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

Cédric Join\textsuperscript{1,2,3}, Cyril Voyant\textsuperscript{4,5}, Michel Fliess\textsuperscript{2,6},
Marc Muselli\textsuperscript{5}, Marie-Laure Nivet\textsuperscript{5}, Christophe Paoli\textsuperscript{7} and Frédéric Chaxel\textsuperscript{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,\textsuperscript{1} but

\begin{itemize}
  \item 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]).\textsuperscript{2}
\end{itemize}

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

\textsuperscript{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, 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/m² and irradiation in Wh/m².

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].
• 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.

II. TIME SERIES WITHOUT ANY EXPLICIT MODEL

A. Nonstandard analysis and the Cartier-Perrin theorem

Take the time interval [0, 1] ⊂ R and introduce as often in nonstandard analysis the infinitesimal sampling

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

where \( t_{i+1} - t_i, 0 \le i < N, \) is infinitesimal, i.e., “very small” ([5], [6]). A time series \( X(t) \) is a function \( X: \mathbb{T} \to \mathbb{R} \).

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 \( \mathbb{T} \) is the function \( \ell \) defined on \( \mathbb{T} \setminus \{1\} \) by \( \ell(t_i) = t_{i+1} - t_i \). The measure of any interval \([c, d], 0 \le c \le d \le 1,\) 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 \( t \in \mathbb{T} \) if, and only if, \( f(t_i) \simeq f(\tau) \) when \( t_i \simeq \tau \). \( X \) is said to be almost continuous if, and only if, it is \( S \)-continuous on \( \mathbb{T} \setminus R \), where \( R \) is a rare subset.

\( X \) is Lebesgue integrable if, and only if, it is \( S \)-integrable and almost continuous.

A time series \( X': \mathbb{T} \to \mathbb{R} \) is said to be quickly fluctuating, or oscillating, if, and only if, it is \( S \)-integrable and \( \int_A X \, d\tau \) is infinitesimal for any quadrable subset.

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

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

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,

3Vaisala Weather Transmitter WXT520 with an additive photosynthesis measuring probe head FLA613GS and the data loggers AHLBORN.
4It might be interesting here to emphasize here that this viewpoint is also useful for greenhouses [18].
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
\( p_1(t) = a_0 + a_1t, \ t \geq 0, \ a_0, a_1 \in \mathbb{R}. \) Rewrite thanks to classical
operational calculus (see, e.g., \([34]\)) \( p_1 = \frac{ds}{dt}. \) Multiply both sides by \( s^2: \)
\[ s^2 P_1 = a_0 s + a_1 \]  
(2)

Take the derivative of both sides with respect to \( s, \) which corresponds in the time domain to the multiplication by \(-t: \)
\[ s^2 \frac{dP_1}{ds} + 2sP_1 = a_0 \]  
(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 \( s^2 P_1, \ s^2 P_1, \) 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 \( f : [0, +\infty) \rightarrow \mathbb{R}, \) take a
suitable truncated Taylor expansion around a given time in-
stant \( t_0, \) 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 \( X_{est}(t + T), \) at time \( t + T, \ \ T > 0, \) of
the time series \( X(t) \) may be written
\[ X_{est}(t + T) = X_{trend}(t) + [\dot{X}_{trend}(t)],_c T \]
where \( [\dot{X}_{trend}(t)],_c \) 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 \( \dot{X}_{trend} \) are calculated with different time
windows, i.e., respectively 10 minutes and 75 minutes.
- What we predict is the trend and not the quick fluctu-
ations (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:

1) irradiation is obtained according to the compu-
tation of the irradiance mean every hours, thus
radiation = mean(radiance(t − 59), radiance(t −
58), ..., radiance(t)) with \( t = 60 \times k \) minutes,
\( k = \{0, 1, 2, \ldots\}. \)
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 \( L^2 \) norm. It becomes fourth
with respect to the normalized \( L^1 \) norm.

C. Irradiance forecasting

We try now to forecast the irradiance, i.e., the instanta-
eous 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
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]).
<table>
<thead>
<tr>
<th>P</th>
<th>WM</th>
<th>SP</th>
<th>MLP</th>
<th>CSI-MLP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Irradiation</strong></td>
<td>0.3373</td>
<td>0.2393</td>
<td><strong>0.1962</strong></td>
<td>0.2445</td>
</tr>
<tr>
<td><strong>Irradiance</strong></td>
<td>0.4700</td>
<td><strong>0.2665</strong></td>
<td>0.3449</td>
<td>0.4705</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>P</th>
<th>WM</th>
<th>SP</th>
<th>MLP</th>
<th>CSI-MLP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Irradiation</strong></td>
<td>0.6200</td>
<td>0.4952</td>
<td><strong>0.44276</strong></td>
<td>0.4763</td>
</tr>
<tr>
<td><strong>Irradiance</strong></td>
<td>1.0177</td>
<td><strong>0.5126</strong></td>
<td>0.8943</td>
<td>0.8860</td>
</tr>
</tbody>
</table>

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

---

**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**


