Short-term solar irradiance and irradiation forecasts via different time series techniques: A preliminary study

Cédric Join, Cyril Voyant, Michel Fliess, Marc Muselli, Marie-Laure Nivet, Christophe Paoli, Frédéric Chaxel

Short-term solar irradiance and irradiation forecasts via different time series techniques: A preliminary study

Cédric Join[1,2,3], Cyril Voyant[4,5], Michel Fliess[2,6],
Marc Muselli[5], Marie-Laure Nivet[5], Christophe Paoli[7] and Frédéric Chaxel[1]

1.- CRAN (CNRS, UMR 7039), Université de Lorraine
BP 239, 54506 Vandœuvre-lès-Nancy, France
{cedric.join, frederic.chaxel}@univ-lorraine.fr
2.- AL.I.E.N. (ALgèbre pour Identification & Estimation Numériques)
24-30 rue Lionnois, BP 60120, 54003 Nancy, France
{cedric.join, michel.fliess}@alien-sas.com
3.- Projet Non-A, INRIA Lille – Nord-Europe, France
4.- Hôpital de Castelluccio, Unité de Radiothérapie
BP 85, 20177 Ajaccio, France
5.- SPE (CNRS, UMR 6134), Università di Corsica Pasquale Paoli
20250 Corte, France
{voyant, marc.muselli, marie-laure.nivet}@univ-corse.fr
6.- LIX (CNRS, UMR 7161), École polytechnique
91128 Palaiseau, France
Michel.Fliess@polytechnique.edu
7.- Département de Génie Informatique, Université Galatasaray
No:36 34357 Ortaköy, Istanbul, Turkey
cpaoli@gsu.edu.tr

***Abstract*—This communication is devoted to solar irradiance and irradiation short-term forecasts, which are useful for electricity production. Several different time series approaches are employed. Our results and the corresponding numerical simulations show that techniques which do not need a large amount of historical data behave better than those which need them, especially when those data are quite noisy.**

***Keywords*-Meteorological forecasts, solar irradiance and irradiation, electricity production, time series, trends, quick fluctuations, machine learning, multilayer perceptron, big data.**

I. Introduction

A. Generalities

The aim of this communication is twofold:

1) It is devoted to solar irradiance and irradiation forecasts during rather short time intervals. As already noted in several publications (see, *e.g.*, [30], and [26], [27]) such predictions turn out to be very useful for electricity production by some systems like thermal and photovoltaic ones, where responses to solar variations are fast and complex.
2) The above forecasts are achieved via different time series techniques which are compared. Although such an endeavor is of course not new (see, *e.g.*, [15], [20], [29]), this communication might be one of the very first which takes into account the implementation issues by evaluating quite closely their need of large historical data (see, also, [14]).

Remember that times series (TS), which play an important rôle in numerical weather prediction (see, *e.g.*, [7], [13], and the references therein), are also utilized in many other domains, like, for instance, econometrics and biology. We will not be employing here the classic setting for time series,[1] but

- various topics from computer science, like artificial neural networks, computational intelligence and machine learning (see, *e.g.*, [1], [3], [4], [16], [17], [22], [28], [33]).[2]

[1] See [23] for a most interesting account and an analysis of the connections with econometrics and finance.

[2] The literature on the application to meteorological forecasts is already quite important. Combined with a lack of space it explains the absence of any review of those approaches in this communication.

- a new viewpoint on time series, which started in financial engineering (see, *e.g.*, [9], [11]), where
  - no explicit mathematical model is needed,
  - the notions of *trends* and *quick fluctuations* are key ingredients.

Our time series were recorded every minute thanks to a meteorological station,[3] which is on the roof of the *Institut Universitaire de Technologie Nancy-Brabois*. In this paper, we are focusing on the solar irradiance in W/m^2 and irradiation in Wh/m^2.

B. Overview of the various techniques

Our TS analysis may be divided into two groups, *i.e.*, with or without a need of big data.

1) Settings without large historical data:

- The setting *without model* (WM), which was sketched in Section I-A, should be related from an engineering viewpoint to the success of *model-free control* [10].[4]
- The *persistence* (P) method [28], which is a trivial machine learning viewpoint, assumes no change between the forecast and the last measure.

2) Settings with large historical data: A rather precise modeling is needed, which is quite greedy. We are considering four cases:

- A *MultiLayer Perceptron* (MLP), which is a standard artificial neural network, is among the most popular tool for analyzing meteorological TS (see, *e.g.*, [1], [3], [31]).
- If we take into account a probabilistic description of the TS, a stationary hypothesis is often necessary. It yields a CSI-MLP, *i.e.*, a *clear sky index* in connection with the MLP in order to deal with stationary TS [32]. The broadband Solis model [24] is used: it allows generating the global irradiation without clouds. The ratio between measurement and clear sky solar radiation defines the clear sky index.
- The *scaled persistence* (SP) [32] adds to the classical persistence P a clear sky procedure like the above one. This methodology is equivalent to clear sky index persistence

After a short presentation of time series without model in Section II, various experiments are reported in Section III. Some concluding remarks may be found in Section IV.

[3]Vaisala Weather Transmitter WXT520 with an additive photosynthesis measuring probe head FLA613GS and the data loggers AHLBORN.

[4]It might be interesting here to emphasize here that this viewpoint is also useful for greenhouses [18].

II. Time series without any explicit model

A. Nonstandard analysis and the Cartier-Perrin theorem

Take the time interval $[0,1] \subset \mathbb{R}$ and introduce as often in *nonstandard analysis*[5] the infinitesimal sampling

$$\mathfrak{T} = \{0 = t_0 < t_1 < \cdots < t_N = 1\}$$

where $t_{\iota+1} - t_\iota$, $0 \leq \iota < N$, is *infinitesimal*, *i.e.*, "very small" ([5], [6]). A *time series* $X(t)$ is a function $X : \mathfrak{T} \to \mathbb{R}$.

Remark 1: The reader, who is not familiar with nonstandard analysis, should not be afraid by the wording *infinitesimal sampling*. It just means in plain words that the sampling time interval is "small" with respect to the total recording time. Let us also stress that several time scales are most natural within this formalism.

The *Lebesgue measure* on $\mathfrak{T}$ is the function ℓ defined on $\mathfrak{T}\backslash\{1\}$ by $\ell(t_i) = t_{i+1} - t_i$. The measure of any interval $[c,d]$, $0 \leq c \leq d \leq 1$, is its length $d - c$. The *integral* over $[c,d]$ of the time series $X(t)$ is the sum

$$\int_{[c,d]} X d\tau = \sum_{t \in [c,d]} X(t)\ell(t)$$

X is said to be S-*integrable* if, and only if, for any interval $[c,d]$ the integral $\int_{[c,d]} |X| d\tau$ is *limited*, *i.e.*, not infinitely large, and also infinitesimal, if $d - c$ is infinitesimal.

X is S-*continuous* at $t_\iota \in \mathfrak{T}$ if, and only if, $f(t_\iota) \simeq f(\tau)$ when $t_\iota \simeq \tau$.[6] X is said to be *almost continuous* if, and only if, it is S-continuous on $\mathfrak{T} \setminus R$, where R is a *rare* subset.[7] X is *Lebesgue integrable* if, and only if, it is S-integrable and almost continuous.

A time series $\mathcal{X} : \mathfrak{T} \to \mathbb{R}$ is said to be *quickly fluctuating*, or *oscillating*, if, and only if, it is S-integrable and $\int_A \mathcal{X} d\tau$ is infinitesimal for any *quadrable* subset.[8]

Let $X : \mathfrak{T} \to \mathbb{R}$ be a S-integrable time series. Then, according to the Cartier-Perrin theorem [2], the additive decomposition

$$\boxed{X(t) = E(X)(t) + X_{\text{fluctuat}}(t)} \tag{1}$$

holds where

- the *mean* $E(X)(t)$ is Lebesgue integrable,
- $X_{\text{fluctuat}}(t)$ is quickly fluctuating.

The decomposition (1) is unique up to an additive infinitesimal.

Remark 2: See [19] for a less demanding presentation of the Cartier-Perrin theorem.

Remark 3: The above quick fluctuations should be viewed like corrupting noises in engineering [8]. Determining the trend is therefore similar to noise attenuation. According to the mathematical definition of quick fluctuations,

[5]See, *e.g.*, [5], [6] and [19] for an introduction to this fascinating domain which is stemming from mathematical logic.

[6]$a \simeq b$ means that $a - b$ is infinitesimal.

[7]See [2] and [19] for this technical definition.

[8]A set is *quadrable* [2] if its boundary is rare.

this may be achieved by integrating on a short time window or, more generally, by any low pass filter.

B. On the numerical differentiation of a noisy signal

Let us start with the first degree polynomial time function $p_1(t) = a_0 + a_1 t, t \geq 0, a_0, a_1 \in \mathbb{R}$. Rewrite thanks to classic operational calculus (see, *e.g.*, [34]) p_1 as $P_1 = \frac{a_0}{s} + \frac{a_1}{s^2}$. Multiply both sides by s^2:

$$s^2 P_1 = a_0 s + a_1 \tag{2}$$

Take the derivative of both sides with respect to s, which corresponds in the time domain to the multiplication by $-t$:

$$s^2 \frac{dP_1}{ds} + 2sP_1 = a_0 \tag{3}$$

The coefficients a_0, a_1 are obtained via the triangular system of equations (2)-(3). We get rid of the time derivatives, i.e., of sP_1, $s^2 P_1$, and $s^2 \frac{dP_1}{ds}$, by multiplying both sides of Equations (2)-(3) by s^{-n}, $n \geq 2$. The corresponding iterated time integrals are low pass filters which, according to Remark 3, attenuate the corrupting noises. A quite short time window is sufficient for obtaining accurate values of a_0, a_1. Note that estimating a_0 yields the trend according to Remark 3.

The extension to polynomial functions of higher degree is straightforward. For derivatives estimates up to some finite order of a given smooth function $f : [0, +\infty) \to \mathbb{R}$, take a suitable truncated Taylor expansion around a given time instant t_0, and apply the previous computations. Resetting and utilizing sliding time windows permit to estimate derivatives of various orders at any sampled time instant.

Remark 4: See [12], [21] for more details.

C. Application of the above calculations

A "good" forecast $X_{\text{est}}(t+T)$, at time $t+T$, $T > 0$, of the time series $X(t)$ may be written

$$\boxed{X_{\text{est}}(t+T) = X_{\text{trend}}(t) + [\dot{X}_{\text{trend}}(t)]_e T}$$

where $[\dot{X}_{\text{trend}}(t)]_e$ denotes the derivative which is estimated as indicated in Remark 3 and Section II-B. In this paper, $T = 60$ minutes or 15 minutes.

Let us stress that

- X_{trend} and $\dot{X}_{\text{trend}}$ are calculated with different time windows, *i.e.*, respectively 10 minutes and 75 minutes.
- What we predict is the trend and not the quick fluctuations (see also [9], [11]).

III. Experiments

A. Data

We benefit from solar irradiance data, which were recorded during daylight every minute during three years, *i.e.*, 2011, 2012 and 2013. Two major types of experiments are distinguished:

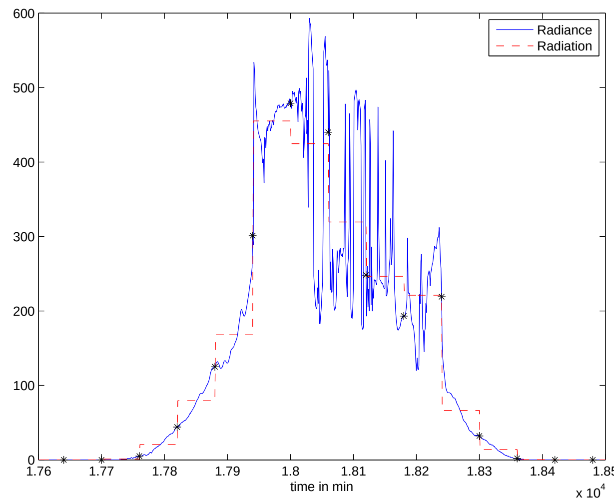

Figure 1. Irradiance in W/m^2 (blue $-$), Irradiation in Wh/m^2 (red $--$) and Irradiance each hour in W/m^2 (black $*$) on 13 February 2013

1) irradiation is obtained according to the computation of the irradiance mean every hours, thus radiation $=$ mean(radiance($t-59$), radiance($t-58$), ..., radiance(t)) with $t = 60 \times k$ minutes, $k = \{0, 1, 2, ...\}$.
2) Instantaneous measurements are used.

See the differences in Figure 1. Set 60 minutes for the forecasting horizon.

B. Irradiation forecasting

All forecasting methods, namely P, SP, MLP and CSI-MLP, use directly irradiation measures, but not WM, for which records every minute permit better performances. Figures 2 and 3 display the results for a 60min forecast. More details may be found in the first lines of Tables I and II. The best method turns out to be SP. Even if the mathematical function "clear sky" plays a crucial rôle, his simplicity is remarkable. Note also that WM method is third with respect to the normalized L^1 norm. It becomes fourth with respect to the normalized L^2 norm.

C. Irradiance forecasting

We try now to forecast the irradiance, *i.e.*, the instantaneous hourly values (see the stars $*$ in Figure 1). Contrarily to the previous calculations, averages are no more useful. Corrupting noises play therefore a much more important rôle. Only instantaneous values are used for P, SP, MLP and CSI-MLP methods. For MLP and CSI-MLP, three learning data sizes were considered, *i.e.*, one, two, and three years. The errors (normalized values) are related to the lower values computed among seven runs. Figures 4 and 5 demonstrate that not only WM yields the best results but is also more robust with respect to the noises. WM is able moreover to

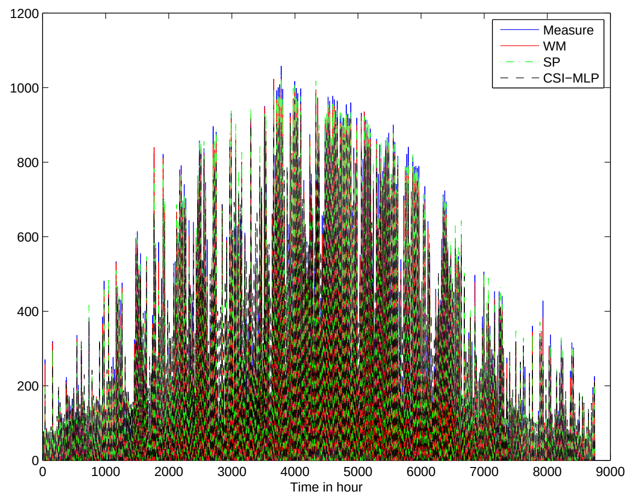

Figure 2. Irradiation forecasting in Wh/m^2, year 2013

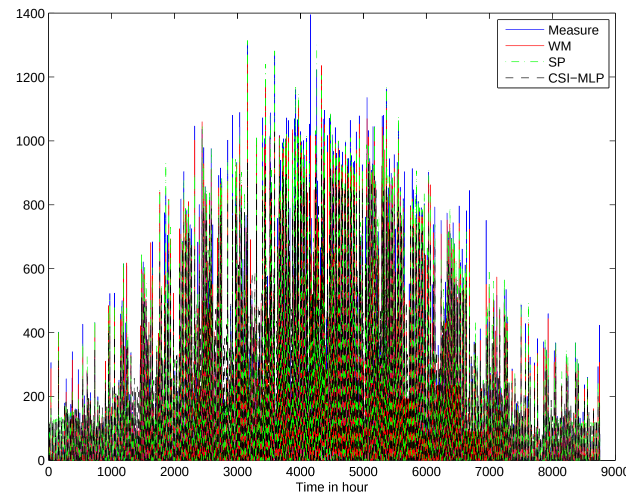

Figure 4. Radiance forecasting in W/m^2, year 2013

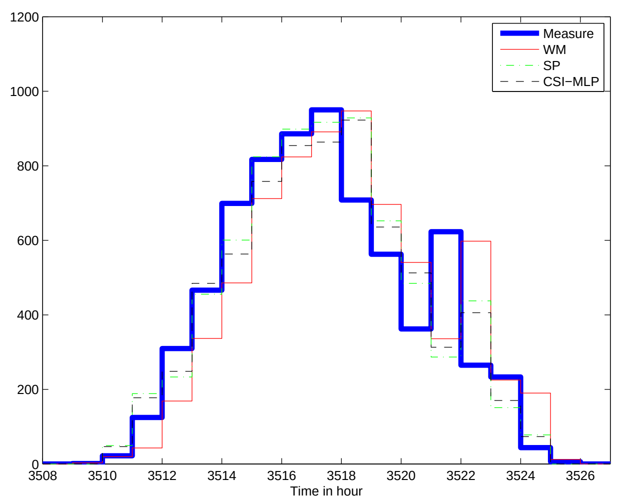

Figure 3. Zoom of Figure 2

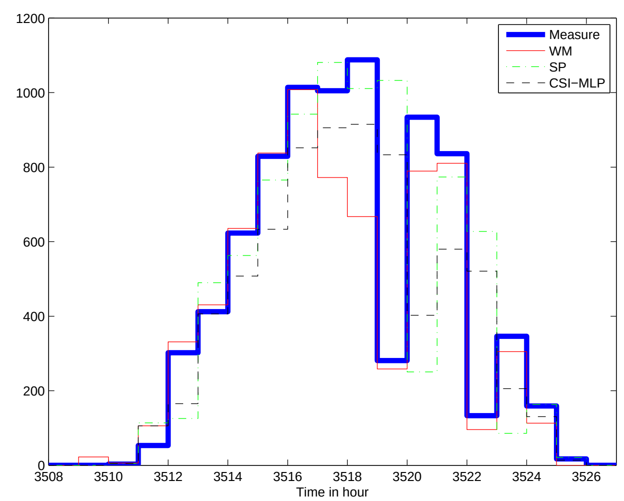

Figure 5. Zoom on figure 4

take advantage of minute data: this a true advantage. Tables I and II confirm those statements.

D. Extension to shorter time forecasts

Switch now to a 15 minutes forecast for the irradiation. The two most efficient methods are compared, *i.e.*, SP and WM. The normalized norm L^2 gives, for a time step of 5 minutes and a 15 minutes time horizon, respectively for SP and WM, 0.5624 and 0.5399. The superiority of the behavior of WM may be given by the percentage 4.2%. Figures 6 and 7 show the predictor behavior predictor during one day in winter and in summer with slightly different time scales.

IV. Conclusion

It is noteworthy to stress that methods with no need of a large amount of historical data give often better results, especially with noisy time series. Further investigations will hopefully confirm and precise those preliminary results. They should moreover yield new elements for epistemological discussions on the nature of "good" weather forecasts (see, *e.g.*, [25]).

	P	WM	SP	MLP			CSI-MLP		
				1 year	2 years	3 years	1 year	2 years	3 years
Irradiation	0.3373	0.2393	**0.1962**	0.2445	0.2512	0.2432	0.2251	0.2240	0.2236
Irradiance	0.4700	**0.2665**	0.3449	0.4705	0.4153	0.4177	0.3654	0.3544	0.3575

Table I
NORMALIZED L^1 NORM OF THE FIVE PREDICTORS FOR A 60 MINUTES HORIZON AND A TIME STEP OF 60 MINUTES. BEST RESULTS IN BOLD

	P	WM	SP	MLP			CSI-MLP		
				1 year	2 years	3 years	1 year	2 years	3 years
Irradiation	0.6200	0.4952	**0.44276**	0.4763	0.4751	0.4690	0.4654	0.4626	0.44291
Irradiance	1.0177	**0.5126**	0.8943	0.8860	0.8487	0.8447	0.7917	0.7872	0.7842

Table II
NORMALIZED L^2 NORM FOR THE FIVE PREDICTORS FOR A 60 MINUTES HORIZON AND A TIME STEP OF 60 MINUTES. BEST RESULTS IN BOLD

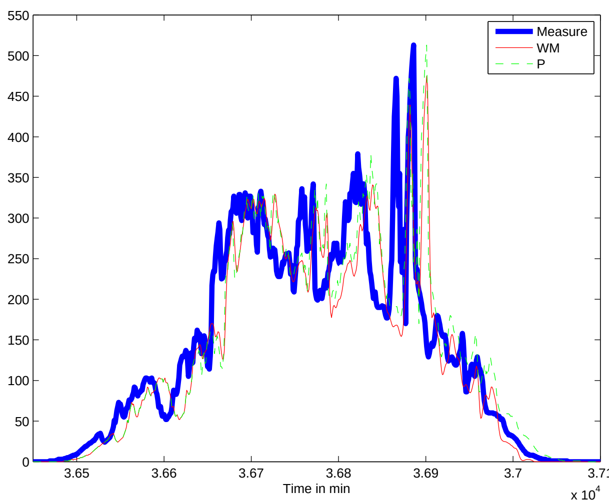

Figure 6. Winter day (W/m^2)

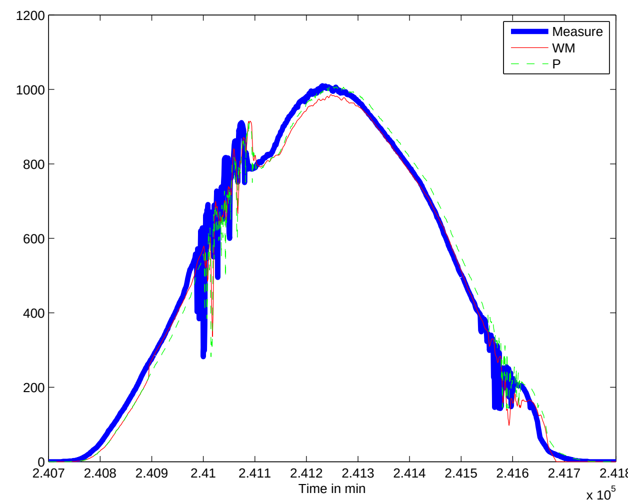

Figure 7. Summer day (W/m^2)

ACKNOWLEDGMENT

The installation of the meteorological station was made possible by the project E2D2, or *Énergie, Environnement & Développement Durable*, which is supported by the *Université de Lorraine* and the *Région Lorraine*.

REFERENCES

[1] M. A. Behrang, E. Assareh, A. Ghanbarzadeh, A. R. Noghrehabadi, "The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data," Solar Energy, vol. 84, pp. 1468–1480, 2010.

[2] P. Cartier, Y. Perrin, "Integration over finite sets," in F. & M. Diener, Eds: Nonstandard Analysis in Practice, Springer, 1995, pp. 195-204.

[3] S.F. Crone, "Stepwise selection of artificial neural networks models for time series prediction," J. Intelligent Syst., vol. 14, pp. 99–122, 2005.

[4] J. G. De Gooijer, R. J. Hyndman, "25 years of time series forecasting," Int. J. Forecasting, vol. 22, pp. 443–473, 2006.

[5] F. Diener, M. Diener, Tutorial, F. & M. Diener, Eds: Nonstandard Analysis in Practice, Springer, 1995, 1–21.

[6] F. Diener, G. Reeb, Analyse non standard, Hermann, 1989.

[7] C. Duchon, R. Hale, Time Series Analysis in Meteorology and Climatology: An Introduction, Wiley-Blackwell, 2012.

[8] M. Fliess, "Analyse non standard du bruit," C.R. Acad. Sci. Paris Ser. I, vol. 342, pp. 797–802, 2006

[9] M. Fliess, C. Join, "A mathematical proof of the existence of trends in financial time series," in A. El Jai, L. Afifi, E. Zerrik, Eds: Systems Theory: Modeling, Analysis and Control, Presses Universitaires de Perpignan, 2009, pp. 43–62. Available at http://hal.archives-ouvertes.fr/inria-00352834/en/

[10] M. Fliess, C. Join, "Model-free control," Int. J. Control, vol. 86, pp. 2228–2252, 2013. Available at http://hal.archives-ouvertes.fr/hal-00828135

[11] M. Fliess, C. Join, F. Hatt, "A-t-on vraiment besoin d'une approche probabiliste en ingénierie financière?," in Proc. Conf. Médit. Ingénierie Sûre Systèmes Complexes, Agadir, 2011. Available at http://hal.archives-ouvertes.fr/hal-00585152/en/

[12] M. Fliess, C. Join, H. Sira-Ramírez, "Non-linear estimation is easy," Int. J. Modelling Identif. Control, vol. 4, pp. 12–27, 2008. Available at http://hal.archives-ouvertes.fr/inria-00158855/en/

[13] F. O. Hocaoglu, F. Karanfil, "A time series-based approach for renewable energy modeling," Renew. Sustain. Energy Rev., vol. 28, pp. 204–214, 2013.

[14] L. Hontoria, J. Aguilera, P. Zufiria, "An application of the multilayer perceptron: Solar radiation maps in Spain," Solar Energy, vol. 79, pp. 523–530, 2005.

[15] M. Hossain, A. Oo, A. Ali, "Hybrid prediction method for solar power using different computational intelligence algorithms," Smart Grid Renew. Energy, vol. 4, 2013, pp. 76–87; 2013.

[16] N.K. Kasabov, Q. Song, "DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction," IEEE Fuzzy Syst., vol. 10, pp. 144–154, 2002.

[17] S. Kalogirou, "Artificial neural networks in renewable energy systems applications: A review," Renew. Sustain. Energy Rev., vol. 5, pp. 373–401, 2001.

[18] F. Lafont, J.-F. Balmat, N. Pessel, M. Fliess, "Model-free control and fault accommodation for an experimental greenhouse," in Proc. Int. Conf. Green Energy Environ. Engin., Sousse, 2014. Available at http://hal.archives-ouvertes.fr/hal-00978226/en/

[19] C. Lobry, T. Sari, "Nonstandard analysis and representation of reality," Int. J. Control, vol. 81, pp. 519–536, 2008.

[20] L. Martín, L.F. Zarzalejo, J. Polo, A. Navarroa, R. Marchante, M. Cony, "Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning," Solar Energy, vol. 84, pp. 1772–1781, 2010

[21] M. Mboup, C. Join, M. Fliess, "Numerical differentiation with annihilators in noisy environment," Num. Algo., vol. 50, pp. 439–467, 2009.

[22] A. Mellit, S. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," Prog. Energy Combustion Sci., vol. 34, pp. 574–632, 2008.

[23] V. Meuriot, Une histoire des concepts des séries temporelles, Harmattan–Academia, 2012.

[24] R.W. Mueller, K.F. Dagestad, P. Ineichen, M. Schroedter-Homscheidt, S. Cros, D. Dumortier, R. Kuhlemann, J.A. Olseth, G. Piernavieja, C. Reise, L. Wald, D. Heinemann, "Rethinking satellite-based solar irradiance modelling: The SOLIS clear-sky module," Remote Sensing Environ., vol. 91, pp. 160–174, 2004.

[25] A.H. Murphy, "What is a good forecast? An essay on the nature of goodness in weather forecasting," Weather Forecast., vol. 8, pp. 281–293, 1993.

[26] G. Notton, C. Paoli, S. Diaf, "Estimation of tilted solar irradiation using artificial neural networks," Energy Proc., vol. 42, pp. 33–42, 2013.

[27] G. Notton, C. Paoli, L. Ivanova, S. Vasileva, M.-L. Nivet, "Neural network approach to estimate 10-min solar global irradiation values on tilted planes," Renewable Energy, vol. 50, pp. 576–584, 2013.

[28] C. Paoli, C. Voyant, M. Muselli, M.-L. Nivet, "Forecasting of preprocessed daily solar radiation time series using neural networks," Solar Energy, vol. 84, pp. 2146–2160, 2010.

[29] G. Reikard, "Predicting solar radiation at high resolutions: A comparison of time series forecasts," Solar Energy, vol. 83, pp. 342–349, 2009.

[30] G. Vijayakumar, M. Kummert, S.A. Klein, W.A. Beckman, "Analysis of short-term solar radiation data," Solar Energy, vol. 79, pp. 495–504, 2005.

[31] C. Voyant, M. Muselli, C. Paoli, M.-L. Nivet, "Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation," Energy, vol. 36, pp. 348–359, 2011.

[32] C. Voyant, C. Paoli, M. Muselli, M.-L. Nivet, "Multi-horizon solar radiation forecasting for Mediterranean locations using time series models," Renew. Sustain. Energy Rev., vol. 28, pp. 44–52, 2013.

[33] A. Yona, T. Senjyu, T. Funabashi, P. Mandal, C.-H. Kim, "Decision technique of solar radiation prediction applying recurrent neural network for short-term ahead power output of photovoltaic system," Smart Grid Renew. Energy, vol. 4, pp. 32–38, 2013.

[34] K. Yosida, Operational Calculus (translated from the Japanese), Springer, 1984.