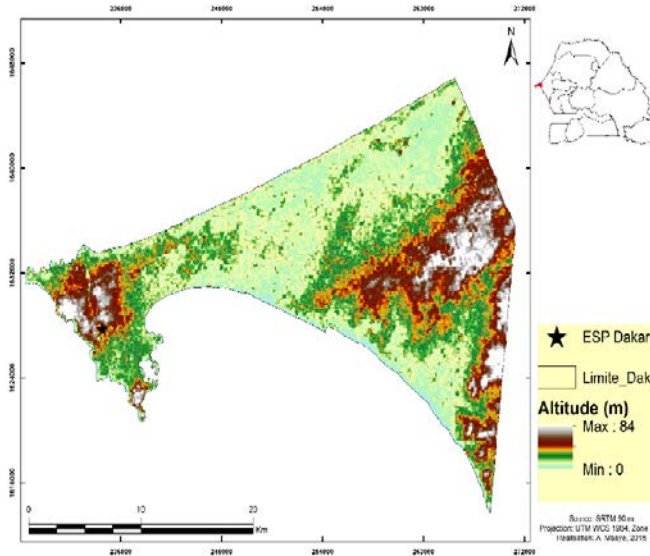
methodologies exist in the literature after [5] and [6]. We chose to work with a model based on the kalman filter which has already been widely applied in other sites around the world [7], [8], [9], [10], [11]. The Kalman filter model is convenient for forecasts at different horizons. We want to contribute to the short-term prediction of the solar potential of solar PV plants over the time horizon 20 minutes. Section 2 describes the methodology and tools used, Section 3 presents the results of simulations and their interpretation, and the last section is reserved for conclusion.

## 2. Materials and Methods

### 2.1. Study area

Dakar is located in the extreme west of the peninsula of Cape Verde, at the edge of the Atlantic Ocean, between latitude 14 ° 75 N and longitude -17.33 ° W with average altitude of 16m (figure1). It houses some meteorological stations including the measuring station of the International Center for Training and Research in Solar Energy (CIFRES measuring at the Advanced school of Engineering, in Cheikh Anta Diop University, Dakar, Senegal. This station measures global, direct and diffuse solar radiation, ambient temperature, relative humidity,

rainfall, wind speed and direction. In this work, we used solar radiation, average temperature as well as relative humidity data from CIFRES station. The data *series* runs from October 1st, 2016 to September 30th, 2017 and measurements are made every 10 seconds. A Campbell Scientific model CR800 datalogger is used.



**Figure 1.** Location of the Dakar-ESP area (Senegal)

## 2.2. Application of the filter on the study site

We build our prediction model assuming that measured solar power is governed by a hidden signal as in the case of a linear dynamic system following this equation:

$$Y_t = A_t X_t + V_t \quad (1)$$

$Y_t \in R^m$  denotes a stochastic process containing the measurement variable (observation) at time  $t$ , it is composed of  $m$  variables corresponding to the measured solar irradiation or photovoltaic power at different points. Generally, due to the intrinsic imperfection of measurement devices, errors occur more often during data collection.  $V_t$  denotes a stochastic process representing the measurement error. The stochastic process  $X_t \in R^n$  is the hidden state of the system that is transformed into an output by the matrix  $A_t \in R^{m \times n}$  that controls the sources of the measurements (solar radiation, ambient temperature, relative humidity). The goal is to find a model for the stochastic process  $X_t$  that is able to track solar radiation. We propose a simple stochastic model where  $X_t$  at discrete time  $t$ , is represented by the linear combination of two components (see equation 2).

$$X_t = C_t \hat{X}_t + \xi_t \quad (2)$$

$\hat{X}_t$  is a predictable component,  $\xi_t$  is a random noise component and  $C_t \in R^{n \times n}$ , is a parameter matrix describing the dynamics of the state.

Generally, the prediction model can have a structure and the noise process can be generated by any distribution. However, linear stochastic predictive models combined with Gaussian noise have a long record of successful applications in a very wide range of engineering techniques. Our idea is to connect the predictable components with the state model. These

variables are not directly observable by measuring devices; we refer to them as hidden states of the system. As we want to estimate the variables  $X_t$  as the state variable at the date  $t$ , we can use a linear stochastic dynamic system based on state space models.

We then build a temporal model that relates from  $X_{t+1}$  to  $X_t$  with the following difference equation (3):

$$X_{t+1} = C_t X_t + B_t U_t + W_t \quad (3)$$

The  $n \times n$  matrix connects the previous state  $t$  to the current state  $t + 1$ ;

Compared to our study site,  $C_t$  describes the atmospheric conditions since the value of the solar radiation at time  $t$  depends on the ambient temperature, the relative humidity. The value of the solar power between  $t$  and  $t+1$  could change according to these atmospheric conditions. We can therefore estimate  $C_t$  given a set of measured data that contains all types of weather conditions to form our forecast model. Indeed, to see the influence of these atmospheric variables on the measurements, it is necessary to take into account the terms  $B_t$  and  $U_t$  which are deterministic matrices where matrices of parameters describing the dynamic state of

$R^{n \times p}$  and  $R^p$  size ( $B_t \in R^{n \times p}$  and  $U_t \in R^p$  with  $p$  exogenous variables), or  $W_t$  is the intrinsic noise process.

To form the complete state space model for the specification of our prediction framework, equations (1) and (3) are now combined to form equation (4). This equation is the classical form of a linear dynamic system with inputs. In this model, we assume that the state noise  $W_t$  and the measurement noise  $V_t$  are gaussian white noise processes at zero mean that are not correlated with covariance matrices  $Q_t \in R^{n \times n}$  and  $R \in R^{m \times m}$ , respectively. The identification of the parameters  $C_t$ ,  $A_t$ ,  $Q_t$  and  $R_t$  is necessary to properly calibrate the system. The method used for the identification of the parameters is based on the Expectation-Maximization (EM) algorithm described in [7].

The filter also functions as a predictor-corrector algorithm. As an iterative analyzer, he estimates the state of the system using two steps: the forecast intervenes in the phase of updating the time and the correction in the phase of updating the measurement.

## 2.3. Model Performance Criteria

Within the prediction context of global solar radiation or PV power, there are many methods to check whether a predictor is effective [7]. In this paper, the tools we used are the NRMSE, NMAE, NMBE.

$$NRMSE = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - \hat{P}_i)^2}}{\max(P_i) - \min(P_i)} \quad (4)$$

$$NMAE = \frac{\frac{1}{N} \sum_{i=1}^N |P_i - \hat{P}_i|}{\max(P_i) - \min(P_i)} \quad (5)$$

$$NMBE = \frac{\frac{1}{N} \sum_{i=1}^N (P_i - \hat{P}_i)}{\max(P_i) - \min(P_i)} \quad (6)$$

### 3 Results

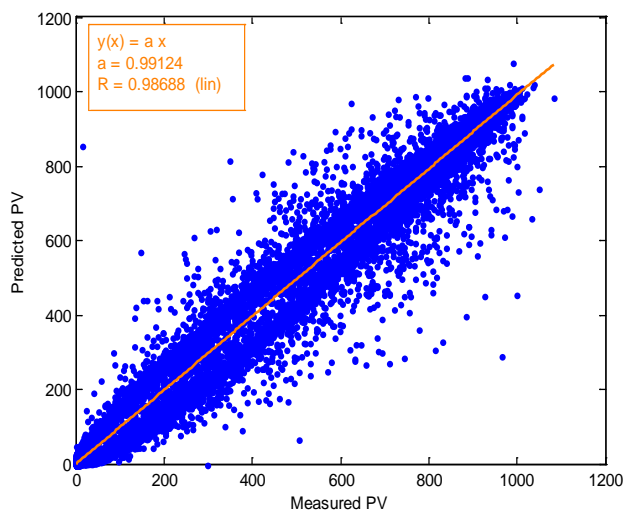
#### 3.1 Results obtained

**Table 1.** Performance of prediction model for time horizon 20 minute.

Model	NRMSE	NMBE	NMAE	Time horizons
EM	0.048	0.004	2.718	20

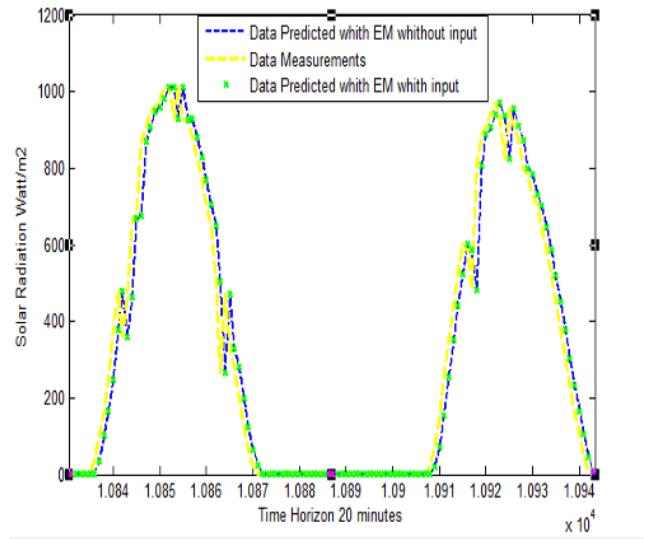
The table 1 summarizes the performance results of the EM model evaluated for horizon time of 20 minutes. The values of NRMSE and NMAE are equal to 0.048 and 2.718 respectively: this means that the model is ideal. The NMBE values for the EM algorithm-based filter model are positive (0.004) for the 20-minute horizon. The model tends to overestimate the values measured at the study site.

Figure 2 shows the correlation between the measurements and the prediction obtained with the EM algorithm for 20-minute time bin. It is possible to note most of the data points are concentrated around the straight line of the linear function. This graph tells us that there's a good agreement (R=0.99) between the predicted data and the measurements on the study site.



**Figure 2.** Correlation between the measurements made on the study site and those predicted by the filter model based on the EM algorithm.

Figure 3 presents the evolution of the values observed on the study site (yellow color), values predicted by the filter model based on the EM algorithm taking into account the impact of temperature and humidity (curve in green) and those predicted by the same model ignoring these atmospheric parameters (curve in blue). The filter model with and without impact of the atmospheric parameters (temperature and relative humidity) follows well the measured data.



**Figure 3.** Evolution of the data measured and predicted by the Kalman filter via EM with input and without input.7

### 4 Conclusion

As part of this work we applied the kalman filter model to measured data on the Dakar (Senegal) site in order to make a short-term prediction of solar potential. The prediction is based on the Kalman filter model combined with the EM algorithm for estimating the parameters needed to execute the filter. We evaluated the model on the time horizon 20 minutes. The precision of this technique is studied with a set of standardized performance criteria taking into account the square root of the mean squared error (NRMSE), the absolute mean of the error (NMAE), the mean error of bias (NMBE). The simulation results following the 20-minute horizon show a good correlation between the prediction and the measurement with an NRMSE of 4.8%, an NMAE of 0.27% and an NMBE of 0.04%. This model can be used to predict the photovoltaic solar potential and the production of solar power plants for 20-minute time horizon on the DAKAR site in Senegal. It could contribute to the planning of energy providers based on photovoltaic power plant in Sahelo-Sahelian environment.

### Acknowledgments

I am very grateful to the center of excellence in mathematics, computer science and ICT (CEA-MITIC) of Gaston Berger University of Saint-Louis (Senegal) for funding research mobility for my participation in the 2018 3rd International Conference on Sustainable and Renewable Energy Engineering in Zamora, Spain. I thank the world science laboratory: c/o CERN,LAA Building # 29,CH-1211 Geneva 23, Switzerland, E-mail: info@worldlab.ch.

Thanks to the fellowships of the World Laboratory of Scientists during my 3 years of thesis: I was able to do my research quietly without problem.

## References

1. A. K. Yadav and S. S. Chandel. Solar radiation prediction using Artificial Neural Network techniques. *Sustainable Energy*, 2014; 33:772-**781**.
2. M. Ikhsan and al. Study of Renewable Energy Sources Capacity and Loading Using Data Logger for Sizing of Solar-wind Hybrid Power System. *Procedia Technol*, 2013; 11:1048-**1053**.
3. A. Mellit and al. Short-term forecasting of power production in a largescale photovoltaic plant. *Solar Energy*, 2014; 105:**401-413**.
4. H. Madsen and al. A tool for predicting the wind power production of offshore wind plants.in *Proceedings of the Copenhagen Offshore Wind Conference &Exhibition*, 2005.
5. PJ. Brockwell and Davis RA. *Time series: theory and methods*. Springer series in statistics, second edition, 2006.
6. JD.Hamilton. *Times series analysis*, 1994.
7. T.Soubdhan, J Ndong and al. A robust forecasting framework based on the Kalman filtering approach with a twofold parameter tuning procedure: Application to solar and photovoltaic prediction. *Solar Energy*, 2016; 131:**246-259**.
8. DJ.Docimo and al. Extended Kalman Filtering to estimate temperature and irradiation for maximum power point tracking of a photovoltaïque. *Energy*, Elsevier, 2017; 120:47-**57**.
9. J.Antonanzas and al. Review of photovoltaic power forecasting. *Solar Energy*, 2016; 136: **78-111**.
10. G.Lorenzo and al. Day-Ahead Hourly Forecasting of Power Generation from Photovoltaic Plants.*IEEE Transactions on Sustainable Energy*, 2017;**99:1-1**.
11. M.Pierro and al. Data-drivenupscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data. *Solar Energy*, 2017; 158:**1026-1038**.