Structured, physically inspired (gray box) models versus black box modeling for forecasting the output power of photovoltaic plants

Marius Paulescu a, Marek Brabec b, c, Remus Boata d, Viorel Badescu e, f, *

a Physics Department, West University of Timisoara, V. Parvan 4, 300223, Timisoara, Romania
b Department of Nonlinear Modeling, Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodarenkou vezi 2, 182 07, Prague 8, Czech Republic
c Czech Institute of Informatics, Robotics, and Cybernetics, Czech Technical University in Prague, Zikova 1903/4, 166 36, Prague 6, Czech Republic
d Astronomical Institute of the Romanian Academy, Timisoara Astronomical Observatory, Axente Sever Square 1, 300210, Timisoara, Romania
e Candida Oancea Institute, Polytechnic University of Bucharest, Spl. Independentei 313, 060042, Bucharest, Romania
f Romanian Academy, Calea Victoriei 125, Bucharest, Romania

A R T I C L E   I N F O

Article history:
Received 2 June 2016
Received in revised form 5 October 2016
Accepted 4 January 2017
Available online 9 January 2017

Keywords:
Photovoltaic plant
Output power
Forecasting
Fuzzy model
Generalized additive model

A B S T R A C T

Two advanced models for forecasting the output power of photovoltaic plants are discussed in details: a black-box Takagi-Sugeno fuzzy model and a physically inspired, semiparametric statistical model (Generalized Additive Model, GAM) based on smoothing splines. The structure of the two models, their strengths and weaknesses, are presented. The models performance is thoroughly compared with the performance of a simple linear model tested under the frame of the European Cooperation in Science and Technology (COST) Action “Weather Intelligence for Renewable Energies”, as a benchmark used also in the forecasting exercise reported in Sperati et al. Energies 8 (2015) 9594. The models are used to forecasting the output power at time horizons of 1–72 h ahead. The data used during the COST competition are used here as input. The present study extends beyond the traditional evaluation of overall model accuracy. Detailed influences of seasonal effects, sun elevation angle and solar irradiance level upon the models performance are assessed. While the accuracy of the simple linear model is not entirely bad, it differs in important details from the two advanced forecasting models. The results show that a moderate, carefully chosen increase in model structure complexity can improve the predictive performance. Suitable penalty on model complexity can help both to enforce parsimony and improve practical forecasting abilities, to a certain extent. The physically inspired GAM comes out as the best performing model.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Unlike the power generated by the traditional power plants based on fossil or nuclear fuels, the output power of the wind and photovoltaic (PV) plants is highly variable to erratic. This variability is quantified in Ref. [1] based on measured time series of the year 2014 collected from the EU transmission system operators. The major time-dependent phenomena affecting PV plants operation are:

(1) Earth movements. They are deterministic by nature and therefore, during a clear sky operation, the output power of a PV plant can be calculated quite accurately over a variety of time scales.

(2) Cloud shading when clouds pass over the PV plant. This process is random by its nature. Since the response time of a PV plant is very short, the abrupt changes in the irradiance level due to passing clouds may induce abrupt changes in the output power. This strong variability is of concern for electric grid operators, because the unexpected and sharp changes in the output power of large PV plants may be followed by grid perturbations or damages [2].

The modern concept of intelligent grid management targets at adapting to the real time energy production, compensating both for the fluctuations of wind generators and PV plants, as well as for the fluctuations of energy demand. The operational algorithm of a smart grid includes procedures for forecasting the energy

* Corresponding author.
E-mail address: badescu@theta.termo.pub.ro (V. Badescu).

http://dx.doi.org/10.1016/j.energy.2017.01.015
0360-5442/© 2017 Elsevier Ltd. All rights reserved.
production of PV plants, playing an important role in the smooth, safe and efficient integration of these plants into the power dis-
tribution grid. Forecast horizons up to 72 h ahead are of particular
interest for grids (partially) fed by PV plants [3]. The practical
importance of this prediction horizon range is motivated both by
technical reasons (e.g. constraints related to the start-up times of
the conventional power plants used to compensate the power
supply to the grid) and energy market perspectives (see Ref. [4]
and the references therein). A framework for quantifying the integra-
tion costs of PV plants associated with subhourly variability and
uncertainty as well as day-ahead forecasting errors is reported in
Ref. [5].

Several research projects have been running all over the globe to
test new ways for the accurate forecasting of the PV plants output
power. One example of such a large scale project is the European
Cooperation in Science and Technology (COST) Action ES1002
“Weather Intelligence for Renewable Energies” (WIRE) which was
conducted between 2011 and 2014 [6], WIRE gathered members from
27 European countries and five non-COST institutions from USA,
Canada, Australia and Japan. The Action was focused on two main
lines of activity:

(1) To develop dedicated post-processing algorithms of wind
and PV power coupled with weather forecasting models and
measured data.

(2) To investigate the complex relationship between the highly
intermittent weather-dependent wind and PV power pre-
diction and the energy distribution toward the end users.

In the frame of the COST Action WIRE, a benchmarking exercise
was organized with the scope of evaluating the performance of
some state-of-the-art models of short-term forecasting the output
power of wind farms and PV plants. The aim of the exercise was to
bring together and evaluate the merits of forecasts based on
different modeling approaches and input data. The benchmarking
exercise and its results are comprehensively described in Ref. [7].
Forecasting the output power of PV plants is of concern in the
context of our paper. The participants of the COST exercise received
both meteorological and PV output power data covering periods of
1.5 and 2 years, respectively, from two Italian locations (Catania and
Milano) to train and test their forecasting models. During the
testing period, the measured power data were masked for the first
14 days of each month. The masked data were used by the orga-
nizers of the COST exercise to evaluate the models performance. An
overview of the results was published in Ref. [7].

In this paper, we shortly present the forecasting model devel-
oped by our team in the frame of the COST Action “WIRE” bench-
mark exercise. It is a very simple physically motivated regression
model which relates the output power to forecasted solar irradi-
ance and estimated solar cells temperature on the base of the
forecasted air temperature. Our previous positive experience with
forecasting the PV output power was reported in Ref. [8]. The Nu-
merical Weather Predicted (NWP) [9] data (solar irradiance and air
temperature) provided in the training set by the COST exercise
organizers have a large bias. Therefore, a first activity was to cali-
brate the NWP data against the ground measurements. The linear
model constitutes a simple and straightforward way of calibration.
Further details about the simple linear model are presented in
Section 3.1.

Two new advanced forecasting models are presented in this
paper. They were tailored to the needs of the typical PV forecasting.
Thus, both models are able to forecast the PV plants output power
at time horizons of 1–72 h ahead, as asked by grid operators in
many countries. Also, they are based on NWP forecasts of radio-
metric and meteorological data, which are outcome of several
popular regional circulation and weather prediction models. One
of these new models is a black-box model developed in the frame of
fuzzy sets theory [10]. Our previous studies showed that fuzzy sets
theory has the ability to deal with large uncertainty in data: a study
on the estimation of the global solar irradiance based on air tem-
perature is reported in Ref. [11] while in Ref. [12] the fuzzy sets
theory is used to forecast the daily global solar irradiation. This
is particularly useful when scarce time series of forecasted meteor-
ological data are available. Our other model is physically inspired,
being based on the generalized additive model (GAM) theory. Its
formulation is motivated by our previous GAM model applications
for various solar energy forecasting problems reaching from fine
time resolution modeling of binary sunshine indicator [13] to
empirical modeling of temporarily aggregated solar irradiance [14],
and also by a broader positive experience with penalized regression
in a wide range of statistical modeling/forecasting problems in
energy PV production [15], natural gas consumption [16] and public
health modeling [17]. The particular GAM model used in this paper
is purpoputely formulated in a way that favors good forecasting
performance close to the noon (where PV output is the largest) and
accentuates less the early morning and late afternoon performance
(when PV output is generally much smaller). We argue that this is
the direction appropriate for PV power plants operation.

The structure of the paper is as follows. In Section 2, the data
used during the WIRE benchmark exercise are presented. In order
to illustrate how to build a particular model, in Section 3 the models
and their development are described in detail. This part of the
paper shows how to use in practice the NWP forecasted data (solar
irradiance, air temperature and cloud amount) for prediction pur-
poses effectively. A detailed analysis of the performance of all
models is presented in Section 4. The simple linear model tested
during the WIRE exercise is used as a benchmark for assessing
the performance of the two new models. The dependence of models
accuracy in respect to the lead time value is assessed. A sensitivity
analysis to season, solar irradiance level and solar irradiance angle
is performed subsequently. Discussion of the forecasting precision
and how it compares among different model follows. Section 5
contains conclusions of our comparative study.

### Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>Nominal power [kWp]</td>
</tr>
<tr>
<td>GHI</td>
<td>Global solar irradiance [W/m²]</td>
</tr>
<tr>
<td>DHI</td>
<td>Diffuse solar irradiance [W/m²]</td>
</tr>
<tr>
<td>DNI</td>
<td>Direct-normal solar irradiance [W/m²]</td>
</tr>
<tr>
<td>T</td>
<td>Air temperature at two meters high [°C]</td>
</tr>
<tr>
<td>TCELL</td>
<td>Solar cell temperature [°C]</td>
</tr>
<tr>
<td>NOCT</td>
<td>Nominal operating cell temperature [°C]</td>
</tr>
<tr>
<td>C</td>
<td>Cloud cover amount</td>
</tr>
<tr>
<td>β</td>
<td>Tilt angle of the PV modules [deg]</td>
</tr>
<tr>
<td>θ</td>
<td>Incidence angle on the PV module surface [deg]</td>
</tr>
<tr>
<td>ϕ</td>
<td>Geographical latitude [deg]</td>
</tr>
<tr>
<td>h</td>
<td>Sun elevation angle [deg]</td>
</tr>
<tr>
<td>δ</td>
<td>Sun declination angle [deg]</td>
</tr>
<tr>
<td>Xm</td>
<td>Suffix m indicates that X is a measured quantity</td>
</tr>
<tr>
<td>Xf</td>
<td>Suffix f indicates that X is a forecasted quantity</td>
</tr>
</tbody>
</table>
| XA     | Suffix A indicates that X is a forecasted quantity by

NWP adjusted to measured data |
| LT     | Lead time [hours] |
دریافت فوری
متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات