Journal of Atmospheric and Solar-Terrestrial Physics 105-106 (2013) 119–134

Contents lists available at ScienceDirect

Journal of Atmospheric and Solar-Terrestrial Physics

journal homepage: www.elsevier.com/locate/jastp

Assessing the performance of solar radiation computing models and model selection procedures

Viorel Badescu [a,b,*]

[a] Candida Oancea Institute, Polytechnic University of Bucharest, Spl. Independentei 313, Bucharest 060042, Romania
[b] Romanian Academy, Calea Victoriei 125, Bucharest, Romania

ARTICLE INFO

Article history:
Received 17 February 2013
Received in revised form
25 July 2013
Accepted 9 September 2013
Available online 18 September 2013

Keywords:
Solar radiation models
Statistical accuracy indicators
Overall accuracy scores

ABSTRACT

This work has two main objectives. First, the redundancy of statistical indicators is analyzed. Sets of statistical indicators are prepared and their redundancy is analyzed. Selection procedures and model hierarchies are discussed. Statistical indicators based on errors have to be preferred instead of indicators based on relative errors. Minimal sub-sets of statistical indicators may be defined. Two sub-sets of indicators are recommended, i.e. (i) Mean Bias Error, Mean Absolute Error and the slope s of the best-fit line and (ii) Mean Bias Error, Root Mean Square Error and s. The t-statistics and Willmott's index of agreement may be added to these sets. Second, several procedures for models selection are analyzed. Different selection procedures and/or different input databases yield different hierarchies among models of comparable performance. The problem of the "best model" seems to have no solution. A reasonable approach is to classify models in "good" and other lower performance categories.

1. Introduction

Many models have been proposed to provide solar radiation estimates for areas where measurements are not carried out, or for situations when gaps in the measurement records occurred. The correct validation and comparison of these models raise specific issues. For instance, different statistics may be used to evaluate the bias and random differences between the computed and measured data series (Iziomon and Mayer, 2002). Also, various ranking procedures can be used to compare the models' accuracy (Gueymard and Myers, 2008).

Two accuracy indicators are widely used by the community of solar radiation modelers, i.e. the mean bias error (MBE) and the root mean square error (RMSE). However, these are not the only statistical indicators used to assess models performance. The literature is quite abundant. For instance, Jeter and Balaras (1986) and Ianetz et al. (2007) have used the coefficient of determination. Willmott proposed a set of indices of agreement (Willmott 1981, 1982a, 1982b; Willmott et al., 1985). The slope s of the best-fit line has been used by Celik et al. (2010) while the Kolmogorov Smirnoff Integral has been used by Beyer et al. (2009) and Espinar et al. (2009) as part of European and International Energy Agency (IEA) tasks (MESOR, 2002–2010; IEA, 2011). Also, combinations of statistical indicators have been used by several authors (Jeter and Balaras, 1986; Jacovides and Kontoyiannis, 1995; Jacovides et al., 1996; Jacovides, 1998; Alados-Arboledas et al., 2000; Ianetz et al., 2007; Gueymard, 2012; Li et al., 2013).

A problem very often encountered in practice is the following. A researcher uses a set of statistical indicators and a set of input data to test a group of models. A performance hierarchy is induced among these models. Users select for practical implementation some of the "best" models, according to that hierarchy. However, the input data are different in practice from those used by the researcher. The question is: are these "best" models still among the "best", for the new set of input data? In most practical cases the users are not in a position to provide the answer. They are simply acting by faith.

This work has two main objectives. First, the redundancy of statistical indicators used to assess models performance is analyzed in Section 4. Sets of recommended statistical indicators are prepared and the redundancy of these sets is analyzed. Second, different models selection procedures are considered and their results are compared in Section 5. The accuracy of 54 clear sky solar global radiation models has been studied in Badescu et al. (2012). The main purpose was to test the models and classify them in "good", "good enough" and "poor". A hierarchy has been finally obtained for these set of models. In this paper we re-analyse the same 54 clear sky models, this time by using other selection procedures. The purpose is to see to what extent the models' hierarchy obtained in Badescu et al. (2012) is conserved or not.

Three features make the present approach to be unique: we use the largest set of clear sky models ever studied (54), we use a large

* Tel.: +40 21 402 9339; fax: +40 21 318 1019.
E-mail address: badescu@theta.termo.pub.ro

1364-6826/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jastp.2013.09.004

number of statistical indicators (15) and the models are tested by using a large set of different input meteorological databases (42).

2. Set of clear sky solar radiation models

A set of 54 broadband models for the prediction of clear sky solar global irradiance on horizontal surfaces are tested here. These models are listed in Table 1 using a call number (G001–G054) for further reference. The models are briefly described in Badescu et al. (2012).

The 54 solar radiation models have different requirements for input data, grouped in astronomical data, geographical data, surface and column integrated meteorological data and data related to atmospheric turbidity (see Table 1 of Badescu et al. (2012)). Table 2 in that paper shows the entries needed by each model.

Table 1
Clear sky models for computing global solar irradiance. Detailed description may be found in Badescu et al. (2012).

Model number	Model name	Short description
G001	ASHRAE72	ASHRAE model 1972
G002	ASHRAE05	ASHRAE model 2005
G003	Badescu	Badescu model
G004	Basha	Composite model of Bashahu and Laplaze
G005	BCLSM	Model by Barbaro et al.
G006	Biga	Biga and Rosa model
G007	Bird	Bird model
G008	CEM	Atwater and Ball model
G009	Chandr	Chandra model
G010	CLS	Cloud Layer Sunshine model by Suckling and Hay
G011	CPCR2	CPCR2 model
G012	Dognio	Dogniaux's model
G013	DPPLT	Daneshyar–Paltridge–Proctor model
G014	ESRA1	ESRA model – first version
G015	ESRA2	ESRA model – second version
G016	ESRA3	ESRA model – third version
G017	ESRA4	ESRA model – fourth version
G018	HLJ	Hottel model
G019	Ideria	Ideriah's model
G020	Ineich	Ineichen's model
G021	IqbalA	Iqbal's Model A
G022	IqbalB	Iqbal's Model B
G023	IqbalC	Iqbal's Model C
G024	Josefs	Model by Josefsson
G025	KASM	Modified Kasten model
G026	Kasten	Kasten model
G027	King	King and Buckius model
G028	KZHW	Model by Krarti
G029	MAC	McMaster model
G030	Machlr	Model by Machler and Iqbal
G031	METSTAT	METSTAT model
G032	MRM4	MRM model version 4
G033	MRM5	MRM model version 5
G034	Nijego	Nijegorodov et al. model
G035	NRCC	NRCC model
G036	Paltri	Paltridge and Platt empirical model
G037	Perrin	Perrin de Brichambaut model
G038	PR	Psiloglou revised model
G039	PSIM	PSI Model
G040	REST250	Gueymard's REST2 model version 5.0
G041	Rodger	John Page's team model
G042	RSC	Composite model described by Carroll
G043	Santam	Santamouris model
G044	Schulz	Schulze model
G045	Sharma	Sharma and Pal empirical model
G046	Watt	Watt model
G047	WKB	Wesely and Lipschutz model
G048	Yang	Yang model
G049	Zhang	Model by Zhang
G050	HS	Combination Hourwith–Schulze models
G051	ABCGS	Adnot–Bourges–Campana–Gicquel–Schulze model
G052	Paulescu	Model by Paulescu and Schlett
G053	Janjai	Model by Janjai
G054	REST281	Gueymard's REST2 model version 8.1

Meteorological data measured during 2009 at two Romanian meteorological stations (Cluj and Bucharest) are used as input during this study. Details about measurement procedures may be found in Badescu et al. (2012). Satellite derived data such as ground albedo, Angstrom turbidity coefficients, aerosol single-scattering albedo and aerosol optical depth are also used as input. Since the input datasets come from various sources, compatibility procedures have been applied in Badescu et al. (2012).

The input data have been organized in several input databases, corresponding to different situations that model users may encounter in practice. There are 21 input databases for Cluj and a similar number for Bucharest, i.e. a total of 42 input databases. Despite data from two sites only are used, the large number of databases makes the results able to be generalized. The set of models has been run for each input database in part. Each run is called *a stage*. A first run (i.e. stage 1) consisted of checking the procedure. Other runs have been also performed. In Badescu et al. (2012), 21 runs for Cluj (i.e. stage 2–stage 22) and 21 runs for Bucharest (stage 32–stage 52) has been analyzed. The same 42 input databases and 42 stages are analyzed in this work. Stage 2 for Cluj and stage 32 for Bucharest are associated to the most accurate available input data.

Computed solar global irradiance values have been compared with measurements. The solar radiation is measured by using CM6B Kipp & Zonen radiometers in Cluj and Kipp & Zonen CM11 radiometers at Bucharest. Details about the measurement uncertainty can be found in Badescu et al. (2012). The device calibration, the measurement methodology and the maintenance are provided by standard procedures prepared at the Romanian National Meteorological Administration, as described in Badescu et al. (2012).

Two accuracy indicators have been used in Badescu et al. (2012), the mean bias error (MBE) and the root mean square error (RMSE). Results concerning the performance of all 54 models when applied at Cluj and Bucharest for all 42 input databases may be found in Badescu et al. (2012, 2013).

3. Statistical accuracy indicators

The statistical accuracy indicators considered here are classified as: (i) indicators based on error measures and (ii) other indicators. Both categories are described next. Examples of indicators for overall accuracy scores are also shown.

3.1. Statistical indicators based on error measures

An individual error is a measure for the difference between a predicted value and the corresponding "true value". The true value is unknown but measured values act usually as approximations of the true value.

One considers n couples of measured and computed values, denoted m_i and c_i ($i=1,n$), respectively. Also, the mean values of measured and computed values are defined as

$$\overline{m} \equiv \frac{1}{n} \sum_{i=1}^{n} m_i \tag{1}$$

$$\overline{c} \equiv \frac{1}{n} \sum_{i=1}^{n} c_i \tag{2}$$

Two ways of defining the error may be envisaged, denoted (i) and (ii) below.

(i) The first way of defining the error e_i of the ith computed value is:

$$e_i \equiv c_i - m_i \tag{3}$$