Available online at www.sciencedirect.com

ScienceDirect

Solar Energy xxx (2015) xxx–xxx

SOLAR ENERGY

www.elsevier.com/locate/solener

Review

Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance

Christian A. Gueymard [a,*], Jose A. Ruiz-Arias [b]

[a] *Solar Consulting Services, Colebrook, NH, USA*
[b] *Department of Applied Physics I, University of Málaga, Málaga, Spain*

Received 15 September 2015; received in revised form 4 October 2015; accepted 10 October 2015

Communicated by: Associate Editor Jan Kleissl

Abstract

A comprehensive evaluation study of the performance of 140 separation models selected from the literature to predict direct normal irradiance (DNI) from global horizontal irradiance (GHI) is presented here. The assessment is conducted using high-quality 1-min data of GHI and DNI at 54 research-class stations from 7 continents. The observational dataset provides (after *a posteriori* quality control) more than 25 million valid data points, thereby representing an unprecedented level of effort. The stations are grouped into 4 distinct climate zones: arid, temperate, tropical and high-albedo. To evaluate the performance of each model at each site, three summary statistics are calculated. Additionally, with the emphasis on selecting models that perform consistently well under the general conditions of each climate zone, the robustness of each model is evaluated using a few consistency criteria.

It is found that, for all models, the errors are exacerbated by cloud enhancement and high-albedo induced effects. A higher number of predictors used by a model appears to improve its performance, but not in a consistent way, since there are many exceptions. These are attributed to possible excessive model localization and/or overfitting. In general, models that consider both a variability predictor and an estimate of coincident clear-sky irradiance tend to perform better. No model performs consistently well over the high-albedo zone, even those rare ones that do consider ground albedo as a predictor. Over the arid, temperate and tropical zones, two models consistently deliver the best predictions. One of them is recommended as a "quasi-universal" model for general use for 1-min DNI prediction wherever and whenever low- to moderate-albedo conditions prevail.

Keywords: Direct–diffuse separation; DNI; Irradiance variability; Cloud enhancement; Validation; Albedo

Contents

* Corresponding author.
E-mail address: Chris@SolarConsultingServices.com (C.A. Gueymard).

http://dx.doi.org/10.1016/j.solener.2015.10.010
0038-092X/

1. Introduction

The correct design and energy performance simulation of solar power systems, as well as various different applications in other scientific fields, require precise solar radiation data in terms of both direct normal irradiance (DNI) and diffuse horizontal irradiance (DIF). (The acronym DIF is purposefully used in lieu of DHI in this context to avoid possible confusion with the latter's alternative meaning of direct horizontal irradiance.) A ubiquitous type of calculation in solar applications consists in deriving the global tilted irradiance (GTI) on the plane of array of flat-plate solar collectors, which involves the separate modeling of the direct and diffuse tilted components. The DNI/DIF separation process in such methods is typically the major source of error in GTI (Gueymard, 2009). At locations with significant solar resource, DNI is normally the dominant component, hence the importance of its correct determination. Moreover, DNI is essential for concentrating solar power (CSP) or concentrating PV (CPV) systems, since this is the only solar radiation component that they can utilize. One difficulty is that DNI observations are relatively rare, particularly compared to those of global horizontal irradiance (GHI). Hence, in most cases, DNI is derived from measured or modeled GHI by performing its "separation" or "decomposition" into its two components, DNI and DIF. This is also done systematically, for instance, to produce time series of DNI when GHI is derived from satellite imagery with the common "cloud index" method (Perez et al., 2002; Polo et al., 2014). The separation process contributes very importantly to the overall uncertainty in such databases (Cebecauer et al., 2011).

Publications proposing a statistical separation equation based on observational data have proliferated since the very first, and seminal, study from (Liu and Jordan, 1960), hereafter LJ60, now more than 55 years old. The usual lack of science and extreme localization in this class of models has pushed the adoption of strict guidelines by at least one archival journal, aimed at restricting their publication (Gueymard et al., 2009; Kasten and Duffie, 1993). Still, such models continue to be developed and used, however with a lack of evidence about which one can provide the best possible results at any specific location where no DNI or DIF measurement exists. One important difficulty here is that the current separation models are empirically derived from site-specific measurements, and cannot be attributed a precise uncertainty without extensive evaluation. Validation studies do exist (e.g., Battles et al., 2000; Bertrand et al., 2015; De Miguel et al., 2001; Dervishi and Mahdavi, 2012; Engerer, 2015; Ineichen, 2008; Jacovides et al., 2010; Karatasou et al., 2003; Kuo et al., 2014; Perez et al., 1990b; Ruiz-Arias et al., 2010; Skartveit et al., 1998; Spencer, 1982; Tapakis et al., 2015; Torres et al., 2010; Vick et al., 2012; Yao et al., 2013), but are inherently limited in scope to a small number of models and test stations. Moreover, most of them aim at validating DIF rather than DNI. The validation of DNI predicted from a larger number of models has been considered in recent studies from the present authors (Gueymard, 2010; Gueymard and Ruiz-Arias, 2014), but the number of stations was still limited in number and/or climatic conditions, thus making generalization of the results difficult.

From another perspective, the temporal resolution of solar radiation data has considerably improved since the early days of solar energy development, represented here by the LJ60 study. During the last few decades, the reporting of solar radiation data from modern radiometric stations has moved from a hourly time step to much shorter steps, generally 1- to 10-min intervals, and sometimes even shorter. In parallel, the proper energy simulation of CSP projects requires solar radiation data at time steps shorter than the customary hourly interval. This is because of the non-linear and transient effects that substantially affect

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

those systems, for which an ideal simulation time step would be of the order of 10 min or less (Hirsch et al., 2010). To respond to this demand, some commercial providers of satellite-derived irradiance time series now offer databases using a 10–15-min time step, in addition to the customary hourly frequency. At these sub-hourly time scales, they still use some of the separation models that are the object of this study. In the case of photovoltaic (PV) systems, ramping effects (which are caused by rapid cloud-induced irradiance fluctuations) are being studied with time steps of 3-s or shorter (Cronin et al., 2013; Sengupta and Keller, 2012). These various developments underline the present need for DNI data at much shorter intervals than the conventional hourly time scale. Since sub-hourly measured DNI datasets are still scarce, reliable estimates based on GHI observations are required, most generally.

Considering this state of affairs, a systematic validation study of the currently available separation models of the literature over a large range of climatic conditions at sub-hourly time scale appears desirable. With only a few exceptions (e.g., Engerer, 2015), these models have been developed using hourly data. A thorough literature search has returned 140 separation models, which are all analyzed here. This constitutes an unprecedented level of effort compared to earlier validation studies—typically an order of magnitude more models. This study describes the models, the sources of data, the validation methodology, and provides a series of results using various statistical indicators, with an emphasis on site-to-site consistency.

To emphasize the trend toward shorter time steps described above, the present study considers 1-min time steps, using solar irradiance observations from high-quality research-class stations. Data from 54 sites in widely diverse climatic environments over seven continents and at various elevations are used, resulting in more than 25 million valid data points. This is another unprecedented level of effort, aimed at providing as general conclusions as possible.

The objectives of this study are multiple: (i) Validate the prediction of DNI using existing separation models and a 1-min data time step; (ii) Determine which model(s) could potentially be of general or broad validity; (iii) Determine whether models with a greater number of input variables (predictors) systematically outperform models with less predictors; (iv) Investigate the impact of high GHI values on the DNI prediction accuracy; and (v) Evaluate whether or how the performance results follow any pattern related to specific geographic, climatic or environmental conditions. Lastly, this study has required the development of a large database of high-quality 1-min measurements, which could be used in the near future for further model developments.

2. Separation models

A first question that needs to be addressed is whether 1-min DNI predictions can be reliably obtained using separation models that were specifically designed for hourly data, or if a new generation of "minutely models" would rather be necessary. A previous study (Gueymard and Ruiz-Arias, 2014) showed that, indeed, hourly models could generally be used with 1-min data, but also that caution needed to be exerted regarding two issues: (i) Random errors are much larger than with hourly data (which could be expected); and (ii) Hourly models do not generally correctly support transient situations when GHI is exceptionally large or even surpass the extraterrestrial value due to "cloud enhancement" effects, which may be frequent in 1-min data. These important issues are investigated here in more detail. It is worth noting that 1-min separation models have started to appear only recently, and are currently less than a handful. It is thus critical to have the possibility to rely on hourly models to generate 1-min time results until more specific minutely models appear in the literature. It is stressed, however, that the comparative model performance results presented here might have been different in relative terms if the models had been used with hourly data.

2.1. Models of the literature

The literature shows that different kinds of separation models have been proposed for sub-hourly, hourly, daily and monthly time scales. The two first models ever proposed (Liu and Jordan, 1960) were for *daily* and *monthly* time scales, separately. Interestingly, the *daily* LJ60 model has been frequently used later for *hourly* energy simulations (e.g., Heller and Dahm, 1999; Shen et al., 2008), particularly with the TRNSYS computer simulation tool (Klein et al., 1975). For that reason, and also because of its stature and historical significance, it is considered here as a benchmark, which can be used to evaluate the progress made during the intervening 55 years. Consequently, in what follows, the literature has been extensively searched for separation models using GHI at *hourly* or *sub-hourly* time steps. (No daily- or monthly-type model is considered here, with the exception of that of LJ60.)

In most cases, the diffuse fraction, K, (i.e., the DIF/GHI ratio) is first evaluated from the clearness index, K_T. The latter is the ratio of GHI to its extraterrestrial counterpart, calculated here from a recent determination of the solar constant, 1361.2 W/m^2 (Gueymard, 2012b). The component separation avenue just described is frequently referred to as the "diffuse fraction" approach. As an example, consider the LJ60 model first. No equation was provided for the curve shown in Fig. 7 of the original paper (Liu and Jordan, 1960). After digitization of this figure, the following best-fit equation is obtained:

$$K = 1 + 0.006381\ K_T - 3.2315\ K_T^2 + 2.2448\ K_T^3 + 0.081882\ K_T^4,\ K_T \le 0.75$$

$$K = 0.16, K_T > 0.75. \quad (1b)$$

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

In what follows, this model will also be refereed to as Liu. More generally, for clarity and conciseness, all models will appear as an acronym with a consistent convention: First author in small caps, optionally followed by a number if there is more than one model attributed to this author.

Alternatively, other models rather use K_T to evaluate DNI directly. The first model of this type to appear in the open literature can be attributed to Boes (1975), who generalized an earlier equation developed at Aerospace Corporation, such as:

$$\mathrm{DNI} = 0, K_T < 0.30 \quad (2a)$$

$$\mathrm{DNI} = \mathrm{Max}(0, -550 + 1790\ K_T), 0.30 < K_T \leq 0.85 \quad (2b)$$

$$\mathrm{DNI} = 1000, K_T > 0.85 \quad (2c)$$

where, from now on, DNI is expressed in W/m^2.

Later "direct fraction" models rather evaluated the direct transmittance of the atmosphere, K_n (i.e., DNI normalized by its extraterrestrial counterpart) from K_T, as an intermediate step. This latter approach is preferable to that of Boes since it automatically takes care of the sun–earth distance seasonal variation, which otherwise introduces systematic errors in equations similar to Eq. (2) that do not normalize DNI. All models proposed since Boes use functions of the type $\mathrm{Kn} = \mathrm{f}(K_T)$, with the exception of Boland5 (see Table 2), which is of the type $\mathrm{DNI} = \mathrm{f}(K_T)$, like Boes.

As far as the determination of DNI is concerned, the use of "diffuse fraction" or "direct fraction" models is mathematically equivalent by consequence of the fundamental closure relationship: $K_n = K_T(1 - K)$.

2.2. Model selection

Following a thorough literature search, 139 separation models have been found since the pioneering LJ60 in scientific journals, conference proceedings or reports, which is an indication of the importance of this topic, and of its vitality since the 1960s. This large number does not mean it reflects an exhaustive inventory, since other models have most likely been proposed in semi-public reports or local conferences and might not appear in a literature search. Note also that a few hourly separation models that require input variables that are difficult to obtain or are based on subjective human observations, such as sunshine duration or cloud oktas, have been deliberately excluded.

No new model is proposed here to maintain the focus on the large volume of existing information, with however one exception. The model noted Louche2 is a simple modification to the original Louche1 (Louche et al., 1991), proposed here to correct a problem at high K_T. This problem, which would not occur with hourly data, was described recently (Gueymard and Ruiz-Arias, 2014) when Louche1 is applied to 1-min data. Most importantly, Louche1's prediction of K_n becomes unphysically negative for $K_T > 1.05$, i.e., under the transient cloud enhancement situations mentioned earlier. Considering that the original K_n function reaches a maximum of 0.7496 when $K_T = 0.8592$, the proposed Louche2 correction simply consists in keeping K_n constant at 0.7496 when $K_T > 0.8592$.

Due to space limitation, all 140 models are only briefly described in Tables 1 and 2. Most models simply use K_T as sole predictor (Table 1). Those models that additionally require other predictors are rather listed in Table 2. In the latter case, they are categorized as a function of the number of predictors. The way these additional predictors may have to be indirectly obtained in the present context, and a description of the sources of ancillary data used here, are offered in Section 3.4.

Since the pioneering work of Perez et al. (1992), an increasing number of models make use of some form of "variability index", for which different definitions exist. For simplicity, only a generic notation for this index, V, appears in Table 2. More details are provided in Section 3.5.

2.3. Model corrections

Some of the models examined here have been found to suffer from typographical errors in their published equations or coefficient values, making their predictions inconsistent or sometimes even unphysical. The corrections were directly obtained from their respective authors, as detailed in the Appendix A.

Authors of a few other models afflicted by curiously high deviations have been queried unsuccessfully, so that no correction could be applied.

3. Solar radiation data and ancillary observations

3.1. Experimental limitations

All decomposition models are empirically derived from irradiance measurements, with two notable exceptions (Hollands, 1985; Hollands and Crha, 1987), which use a simplified physical approach. All irradiance measurements are obtained with various radiometers, using various calibration processes and maintenance or quality control procedures. A recent study (Gueymard and Myers, 2009) has shown that different models of pyranometers, for instance, could generate significantly differing data series even when installed side-by-side, due to their cosine errors, optical characteristics, more or less pronounced thermal offset, etc. This means that empirical models necessarily embed the limitations and experimental errors contained in the original data. These errors can be only minimized if careful corrections and filtering are exerted on the raw data first. Such refinements have been attempted in various ways in recent studies, at least to some extent; e.g., (Boland et al., 2008; Clarke et al., 2007; De Miguel et al., 2001; Muneer et al., 2007; Posadillo and Lopez Luque, 2009; Ruiz-Arias et al., 2010; Tapakis et al., 2015; Younes et al., 2005). It is however difficult to know whether the various levels of quality control in current use are

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

Table 1
Description of separation models whose sole predictor is K_T.

Acronym	Author	Notes	Acronym	Author	Notes
ALRIAHI	Al-Riahi et al. (1992)		LOPEZ3	Lopez et al. (2000)	"Model 3"
BAKHSH	Bakhsh et al. (1985)		LOUCHE1	Louche et al. (1991)	
BOES	Boes (1975)	Eq. (2) in this paper	LOUCHE2	Louche et al. (1991); this work	Correction of LOUCHE1 for $K_T > 0.8592$
BOLAND1	Boland et al. (2001)		MADUEKWE1	Maduekwe and Chendo (1997)	
BOLAND3	Boland et al. (2008)	"Average" from their Table 1	MONDOL1	Mondol et al. (2005)	
BOLAND4	Boland and Ridley (2008) and Boland et al. (2008)	Eq. (32) in their paper	MONDOL2	Mondol et al. (2008)	
BOURGES	Bourges (1992)		MORENO	Moreno et al. (2009)	
BRUNO	Bruno (1978)		MUNEER1	Muneer et al. (1984)	
CHANDRASEKARAN	Chandrasekaran and Kumar (1994)		MUNEER2	Muneer and Saluja (1986)	
CHENDO1	Chendo and Maduekwe (1994)		MUNEER3	Muneer et al. (1997)	
CHIKH1	Chikh et al. (2012)	"For Bechar"	OLIVEIRA	Oliveira et al. (2002)	
CHIKH2	Chikh et al. (2012)	"For Tamanrasset"	ORGILL	Orgill and Hollands (1977)	
CHIKH3	Chikh et al. (2012)	"For Alger"	PAGOLA1	Pagola et al. (2009)	"Orgill correlation"
CIBSE	CIBSE (2002)		PAGOLA2	Pagola et al. (2009)	"Erbs correlation"
DEMIGUEL	De Miguel et al. (2001)		PAGOLA4	Pagola et al. (2009)	"Louche correlation"
ELMINIR1	Elminir et al. (2007)	"For Cairo"	PEREZBURGOS	Perez-Burgos et al. (2014)	
ELMINIR2	Elminir et al. (2007)	"For Aswan"	POSADILLO1	Posadillo and Lopez Luque (2009)	Eq. (10); see Appendix A
ELMINIR3	Elminir et al. (2007)	"For South Valley"	POSADILLO3	Posadillo and Lopez Luque (2009)	Eq. (11); see Appendix A
ERBS	Erbs et al. (1982)		REINDL1	Reindl et al. (1990)	
FURLAN	Furlan and Oliveira (2008)		RERHRHAYE	Rerhrhaye et al. (1995)	
GONZALEZ1	González and Calbó (1999)	"Model T1"	RUIZARIAS1	Ruiz-Arias et al. (2010)	"Model G0"
HAWLADER	Hawlader (1984)		SANCHEZ	Sanchez et al. (2012)	
HIJAZIN	Hijazin (1998)		SOARES	Soares et al. (2004)	
HOLLANDS1	Hollands (1985)		SOLMET	SOLMET (1977)	
INEICHEN1	Ineichen et al. (1984)	Fig. 10 in their paper	SPENCER	Spencer (1982)	
INEICHEN2	Ineichen et al. (1984)	Fig. 11 in their paper	TAPAKIS1	Tapakis et al. (2015)	Their Table 4
INEICHEN3	Ineichen et al. (1984)	Fig. 12 in their paper	TORRES1	Torres et al. (2010)	"Model 1"; see Appendix A
JACOVIDES	Jacovides et al. (2006)		TORRES2	Torres et al. (2010)	"Model 2"
JANJAI	Janjai et al. (2010)		TORRES3	Torres et al. (2010)	"Model 3"
JETER	Jeter and Balaras (1986)		TORRES4	Torres et al. (2010)	"Model 4"
KARATASOU	Karatasou et al. (2003)		TSUBO1	Tsubo and Walker (2003)	"Model 1"
LAM1	Lam and Li (1996)	Whole year model	TSUBO2	Tsubo and Walker (2003)	"Model 2"
LAM2	Lam and Li (1996)	Whole year "Hybrid" model	TSUBO3	Tsubo and Walker (2003)	"Model 3"
LEE	Lee et al. (2013)		ULGEN	Ulgen and Hepbasli (2002)	
LI	Li and Lam (2001)	Whole year model	YAO1	Yao et al. (2013)	"Model 1"
LIU	Liu and Jordan (1960)	Eq. (1) in this paper	YAO2	Yao et al. (2013)	"Model 2"
LOPEZ1	Lopez et al. (2000)	"Model 1"	YAO3	Yao et al. (2013)	"Model 3"
LOPEZ2	Lopez et al. (2000)	"Model 2"	YAO4	Yao et al. (2013)	"Model 4"

sufficient, or conversely too stringent (thus eliminating too many good data points). Moreover, the systematic biases due to the thermal offset of pyranometers do not seem to have been corrected in any study.

To evaluate the magnitude of the observational interference in empirical decomposition modeling, a special experiment is conducted here, in which two independent series of measurements obtained with side-by-side instruments are compared. The Solar Radiation Research Laboratory of the National Renewable Energy Laboratory, which is located in Golden, Colorado (latitude 39.7424°N, longitude 105.1787°W, elevation 1830 m; http://www.nrel.gov/midc/srrl_bms/), is probably unique in the world by the large number of radiometers it maintains, and by the long periods of record these instruments have generated until now. This creates the unique possibility to compare the

Please cite this article in press as: Gueymard, C.A., Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

Table 2
Description of separation models with predictors other than K_T. The symbols for these predictors are: C: cloud fraction; E_{bnc}: clear-sky direct normal irradiance; E_c: clear-sky global horizontal irradiance; K_{Tm} mean daily K_T; m: air mass; RH: relative humidity; t: time of day; T: dry-bulb temperature; T_{dp}: dew-point temperature; V: variability index; Z: zenith angle; β: Ðngström turbidity coefficient; ρ: surface albedo.

Acronym	Author	Variables	Notes
1 predictor			
REMUND	Remund et al. (1998)	Z	Only model not dependent on K_T
2 predictors			
BOLAND2	Boland et al. (2001)	K_T, t	
BUGLER	Bugler (1977)	K_T, E_c	Paired with Perez–Ineichen clear-sky model
CHENDO2	Chendo and Maduekwe (1994)	K_T, Z	
CLARKE	Clarke et al. (2007)	K_T, Z	
DEJONG	De Jong (1980)	K_T, Z	
ERUSIAFE	Erusiafe and Chendo (2014)	K_t, Z	
GONZALEZ2	González and Calbó (1999)	K_T, Z	"Model T2"
GONZALEZ3	González and Calbó (1999)	K_T, V	"Model T3"
GONZALEZ4	González and Calbó (1999)	K_T, V	"Model T4"
GONZALEZ5	González and Calbó (1999)	K_T, V	"Model T5"
HELBIG	Helbig et al. (2010)	K_T, Z	
HOLLANDS2	Hollands and Crha (1987)	K_T, ρ	
MACAGNAN	Macagnan et al. (1994)	K_T, Z	
MADUEKWE2	Maduekwe and Chendo (1997)	K_T, Z	
MAXWELL	Maxwell (1987)	K_T, m	Also known as "DISC" model
OUMBE	Oumbe et al. (2012)	K_T, E_c	Paired with Perez–Ineichen clear-sky model
PAGOLA3	Pagola et al. (2009)	K_T, Z	"Reindl correlation"
PEREZ3	Perez et al. (1990b)	K_T, Z	Modification of MAXWELL
POSADILLO2	Posadillo and Lopez Luque (2009)	K_T, Z	Eq. (10) in original paper
POSADILLO4	Posadillo and Lopez Luque (2010)	K_T, Z	Eq. (6) in original paper
POSADILLO5	Posadillo and Lopez Luque (2010)	K_T, V	Eqs. (15) and (16) in original paper
TUOMIRANTA3	Tuomiranta and Ghedira (2015)	K_T, m	"UAE east"
TURNER	Turner and Salim (1984)	K_T, E_c	Paired with HLJ clear-sky model
UDAGAWA	Udagawa and Kimura (1978)	K_T, Z	
WATANABE	Watanabe et al. (1983)	K_T, Z	
ZHANG	Zhang et al. (2004)	K_T, Z	
3 predictors			
CHENDO3	Chendo and Maduekwe (1994)	K_T, Z, T	See Appendix A
CUCUMO	Cucumo et al. (2006)	K_T, m, E_{bnc}	Paired with Perez–Ineichen clear-sky model
GONZALEZ6	González and Calbó (1999)	K_T, Z, V	"Model T6"
GONZALEZ7	González and Calbó (1999)	K_T, Z, V	"Model T7"
GONZALEZ8	González and Calbó (1999)	K_T, Z, V	"Model T8"
HAY	Hay (1976) and Hay and Davies (1980)	K_T, C, ρ	Paired with Perez–Ineichen clear-sky model
MADUEKWE3	Maduekwe and Chendo (1997)	K_T, Z, β	β derived from T_L (Katz et al., 1982)
POSADILLO6	Posadillo and Lopez Luque (2010)	K_T, Z, V	Eq. (20) in original paper
SKARTVEIT2	Skartveit et al. (1998)	K_T, Z, V	Fixed albedo (0.2) assumed
TAMURA	Tamura et al. (2003)	K_T, Z, V	
4 predictors			
ENGERER1	Engerer (2015)	K_T, Z, t, E_c	For clear skies; paired with Perez–Ineichen clear-sky model
PEREZ1	Perez et al. (1992)	K_T, Z, T_{dp}, V	Also known as "DIRINT" model
REINDL3	Reindl et al. (1990)	K_T, Z, T, RH	
SKARTVEIT3	Skartveit et al. (1998)	K_T, Z, V, ρ	Similar to SKARTVEIT2, but with variable albedo
5 predictors			
BOLAND5	Boland et al. (2013)	K_T, Z, t, K_{Tm}, V	See Appendix A
ENGERER2	Engerer (2015)	K_T, Z, t, E_c, V	Paired with Perez–Ineichen clear-sky model

POSADILLO7	Posadillo and Lopez Luque (2009)	K_T, Z	Eq. (9); See Appendix A	ENGERER3	Engerer (2015)	K_T, Z, t, E_c, V	Paired with Perez–Ineichen clear-sky model
REINDL2	Reindl et al. (1990)	K_T, Z		KUO1	Kuo et al. (2014)	K_T, Z, t, K_{Tm}, V	"Model 1"
RIDLEY1	Ridley et al. (2004)	K_T, t		KUO2	Kuo et al. (2014)	K_T, Z, t, K_{Tm}, V	"Model 2"
RUIZARIAS2	Ruiz-Arias et al. (2010)	K_T, m	"Model G2"; See Appendix A	KUO3	Kuo et al. (2014)	K_T, Z, t, K_{Tm}, V	"Model 3"
SKARTVEIT1	Skartveit and Olseth (1987)	K_T, Z		KUO4	Kuo et al. (2014)	K_T, Z, t, K_{Tm}, V	"Model 4"
STAUTER	Stauter and Klein (1980)	K_T, E_c	Paired with HLJ clear-sky model	LAURET	Lauret et al. (2013)	K_T, Z, t, K_{Tm}, V	
SUEHRKE	Suehrcke and McCormick (1988)	K_T, m		MAGARREIRO	Magarreiro et al. (2014)	K_T, Z, t, K_{Tm}, V	
TAPAKIS2	Tapakis et al. (2015)	K_T, Z	Their Table 5; see Appendix A	RIDLEY2	Ridley et al. (2010)	K_T, Z, t, K_{Tm}, V	
TAPAKIS3	Tapakis et al. (2015)	K_T, Z	Their Table 7; see Appendix A	YAO5	Yao et al. (2013)	K_T, Z, t, K_{Tm}, V	"Model 5"
TUOMIRANTA1	Tuomiranta and Ghedira (2015)	K_T, m	"UAE general"	**6 predictors**			
TUOMIRANTA2	Tuomiranta and Ghedira (2015)	K_T, m	"UAE inland"	PEREZ2	Perez et al. (2002)	K_T, Z, T_{dp}, V, E_c, E_{bnc}	Also known as "DIRINDEX" model; paired with Perez–Ineichen clear-sky model

long-term performance of collocated instruments. A previous study (Gueymard and Myers, 2009) concluded that the *optimal* way to obtain GHI, at least at that site, consisted in adding the horizontally-projected direct component measured with a Kipp & Zonen CH1 pyrheliometer to the diffuse component measured with a ventilated and thermal-offset-corrected Kipp & Zonen CM22 pyranometer being shaded from the sun by a tracking shading-ball assembly. (The CM22 is considered a working standard for diffuse measurements (Michalsky et al., 2007).) Another CM22 pyranometer, but unshaded, is used to provide an independent measurement of GHI, thus offering the possibility to detect any discrepancy between the two (direct and indirect) evaluations of GHI. In parallel, a less optimal method, but also less expensive and much more widespread at radiometric stations over the world, is to obtain GHI from an Eppley PSP pyranometer, and the diffuse component from another PSP equipped with a manually adjustable shadow ring.

Based on ≈0.8 million 1-min data points, Fig. 1 shows how the two experimental setups impact the $K = f(K_T)$ relationship differently. The left plot shows the GHI and DIF measurements from the set of two CM22 pyranometers. Similarly, the right plot shows the same kind of relationship, but using the set of two PSP pyranometers. The popular ERBS function for $K = f(K_T)$ is also shown to add perspective on the typical scatter that can be found in measured vs. modeled data. Note the noticeable disconnect between model and measurement when K_T is larger than ≈0.8, which will be discussed further in Section 4.5. A modification in clustering pattern is obvious between the two plots, in addition to the increased scatter around the ERBS equation when using the second source of radiation data (in the right plot).

Beside the direct impact on data quality of the radiometers' performance involved in the derivation of the empirical models or their validation, other aspects of data quality may play an even more critical role: Radiometer calibration, station maintenance, instrument cleaning, *a posteriori* quality assessment and corrections, etc. To decrease the impact of such factors on this analysis, only data from research-class stations are used here. However, the equipment of these stations varies in specifications and performance. Moreover, these stations may experience difficult periods from time to time for a number of reasons. These limitations must be acknowledged, and can only be attenuated through additional quality control, as described in Section 3.3.

3.2. Test stations

A database of high-quality 1-min irradiance measurements has been assembled from 54 stations that independently observe the three components (DNI, DIF and GHI) with thermopile radiometers. This observational redundancy is an essential condition for efficient quality control, as discussed in the next section. A description of

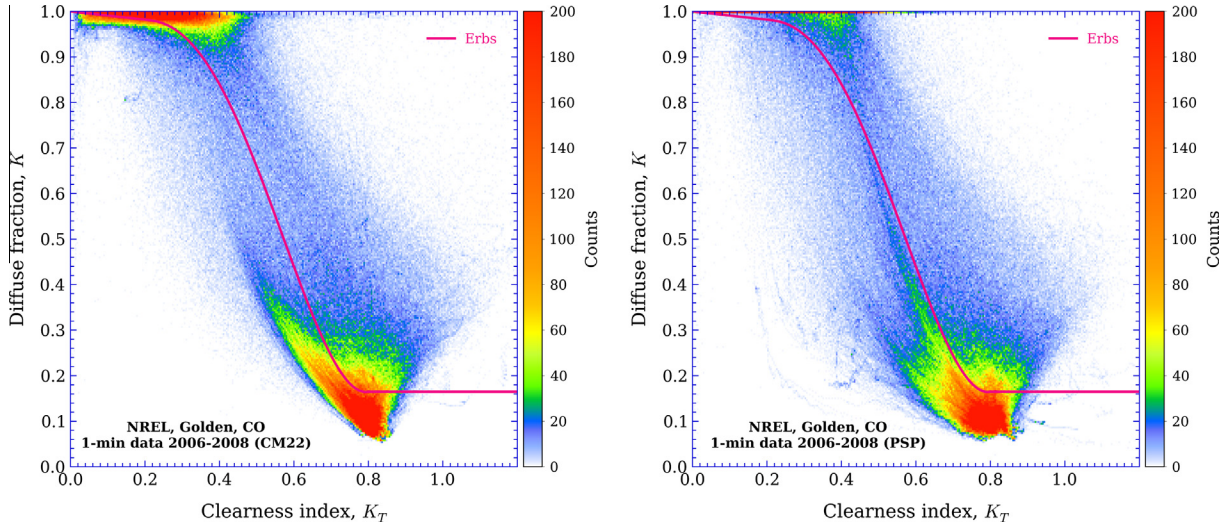

Fig. 1. Observed relationship between 1-min K_T and K at the Golden NREL site, Colorado, based on data observed with: (Left) An unshaded CM22 pyranometer for GHI and another CM22 with a tracking-shade assembly for DIF; (Right) An unshaded PSP pyranometer for GHI and another PSP with a manually adjustable shade-ring assembly for DIF. The Erbs separation function is shown for reference and further discussion.

all stations appears in Table 3, and their geographical location in Fig. 2. Note that the station selection includes extreme latitudes and high elevations. Most stations belong to the Baseline Surface Radiation Network (BSRN, http://bsrn.awi.de/; Ohmura et al., 1998), a project of the Global Energy and Water Cycle Experiment (GEWEX) under the umbrella of the World Climate Research Programme (WCRP).

At all stations, DIF is measured with a shading ball attached to the sun tracker also supporting the pyrheliometer measuring DNI. A special installation existed at Solar Village, where both a high-performance active cavity radiometer (ACR) and a pyrheliometer were used to obtain DNI. DNI was normally obtained from the ACR, except during its regular periods of self-calibration, when the pyrheliometer reading was substituted using a linear correction procedure described in Gueymard and Ruiz-Arias (2015). An ACR is a reference laboratory instrument, normally used only to transfer calibration from the World Radiometric Reference to field pyrheliometers. A windowed ACR can be deployed in the field, and is capable of measuring DNI with long-term uncertainties as low as ≈0.5% (Michalsky et al., 2011), but this never occurs in practice because of the extremely high cost of such instruments. Thus the case of Solar Village is undeniably exceptional.

For each station, the number of available individual data points was first calculated for each year of their complete period of operation. This helped select a period of three years (whenever possible) with the highest possible number of data points. In most cases, these periods were made of consecutive years during the latest 15 years. One station (PSA-DLR) has two years of data, and one station (Masdar) only one year. In the case of the seven stations code-named BON, BOU, DRA, FPE, GCR, PSU and SXF, the availability of 1-min data started in 2009 only. Hence, prior to that, their 3-min data had to be used here, which explains why these stations have comparatively less data points than others. The total number of valid data points, N, is indicated in Table 3. Valid data points are those that remained after the quality control procedure described in the next section. Table 3 also indicates the mean measured values of DNI and GHI (in W/m^2) from all valid measurements.

3.3. Data quality control

There is currently no definitive, ideal, or widely accepted procedure for optimal *a posteriori* quality control of irradiance data. Each institution typically develops its own method, which implies that some may be more stringent than others. For the present purposes, those data points not respecting the following conditions were rejected:

1. $Z < 85°$
2. GHI > 0 and DIF > 0 and DNI $\geqslant 0$
3. DNI $< 1100 + 0.03$ Elev
4. DNI $< E_{0n}$
5. DIF $< 0.95\ E_{0n} \cos^{1.2} Z + 50$
6. GHI $< 1.50\ E_{0n} \cos^{1.2} Z + 100$
7. Abs(Closr) $< 5\%$
8. DIF/GHI < 1.05 for GHI > 50 and $Z < 75°$
9. DIF/GHI < 1.10 for GHI > 50 and $Z > 75°$

where Elev is the elevation (m) from Table 3, E_{0n} is the extraterrestrial irradiance on a normal surface, calculated here with the PSA and SUNAE algorithms for the sun position and the sun–earth distance correction, respectively (Blanco-Muriel et al., 2001; Michalsky, 1988), and Closr $= 100$ [DNI cos $Z +$ DIF $-$ GHI)/GHI] is the closure error in percent. Condition (1) eliminates situations of low irradiance of marginal importance in solar applications, when moreover both instruments and models have low

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

Table 3
Information on the 54 test stations used for validation, including latitude and longitude in degrees and elevation in meters above mean sea level. Acronyms for the data sources are: BSRN (Baseline Solar Radiation Network; NREL (National Renewable Energy Laboratory); and DLR (Deutches Zentrum für Luft- und Raumfahrt, German Aerospace Center). Acronyms for climate zones are: AR (Arid), HA (High albedo), TM (Temperate), and TR (Tropical). The 3-letter station codes are those used by BSRN for stations that belong to that network. *N* is the number of valid data points. The mean measured values of GHI and DNI are in W/m^2.

Code	Station	Lat.	Long.	Elev.	Source	Period	Climate	N	Mean GHI	Mean DNI
ALE	Alert	82.490	−62.420	127	BSRN	2009–2011	HA	198,662	234.5	389.0
ASP	Alice Springs	−23.798	133.888	547	BSRN	2007–2009	AR	176,216	499.5	611.2
BER	Bermuda	32.267	−64.667	8	BSRN	2006–2008	TM	427,070	548.8	448.0
BIL	Billings	36.605	−97.516	317	BSRN	2005–2007	TM	579,418	486.5	504.1
BON	Bondville	40.067	−88.367	213	BSRN	2007–2009	TM	131,852	592.6	533.7
BOU	Boulder	40.050	−105.007	1577	BSRN	2002–2004	TM	416,471	529.8	529.3
BRB	Brasilia	−15.601	−47.713	1023	BSRN	2009–2011	TR	460,969	428.8	361.6
CAR	Carpentras	44.083	5.059	100	BSRN	2003–2005	TM	697,048	399.1	476.8
CNR	Cener	42.816	−1.601	471	BSRN	2010–2012	TM	633,183	403.6	421.0
CLH	Chesapeake Light	36.905	−75.713	37	BSRN	2011–2013	TM	651,979	405.7	417.5
COC	Cocos Island	−12.193	96.835	6	BSRN	2006–2008	TR	539,472	498.2	410.9
DOM	Concordia Station	−75.100	123.383	3233	BSRN	2005–2007	HA	368,679	399.1	846.0
DAR	Darwin	−12.425	130.891	30	BSRN	2009–2011	TR	628,940	531.5	469.6
DWN	Darwin Met Office	−12.424	130.893	32	BSRN	2000–2002	TR	653,803	481.9	450.1
DAA	De Aar	−30.667	23.993	1287	BSRN	2002–2004	AR	570,186	518.7	656.0
DRA	Desert Rock	36.626	−116.018	1007	BSRN	2007–2009	AR	183,795	688.9	794.6
EUR	Eureka	79.989	−85.940	85	BSRN	2009–2011	HA	514,930	221.3	304.1
FLO	Florianopolis	−27.533	−48.517	11	BSRN	2002, 2003, 2005	TM	166,160	446.6	438.6
FPE	Fort Peck	48.317	−105.100	634	BSRN	2007–2009	TM	126,234	583.1	576.8
FUA	Fukuoka	33.582	130.375	3	BSRN	2011–2013	TM	677,367	350.7	273.4
GVN	Georg von Neumayer	−70.650	−8.250	42	BSRN	2011–2013	HA	584,025	308.1	272.4
GOB	Gobabeb	−23.561	15.042	407	BSRN	2012–2014	AR	408,714	582.2	732.1
GOL	Golden-NREL	39.742	−105.180	1829	NREL	2006–2008	TM	640,152	458.2	558.0
GCR	Goodwin Creek	34.250	−89.870	98	BSRN	2007–2009	TM	158,949	564.3	514.6
ILO	Ilorin	8.533	4.567	350	BSRN	1995, 1999, 2000	TR	99,476	256.6	72.1
ISH	Ishigakijima	24.337	124.163	6	BSRN	2011–2013	TM	698,419	367.7	246.2
IZA	Izana	28.309	−16.499	2373	BSRN	2011–2013	AR	671,556	617.9	791.4
KWA	Kwajalein	8.720	167.731	10	BSRN	1998–2000	TR	472,581	544.0	433.5
LAU	Lauder	−45.045	169.689	350	BSRN	2005–2007	TM	660,605	371.7	402.1
LER	Lerwick	60.139	−1.185	80	BSRN	2004–2006	TM	629,312	204.6	131.0
LIN	Lindenberg	52.210	14.122	125	BSRN	2001–2003	TM	669,495	279.9	255.8
MAS	Masdar	24.442	54.617	6	Masdar	2013	AR	25,100	506.0	552.5
MNM	Minamitorishima	24.288	153.983	7	BSRN	2011–2013	TM	718,663	477.7	451.9
MAN	Momote	−2.058	147.425	6	BSRN	2008–2010	TR	615,189	504.2	383.7
NAU	Nauru Island	−0.521	166.917	7	BSRN	2005–2007	TR	595,478	539.5	439.7
NYA	Ny-Alesund	78.925	11.930	11	BSRN	2007–2009	HA	574,322	189.8	209.9
PAL	Palaiseau	48.713	2.208	156	BSRN	2009–2011	TM	676,255	304.7	280.2
PAY	Payerne	46.815	6.944	491	BSRN	2008–2010	TM	555,948	378.0	356.5
PTR	Petrolina	−9.068	−40.319	387	BSRN	2007–2009	TR	416,871	528.7	485.4
PSA	PSA-DLR	37.091	−2.358	500	DLR	2011, 2012	AR	248,680	457.0	550.2
REG	Regina	50.205	−104.713	578	BSRN	2009–2011	TM	660,686	362.9	394.6
PSU	Rock Springs	40.720	−77.933	376	BSRN	2007–2009	TM	132,528	499.5	420.1
SMS	Sao Martinho da Serra	−29.443	−53.823	489	BSRN	2006–2008	TM	594,872	435.4	444.2
SAP	Sapporo	43.060	141.328	17	BSRN	2011–2013	TM	647,992	325.4	267.6
SBO	Sede Boqer	30.860	34.779	480	BSRN	2009–2011	AR	562,982	607.2	642.9
SXF	Sioux Falls	43.730	−96.620	473	BSRN	2007–2009	TM	119,546	582.7	602.3
SOV	Solar Village	24.907	46.397	768	NREL	2000–2002	AR	563,835	567.3	580.4
SON	Sonnblick	47.054	12.958	3109	BSRN	2013–2015	HA	405,565	367.7	296.7
TAM	Tamanrasset	22.790	5.529	1385	BSRN	2006–2008	AR	585,375	589.1	622.5
TAT	Tateno	36.050	140.133	25	BSRN	2008–2010	TM	678,550	332.8	277.2
TIK	Tiksi	71.586	128.919	48	BSRN	2011–2013	HA	346,098	260.2	238.7
TOR	Toravere	58.254	26.462	70	BSRN	2010–2012	TM	384,606	239.6	263.7
TUC	Tucson	32.230	−110.955	786	NREL	2011, 2013, 2014	AR	679,106	557.3	694.9
XIA	Xianghe	39.754	116.962	32	BSRN	2008–2010	TM	445,440	420.8	376.0

accuracy. Conditions (8) and (9) do not systematically exclude data points with DIF/GHI > 1 (even though these should not occur in theory) because they can simply be caused by experimental uncertainty, particularly at low solar elevations and/or under low-irradiance conditions. Most of these cases are implicitly eliminated by condition (1) any-

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

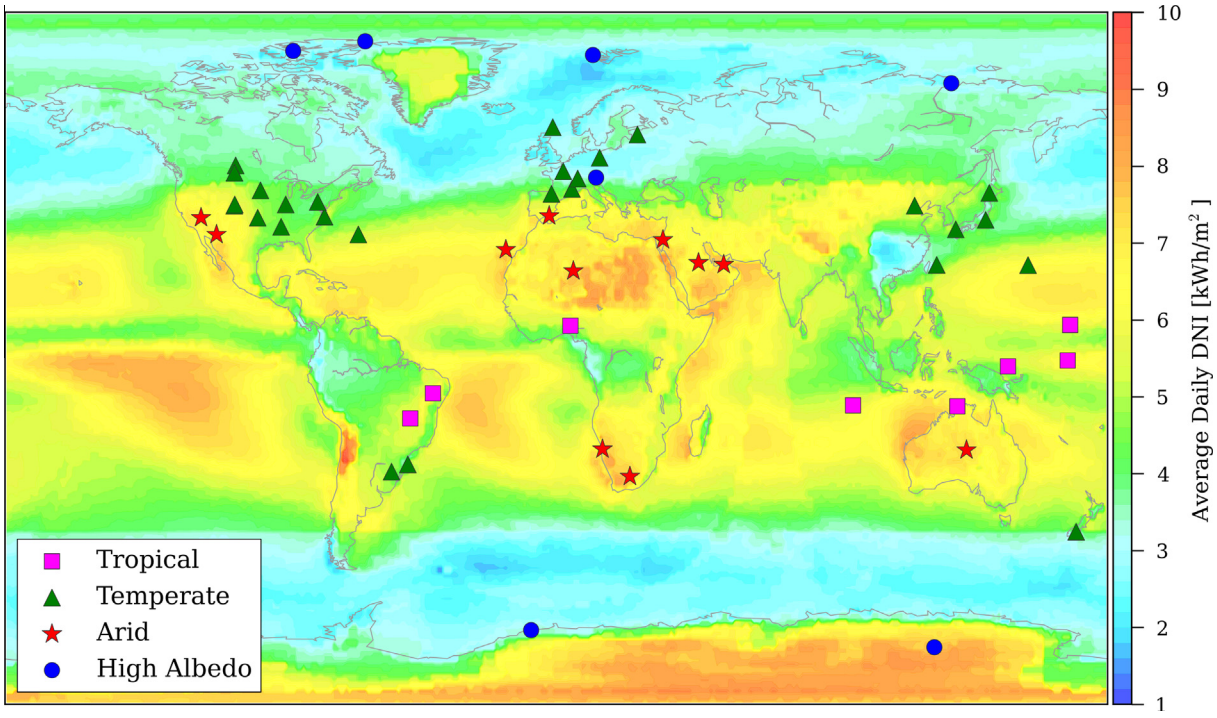

Fig. 2. Geographic distribution of the 54 stations used in this study, superimposed on the mean annual DNI solar resource as derived from the NASA-SSE database (https://eosweb.larc.nasa.gov/sse/), and categorized into 4 climate zones. Red stars: Arid stations; Green triangles: Temperate stations; Purple squares: Tropical stations; Blue circles: High-albedo stations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

way. However, during the main processing and before the calculation of all model results and statistics, both the observed and modeled K values are forced to a maximum value of 1 for consistency. Overall, these tests are inspired from, but in general more stringent than, those recommended for the use of BSRN data (Long and Shi, 2008; Roesch et al., 2011).

3.4. Non-radiometric data

A significant fraction of the models tested here (i.e., those in Table 2) use more inputs than just K_T. A few use Z or air mass, m, which are deterministic, readily available quantities. Some models additionally require temperature, T, and relative humidity, RH. Such variables are not always measured alongside irradiance or at the same frequency, which can cause problems, and ultimately can limit the applicability of the model. In the present case, T and RH are available continuously at only 25 sites, so that REINDL3 and CHENDO3 could not be tested at more than half of the validation sites. At Sede Boker and De Aar, T and RH are only available in 10- and 5-min increments, respectively. At Fukuoka, Ishigakijima, Lindenberg, Minamitorishima, Sapporo and Tateno, they are specified every hour. All of these data series have been interpolated to 1-min steps with a cubic spline. PEREZ1 and PEREZ2 use dew-point temperature, T_{dp}, which is not measured at any of the sites under scrutiny here. Fortunately, T_{dp} can be derived from T and RH with appropriate equations (typically used by meteorological services) for that purpose. These Perez models still work—albeit presumably in less optimum conditions—whenever T_{dp} is not available (Perez et al., 1992). This makes it possible to test them at all sites, including the sites with no T, RH data.

HAY, HOLLANDS2 and SKARTVEIT3 require the surface albedo, ρ, which is rarely measured. This additional variable was introduced in these models to explicitly take backscattering effects between the surface and sky into consideration (see also Section 4.5), particularly over snowy areas. (Backscattering affects both DIF and GHI.) Even if measured locally by, e.g., comparing the reading of a down-facing pyranometer to that of an up-facing pyranometer, this observation would not necessarily represent the far-field albedo required for the models, i.e., an average surface albedo over an area of ≈15 km around the station. A fixed monthly-average value was thus determined for each site, based on a 15-year climatology derived from results of NASA's Modern Era Retrospective-Analysis for Research and Applications (MERRA) reanalysis interpolated to a spatial resolution of 0.5 × 0.5°. Fig. 3 provides a representation of this monthly-mean albedo for all stations. It is obvious that there is much variance compared to the usual assumption of a fixed albedo of 0.2. Temperate sites typically have a slightly lower albedo (0.15–0.18). Coastal and island sites tend to have a much lower albedo (as low as ≈0.07), whereas desert sites may have a constant albedo up to ≈0.37, and cold sites usually experience a high albedo during the snow season—or even all year long

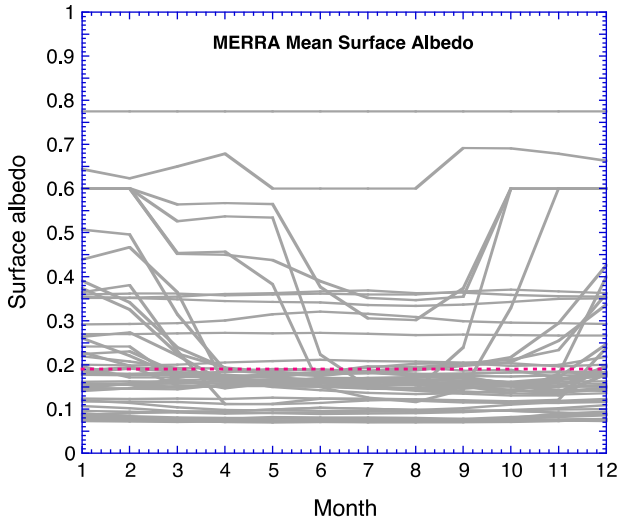

Fig. 3. Estimated mean monthly surface albedo for the 54 test stations. The red dotted line indicates the conventionally assumed constant 0.2 value. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

over high-latitude Arctic and Antarctic regions. The high-elevation Sonnblick station constitutes a special case, since it is located on a mountain top, and is surrounded by glaciers and a lot of snow in winter. Additionally, a low-level cloud deck (of high albedo) frequently forms below the station, particularly during winter. For these reasons, the mean MERRA albedo estimates appeared much too low in winter there, and were significantly increased based on recommendations from the station's principal investigator (Pers. comm. with Dr. Marc Olefs, ZAMG, 2015).

Bugler, Cucumo, Engerer1, Engerer2, Engerer3, Hay, Oumbe, and Perez2 require an evaluation of clear-sky global and/or direct irradiance, which is provided by the Perez–Ineichen model, as suggested and described by Perez et al. (2002). Inputs to this model are provided by the popular monthly-mean, high-resolution Linke turbidity data described by Remund et al. (2003), and available from the SoDa service (http://www.soda-is.com/eng/services/climat_free_eng.php#c5). In the case of Stauter and Turner, the popular HLJ clear-sky model (Gueymard, 2012a; Hottel, 1976; Liu and Jordan, 1960) is rather used, as suggested in the latter's original publication (Stauter and Klein, 1980). The simple Perez–Ineichen and HLJ clear-sky models are well adapted to the extensive calculations involved here, but are known for their limited performance (Gueymard, 2012a, 2014). Therefore, slightly better DNI results could potentially be obtained with more advanced clear-sky models (Vindel et al., 2014).

In addition to the surface albedo, Hay requires the cloud fraction, used to evaluate the sky albedo and the surface-sky backscattering effect. Since no cloud data is available here, the cloud fraction is estimated by reversing the simple Kasten model (Kasten and Czeplak, 1980).

Finally, Maduekwe3 requires the Ångström turbidity coefficient, β. It is estimated here from the Linke turbidity coefficient (Katz et al., 1982).

3.5. Variability indices

Some separation models use information on the temporal variability of GHI through the use of a specific variability index. The significant differences in the way such indices have been defined are worth examining, particularly in the way they can be used here with 1-min data.

The Perez variability index was originally developed using sequences of hourly irradiances immediately before and after the moment being modeled, such as

$$V_1 = 0.5\left[\left|K'_T(t) - K'_T(t+1)\right| + \left|K'_T(t) - K'_T(t-1)\right|\right] \qquad (3)$$

where $K'_T(t)$ is the modified clearness index at time t (Perez et al., 1990a), with special provision in the case one of the $K'_T(t)$ values of the hourly sequence is missing. V_1 is used in both Perez1 and Perez2, with t defined here in 1-min increments.

Skartveit et al. (1998) proposed a more elaborate hourly variability index, V_2, also a function of a special clear-sky index. V_2 is used in Skartveit2 and Skartveit3.

González and Calbó (1999) defined three new indices, V_3–V_5, all based on $K'_T(t)$, like V_1. Their particularity, however, is that they are defined from a sequence of 12 successive 5-min values of $K'_T(t)$ over the preceding hour, with $K'_T(t)$ defined as a different function of K_T and air mass, m, than that of Perez just mentioned. For the present application, the 12 preceding 1-min values of K_T' were used instead.

Tamura et al. (2003) proposed another index, V_6, dubbed "moving function" by its authors, which appears in essence similar to V_5. However, it is calculated from 1-min data over a period of "several minutes". V_6 is also used in Posadillo6. It has been obtained here over a period of 12 min before the time stamp, just like V_3–V_5.

Ridley et al. (2010) introduced a simpler index V_7 for Ridley2, simply defined as the average between $K_T(t-1)$ and $K_T(t+1)$. The previous and next minutes are used here. Later, V_7 was used as input to Boland5, Kuo2, Kuo3, Lauret, and Yao5.

Finally, Engerer (2015) introduced two 1-min indices describing the "instantaneous" departure of the actual sky from ideal clear-sky conditions: $V_8 = K_{Tc} - K_T$ and $V_9 = \mathrm{Max}(0,\ 1 - \mathrm{GHI}_c/\mathrm{GHI})$, where GHI_c is the clear-sky global irradiance calculated with the REST2 model (Gueymard, 2008), and K_{Tc} is the clear-sky equivalent of K_T. For simplification and consistency, GHI_c is rather evaluated here with the Perez–Ineichen clear-sky model.

It should be underlined that the definition of variability indices V_1, V_2 and V_7 requires knowledge of future observations. This is not a problem when dealing with historical time series, like here, but would make them unfit for the operational production of modeled time series, or for nowcasting or forecasting applications, without proper adaptation. Similarly, Boland5, Kuo2, Kuo3, Lauret, Magarreiro, Ridley2 and Yao5 also use the daily-mean K_T as a predictor. This complication is not an issue here

Please cite this article in press as: Gueymard, C.A., Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

(besides the additional computer time), but this requirement would also make them unfit for the aforementioned applications.

3.6. Implementation

The handling of so many datasets and models obviously requires a lot of care to avoid errors, etc., while keeping the process easy to manage and computationally efficient. The flow of calculations consists of two different blocks. In the first block, all data required to run the separation models are exported to a common intermediate format with independent data files for each station and calendar year. The data chosen format is netCDF (http://www.unidata.ucar.edu/software/netcdf/). All the related processing is undertaken with the Python programming language. Each netCDF file contains all valid observational data points (irradiance and ancillary data) at 1-min resolution, as well as all associated deterministic variables, such as extraterrestrial irradiance or solar zenith angle, and quality assurance flags. A fill value is used whenever an observed data value is missing. The second block consists of a single Fortran code file that is called once for each station and calendar year. It ingests the input data and makes additional calculations, such as clear-sky irradiances, daily-average K_T, or variability indices. Then, it successively evaluates the 1-min K and DNI predictions for each model, and finally evaluates all cumulative statistics. Both the model outputs and cumulative statistics are serialized to ASCII files. Simple separation models typically require 3–6 lines of code only. The code for some more complex models (such as those from Perez and Skartveit) was obtained from their authors to avoid misinterpretations.

This extensive calculation process is error-prone. Hence, an important task that requires substantial developing time within the computational framework just described is the implementation of actions directed to assure that errors stay at a minimum. Specifically, both the Python and Fortran source codes embed consistency tests and physical threshold tests for input, intermediate and output data. Whenever a test is not fulfilled a missing value is returned or a fatal error is raised if that is pertinent. After a number of iterations in which initial errors are detected and (subsequently) corrected, all the data can be processed without any issue. A very important design strategy to assure that all the separation models are run from exactly the same set of inputs so that the results are consistent and truly comparable to each other is to implement all the models within the same source file.

Whenever a model yielded suspect results, its code and describing publication were double-checked before their authors were contacted. In most cases, the suspect results could be attributed to typographic errors in numerical coefficients (see Appendix A). Suspect results were also detected with Tamura and Yao5, urging similar requests, but no reply was received from their authors. Finally, an ultimate consistency test consisted in the re-evaluation of the cumulative statistics using Python scripts to double-check those computed with the Fortran program. No differences were found beyond round-off errors.

4. Results and discussion

4.1. Climate clustering

Due to space limitations, it is not possible to report all statistical results for all stations and all models in this report. However, for further reference, these site-specific and model-specific numerical results are provided in the Supplementary Material. Site-specific results in graphical form are also provided in Section 4.3. In any case, such amount of detailed information may be considered of limited value in practice: In most applications, solar analysts would prefer to deal either with only a single "universal" model, or at most with one model per large climate zone. For this reason, four climatic/environmental zones are defined here (see Table 3): (i) Arid (AR, 11 stations); (ii) Temperate (TM, 27 stations); (iii) Tropical (TR, 9 stations); and (iv) High albedo (HA, 7 stations).

A visual inspection of the resulting clustering is possible from Fig. 2. In general, arid sites are synonymous of low cloudiness, high mean annual DNI, which is mapped here based on NASA's Surface meteorology and Solar Energy (SSE, https://eosweb.larc.nasa.gov/sse/) dataset. In contrast, high-latitude sites are impacted by low-sun and generally cloudy conditions, thus have a much lower DNI resource, except over Antarctica. Temperate and tropical sites experience intermediate conditions with respect to DNI.

The clustered results described in the next section clearly indicate that it would not be advisable to merge all these climate zones in an attempt to identify a single "universal" model.

4.2. Statistical indicators

As discussed in Gueymard (2014) and Gueymard and Myers (2008), various statistical indicators can be used to evaluate the performance of DNI predictions. Only a few conventional overall statistics are used here for conciseness: Mean Bias Deviation (MBD), Root Mean Square Deviation (RMSD), and Mean Absolute Deviation (MAD), all expressed in percent of the mean measured DNI in each of the four climate zones. The use of relative statistical results in percent rather than absolute results in W/m^2 follows the recommendations in (Gueymard, 2014). Good models should have low RMSD and MAD values, as well as MBD close to ±0. To stress the fact that the performance of any model is evaluated against observations that are affected by low, but non-negligible, experimental uncertainties, the term "deviation" is used here in lieu of the conventional "error" terminology.

Based on the discussion in the previous section, it is desirable to find which model(s) can have *consistently* high performance over a specific climate zone. A number of criteria have been devised here for this evaluation:

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

(i) The ranges of MBD, RMSD and MAD obtained by aggregating the results from all stations and each model within a given climate zone: Models with shorter ranges indicate better robustness and potential for generalization; in other words, from a practical standpoint, such models offer consistency and repeatability.
(ii) The standard deviations of MBD, RMSD and MAD; again, smaller standard deviations indicate better consistency and repeatability within a given climate zone.
(iii) The frequency of low errors (defined here as |MBD| < 10%, RMSD < 30%, and MAD < 30%) calculated over all individual stations: Models *more frequently* respecting these limits (i.e., at more stations within a given climate zone) are more desirable. The largest possible frequency of occurrence corresponds to the total number of stations in a climate zone. This is an ideal situation, however, so that in general even the best models do not perform well at all stations. This test is considered "not discriminant" if no model performs well enough at more than 75% of the sites in each zone, or, conversely, if more than 14 models (10% of the total population) perform equally well.
(iv) The occurrence of a very low bias (defined here as |MBD| < 5%), when considering the average bias from all stations within a climate zone. The importance of this criterion stems from the fact that any bias in solar resource estimates may have a detrimental effect on the power output simulation, bankability and overall success of large solar energy projects.

Based on these criteria, a single model is recommended for each climate zone. Graphical results are also presented in Sections 4.3 and 4.4.

4.3. Overall statistical results

The overall statistics mentioned in Section 4.2 are compiled in Tables 4–7 for each climate zone separately, considering the arithmetic mean of each statistical indicator for all stations of each zone. These tables also indicate useful information from the additional consistency criteria defined above.

Table 4
Results for the Arid climate zone. N/A: not applicable; StDev: standard deviation.

Criterion	Value	Model
Mean measured DNI (W/m^2)	645.9	N/A
Mean measured GHI (W/m^2)	572.0	N/A
N (million)	4.676	N/A
Number of sites	11	N/A
Separate stations		
Lowest RMSD at a single station (%)	9.4	Perez1
Lowest MAD at a single station (%)	6.0	Udagawa
Smallest Range MBD (%)	12.4	Engerer1
Smallest Range RMSD (%)	7.1	Muneer3
Smallest Range MAD (%)	5.5	Skartveit3
Lowest StDev MBD (%)	3.7	Lee
Lowest StDev RMSD (%)	2.3	Hay
Lowest StDev MAD (%)	1.9	Chandrasekaran
Models reporting |MBD| < 10% most frequently	N/A	Erbs, Gonzalez1, Gonzalez2, Helbig, Kuo3, Oliveira, Perez3, Torres1, Torres2, Torres3, Torres4
Models reporting RMSD < 30% most frequently	N/A	(Not discriminant)
Models reporting MAD < 30% most frequently	N/A	(Not discriminant)
All stations, average		
Low overall bias, |MBD|<5%	N/A	Boland1, Boland4, Boland5, Bourges, DeMiguel, Engerer2, Erbs, Gonzalez1, Gonzalez2, Gonzalez3, Gonzalez6, Hawlader, Hay, Helbig, Hollands1, Hollands2, Ineichen3, Jacovides, Jeter, Kuo2, Kuo3, Lauret, Macagnan, Magarreiro, Oliveira, Orgill, Oumbe, Pagola3, Perez2, Perez3, Reindl1, Ridley1, Ridley2, Skartveit2, Soares, Tapakis1, Torres1, Torres2, Torres3, Torres4, Tsubo3, Tuomiranta2, Zhang
Lowest mean RMSD (%)	15.1	Perez1
Lowest mean MAD (%)	10.0	Perez1
Benchmark MBD (%)	4.5	Liu
Benchmark RMSD (%)	18.4	Liu
Benchmark MAD (%)	13.0	Liu
Recommended model's MBD (%)	−3.2	Engerer2
Recommended model's RMSD (%)	16.0	Engerer2
Recommended model's MAD (%)	11.8	Engerer2

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

Table 5
Results for the temperate zone. N/A: not applicable; StDev: standard deviation.

Criterion	Value	Model
Mean measured DNI (W/m^2)	371.8	N/A
Mean measured GHI (W/m^2)	391.1	N/A
N (million)	13.579	N/A
Number of sites	27	N/A
Separate stations		
Lowest RMSD at a single station (%)	17.7	Perez2
Lowest MAD at a single station (%)	11.2	Perez1
Smallest Range MBD (%)	23.5	Engerer2
Smallest Range RMSD (%)	47.3	Yao2
Smallest Range MAD (%)	16.2	Yao2
Lowest StDev MBD (%)	5.6	Engerer2
Lowest StDev RMSD (%)	9.6	Clarke
Lowest StDev MAD (%)	4.0	Yao2
Models reporting |MBD| < 10% most frequently	N/A	Engerer2, Suehrke
Models reporting RMSD < 30% most frequently	N/A	Engerer2, Reindl2, Skartveit1, Suehrke
Models reporting MAD < 30% most frequently	N/A	(Not discriminant)
All stations, average		
Low overall bias, |MBD| < 5%	N/A	Bakhsh, Chikh1, Chikh2, Chikh3, CIBSE, DeJong, Engerer2, Gonzalez4, Gonzalez5, Gonzalez7, Gonzalez8, Lam2, Mondol1, Muneer1, Muneer2, Muneer3, PerezBurgos, Rerhrhaye, Skartveit1, Suehrke, Tapakis2, Tuomiranta1, Tuomiranta2, Tuomiranta3
Lowest mean RMSD (%)	28.2	Engerer2
Lowest mean MAD (%)	16.8	Engerer2
Benchmark MBD (%)	21.3	Liu
Benchmark RMSD (%)	40.1	Liu
Benchmark MAD (%)	27.0	Liu
Recommended model's MBD (%)	2.1	Engerer2
Recommended model's RMSD (%)	28.2	Engerer2
Recommended model's MAD (%)	16.8	Engerer2

At arid sites (Table 4), the mean measured DNI is large (as could be expected) and significantly larger than the mean GHI. Based on 4.7 million data points collected at 11 stations, the results indicate that Perez1 provides both the lowest absolute mean RMSD (15.1%) and the lowest absolute mean MAD (10.0%). However, it cannot be considered consistent over the arid zone for two reasons: (i) Both its range and standard deviation (SD) of MBD, RMSD and MAD statistics over the individual sites are significantly larger than those of other models; and (ii) Its bias is not always low. Compared to the benchmark (Liu, also referred to as LJ60), the gain of accuracy obtained by the recommended model for this climate zone (Engerer2) appears modest. The station-specific bias of eleven different models is within ±10% at all stations, while 43 models have a mean bias within ±5% over the entire dataset for this climate zone. Conversely, no model achieves the goal of maintaining both an RMSD and an MAD lower than 30% at all 11 sites simultaneously. These results indicate that absolute accuracy at a single site and site-to-site consistency may be two conflicting model features. This finding requires further scrutiny.

The temperate zone (Table 5) has the largest population (27 stations) and the largest number of data points (13.6 millions) of all climate zones considered here. It is also representative of the radiative conditions for which most of the models of the literature have been developed. It can thus be expected that many models perform well under such conditions. Curiously, however, conflicting results between accuracy and consistency, similar to those just discussed, are readily apparent. Whereas Perez1 and Perez2 provide the lowest absolute mean MAD and RMSD, respectively, these two models lack consistency. Engerer2, and to a lower extent, Suehrke, obtain the highest scores in consistency: Their bias is below 10% at 25 of the 27 sites and their RMSD is lower than 30% at 17 sites. Substantially systematic larger deviations at Lerwick make this station appear as an outlier. Its high latitude and cloudy climate makes its DNI resource lower than most other sites (Table 3). Occasional snow may also result in a relatively higher diffuse fraction than usual. MAD is less than 30% at the 26 other sites for no less than 50 models. The latter test cannot therefore be considered as discriminant in this case. Compared to Liu's benchmark results, the gain of accuracy of the recommended model (Engerer2) appears more substantial here than for the arid zone, particularly with regard to MBD.

Results from the tropical zone (Table 6) are based on 9 stations and 4.5 million data points, and do share some features with those of the temperate zone just discussed. For instance, Perez2 obtains the lowest mean MAD and RMSD, and Engerer2 appears relatively consistent (albeit less conclusively than before). Many other models obtain good scores in either accuracy or consistency, but only a

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

Table 6
Results for the tropical zone. N/A: not applicable; StDev: standard deviation.

Criterion	Value	Model
Mean measured DNI (W/m^2)	421.7	N/A
Mean measured GHI (W/m^2)	502.0	N/A
N (million)	4.483	N/A
Number of sites	9	N/A
Separate stations		
Lowest RMSD at a single station (%)	19.6	Perez2
Lowest MAD at a single station (%)	13.0	Perez2
Smallest Range MBD (%)	10.5	Engerer2
Smallest Range RMSD (%)	68.9	Engerer2
Smallest Range MAD (%)	26.9	Perez2
Lowest StDev MBD (%)	3.4	Engerer1
Lowest StDev RMSD (%)	21.7	Kuo1
Lowest StDev MAD (%)	8.1	Perez2
Models reporting |MBD| < 10% most frequently	N/A	Boland5, Engerer2, Ridley1, Skartveit1
Models reporting RMSD < 30% most frequently	N/A	Boland5, Engerer2, Gonzalez2, Kuo3, Perez1, Perez2, Perez3, Reindl2, Skartveit1
Models reporting MAD < 30% most frequently	N/A	(Not discriminant)
All stations, average		
Low overall bias, |MBD| < 5%	N/A	Bakhsh, Boland3, Boland5, Bourges, Chadrasekaran, Chikh1, Chikh2, Chikh3, CIBSE, DeJong, Engerer2, Gonzalez5, Gonzalez7, Gonzalez8, Helbig, Ineichen1, Ineichen2, Lauret, Magarreiro, Mondol1, Muneer1, Muneer3, Perez2, Perez3, Reindl2, Ridley1, Skartveit1, Skartveit2, Skartveit3, Stauter, Turner
Lowest mean RMSD (%)	32.0	Engerer2
Lowest mean MAD (%)	19.2	Perez2
Benchmark MBD (%)	25.9	Liu
Benchmark RMSD (%)	46.4	Liu
Benchmark MAD (%)	32.2	Liu
Recommended model's MBD (%)	−1.3	Engerer2
Recommended model's RMSD (%)	32.0	Engerer2
Recommended model's MAD (%)	19.5	Engerer2

few of them are found consistent here. These differences between the two climate zones may be caused in part by a too low number of stations to reliably represent the climate characteristics of the tropical zone. Four models (Boland5, Engerer2, Ridley1, and Skartveit1) have a bias below 10% at all 9 sites, but none has either RMSD < 30% or MAD < 30% at all sites. As with the temperate zone, Engerer2 is the recommended model, which brings substantial improvement over the benchmark (Liu), particularly with regard to MBD.

The high-albedo zone (Table 7) counts the least number of sites (7) and data points (3 million). It is also the most internally diverse in terms of latitude, elevation and seasonal variability in albedo (some sites having a high albedo in winter but a much lower albedo in summer). In any case, this special zone is of practical importance because some large solar systems are being developed in mountain areas, where snowfall may occur more or less frequently. Furthermore, only 3 models (Hay, Hollands2, and Skartveit3) have been specifically designed to accommodate variable albedo as an input. Of these 3 models, only Hollands2 performs relatively consistently at all sites. Still, it has important issues that would need to be addressed for 7systematic use with 1-min data (see Section 4.5). The lowest mean RMSD and MAD are obtained by Engerer2. The consistent results of Chendo1, Maduekwe1 and Yao3 are also noticeable, but those of Yao2 are even better, which is the reason why it is selected as the recommended model for that zone. Chendo1 and Maduekwe1 have a low bias (<5%) at 5 of the 7 sites, while Engerer2 has a low MAD (<30%) at 5 sites too. In contrast, no model obtains a low RMSD (<30%) at more than one site. Compared to Liu's benchmark results, the gain in accuracy of the recommended model for this zone (Yao2) appears substantial, particularly with regard to MBD. Moreover, a closer look at individual results (see Supplementary Material) shows that most models that perform well in the other three climate zones do not have such success here.

A simplified graphical representation of all results appears in Figs. 4–15. The best performer in each of the 6 model categories (for 1–6 predictors, as described in Tables 1 and 2) is indicated by a color dot. For each station and each of the 3 statistical indicators (MBD, RMSD and MAD), the acronym of the best overall performer is also indicated. Note the significant scatter in the results from all individual sites. A close examination of Figs. 4–15 clearly indicates that two stations appear like outliers within their own climate zone: Lerwick for the Temperate zone, and Ilorin for the Tropical zone. For the former case, potential reasons were proposed above. The case of Ilorin can be

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

Table 7
Results for the high-albedo zone. N/A: not applicable; StDev: standard deviation.

Criterion	Value	Model
Mean measured DNI (W/m^2)	343.6	N/A
Mean measured GHI (W/m^2)	279.3	N/A
N (million)	2.992	N/A
Number of sites	7	N/A
Separate stations		
Lowest RMSD at a single station (%)	17.7	Engerer1
Lowest MAD at a single station (%)	10.9	Watanabe
Smallest Range MBD (%)	41.9	Hollands2
Smallest Range RMSD (%)	63.5	Yao3
Smallest Range MAD (%)	36.7	Hollands2
Lowest StDev MBD (%)	15.4	Hollands2
Lowest StDev RMSD (%)	20.4	Yao3
Lowest StDev MAD (%)	13.1	Hollands2
Models reporting |MBD| < 10% most frequently	N/A	Chendo1, Maduekwe1
Models reporting RMSD < 30% most frequently	N/A	(Not discriminant)
Models reporting MAD < 30% most frequently	N/A	Engerer2
All stations, average		
Low overall bias, |MBD| < 5%	N/A	Chendo2, DeJong, Engerer1, Maduekwe2, Maduekwe3, Yao1, Yao2, Yao3, Yao4
Lowest mean RMSD (%)	53.2	Engerer2
Lowest mean MAD (%)	32.1	Engerer2
Benchmark MBD (%)	49.4	Liu
Benchmark RMSD (%)	75.0	Liu
Benchmark MAD (%)	53.2	Liu
Recommended model's MBD (%)	1.4	Yao2
Recommended model's RMSD (%)	56.4	Yao2
Recommended model's MAD (%)	40.3	Yao2

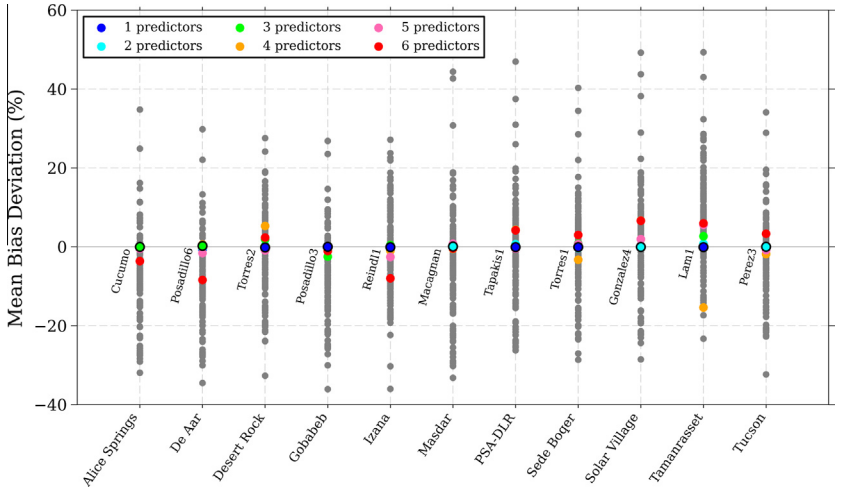

Fig. 4. MBD results for each station of the arid climate zone. For all models with the same number of predictors, the MBD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

explained by the much lower DNI resource than elsewhere (Table 3), which tends to drive relative errors up. Somewhat larger experimental errors are also likely there because the station reportedly suffered difficulties with maintenance due to insufficient funding (WCRP, 2001).

At arid sites, the typically low nebulosity makes results more dependent on the local aerosol climate. This may explain why both RMSD and MAD are larger at sites frequently impacted by high-turbidity conditions (such as Tamanrasset), when the actual diffuse fraction is higher than normal due to aerosol scattering, than at cleaner sites (such as Alice Springs, Gobabeb, or Tucson), as noted previously (Gueymard and Ruiz-Arias, 2014).

A systematically large overestimation of DNI is notable at most high-albedo sites. This can be explained by

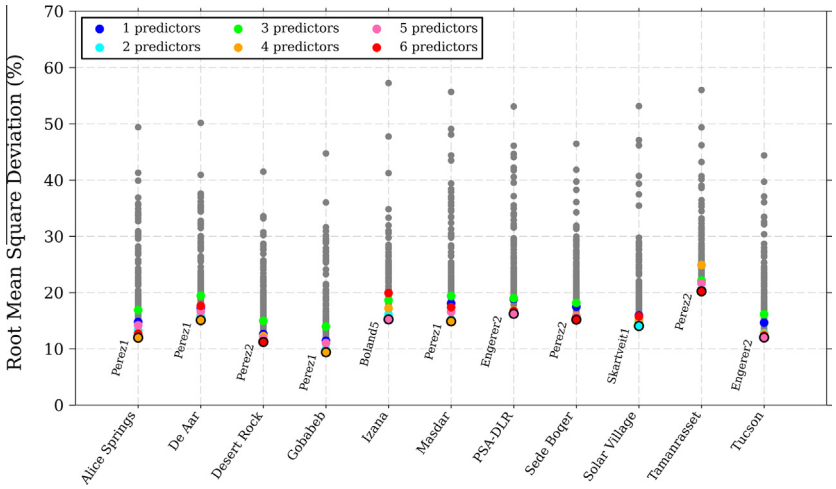

Fig. 5. RMSD results for each station of the arid climate zone. For all models with the same number of predictors, the RMSD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

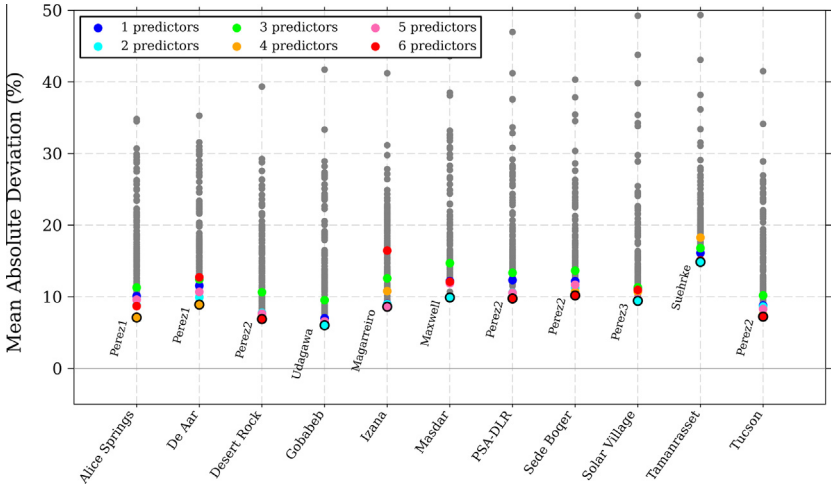

Fig. 6. MAD results for each station of the arid climate zone. For all models with the same number of predictors, the MAD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the enhanced diffuse backscattering induced by a highly reflective surface, particularly under cloudy or overcast skies, which is ignored or incorrectly taken into account by nearly all separation models. An exception to this trend occurs at Concordia, which is surprising since its albedo is the highest and most constant of all stations. It is likely that the high latitude and, most importantly, a low-cloudiness regime (explaining the high mean DNI there) successfully attenuate the albedo effects there.

Another unexpected finding is that single-predictor models are frequently those with the best MBD (close to 0). Conversely, the best RMSD and MAD results are obtained with models having 2 or more predictors, with only a few exceptions.

4.4. Taylor diagrams

Alternatively to the previous analysis, Taylor diagrams (Taylor, 2001) provide a convenient way to visualize and compare the performance of different models (Gueymard, 2014). Fig. 16 shows one such diagram for each of the four climate zones considered here. The 140 models are [1]color coded to denote their number of predictors. A perfect model would be located along the circle of unit value of standardized deviation, indicating it has the same standard deviation as the observed dataset, and over the thick dot on

[1] For interpretation of color in Fig. 16, the reader is referred to the web version of this article.

Please cite this article in press as: Gueymard, C.A., Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

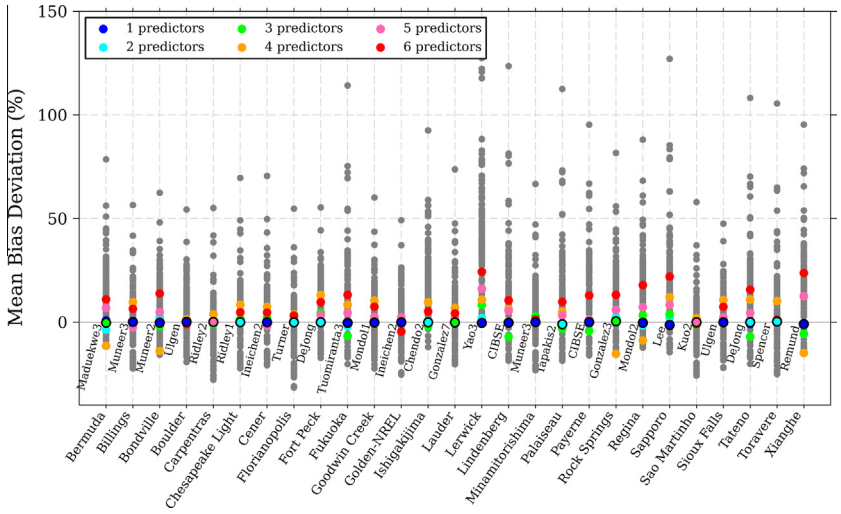

Fig. 7. MBD results for each station of the temperate climate zone. For all models with the same number of predictors, the MBD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

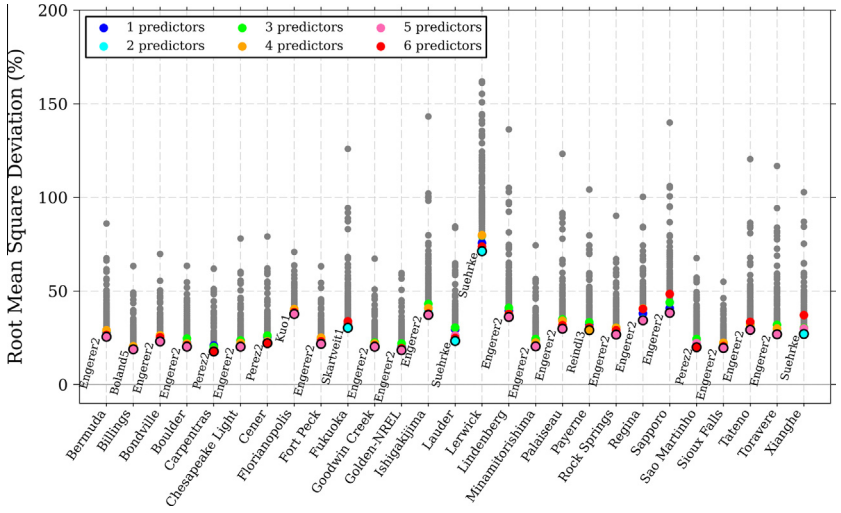

Fig. 8. RMSD results for each station of the temperate climate zone. For all models with the same number of predictors, the RMSD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the X-axis, marking a perfect correlation coefficient of 1.0. The modeled-measured correlation with the best models for all zones is ≈0.95, except the high-albedo zone where it only reaches ≈0.90. Most models with 5 or 6 predictors are closer to the ideal location on the diagrams than models with less predictors. This suggests that adding predictors tends to improve model precision and repeatability. There are exceptions to this observation, however, which are possibly caused by overfitting when adding many predictors multilinearly and without physical basis. Moreover, the distance between models having a different number of predictors is not large, which may explain why Tables 4–7 indicate that simpler models sometimes perform as well or better than more elaborate ones.

4.5. Impact of surface albedo and cloud enhancement

With conventional hourly irradiance data, it is rare to observe K_T values above ≈0.75 at any low-elevation site. As Fig. 1 demonstrates, things are different when using 1-min data. K_T can reach much higher, over-unity values because of intense transient effects usually referred to as "cloud enhancement", "cloud lensing" or "overirradiance". Recent investigations under different types of climate (Almeida et al., 2014; Piacentini et al., 2011; Piedehierro et al., 2014; Tapakis and Charalambides, 2014; Yordanov et al., 2015) generally describe these events as transient situations of high sun surrounded by bright clouds. Detection of such events is important in the context

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

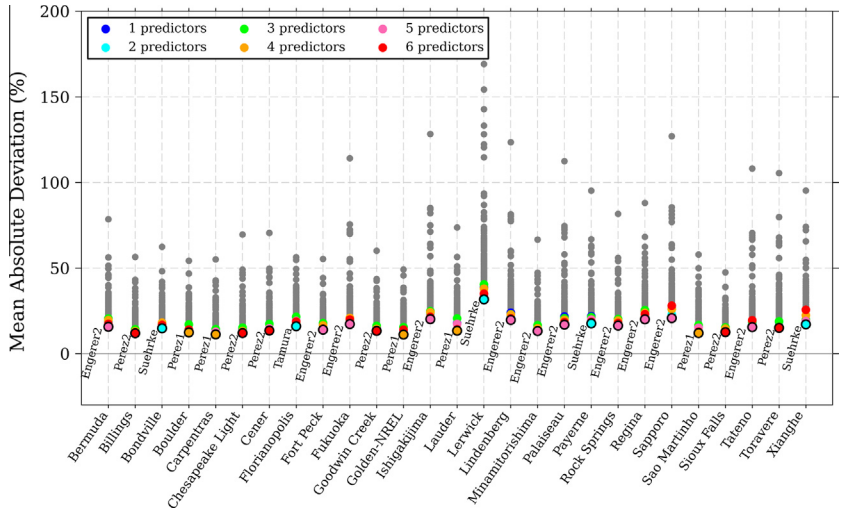

Fig. 9. MAD results for each station of the temperate climate zone. For all models with the same number of predictors, the MAD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

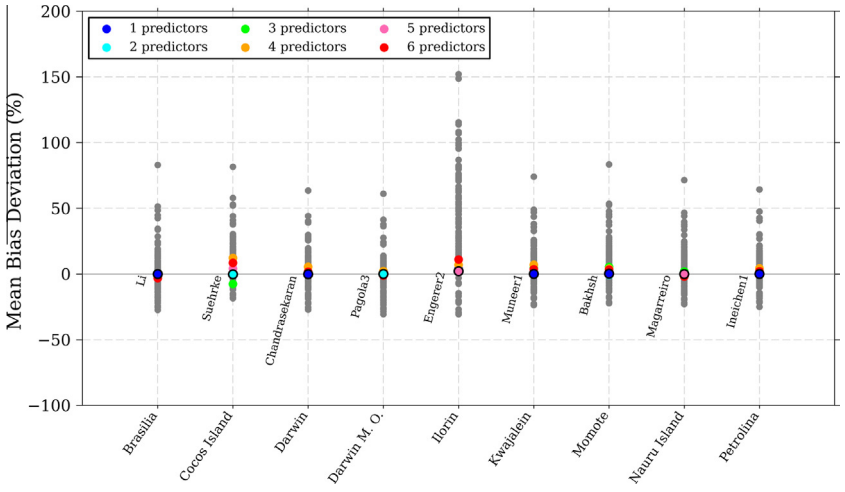

Fig. 10. MBD results for each station of the tropical climate zone. For all models with the same number of predictors, the MBD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of PV applications in particular to correctly size inverters (Luoma et al., 2012) or to characterize ramping effects and their impact on electricity distribution and grid stability (Lave et al., 2015a).

At the 54 sites under scrutiny here, the absolute maximum value of 1-min K_T, noted K_{TM}, is typically between 1.0 and ≈1.5 over each station's entire record, which means that GHI can be occasionally up to 50% larger than its extraterrestrial counterpart—and still be a valid observation. Fig. 17 indicates that K_{TM} tends to be lower at tropical and arid sites, and higher at temperate and high-albedo sites. Some occurrences of extremely high K_{TM} are observed at Lerwick ($K_{TM} = 1.86$) and Sonnblick ($K_{TM} = 2.50$). These findings were unexpected because situations of high sun and bright clouds of high vertical extension are more typical of tropical sites than of higher-latitude sites. Closer examination reveals that, at Sonnblick for instance, a large fraction of the $K_T > 1.5$ events actually occur under *low-sun* conditions. This seems to be caused by a strong increase in DIF due to intense backscattering between highly reflective ground and bright clouds. Some of these cases could also be the result of transient experimental errors, despite the various levels of data quality control involved here. Even though a better understanding of the possible causes behind all high-K_T situations is desirable, it is beyond the scope of this contribution. The existence of low-sun high-K_T events was unexpected, based on the studies previously

Please cite this article in press as: Gueymard, C.A., Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

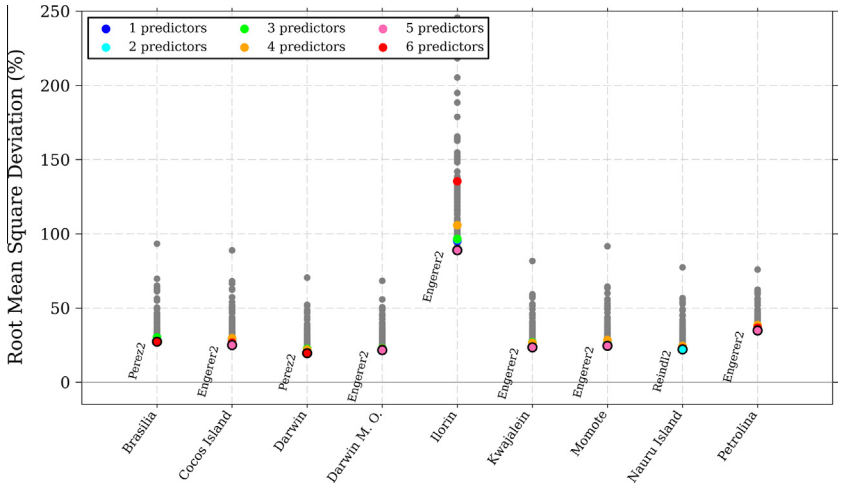

Fig. 11. RMSD results for each station of the tropical climate zone. For all models with the same number of predictors, the RMSD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

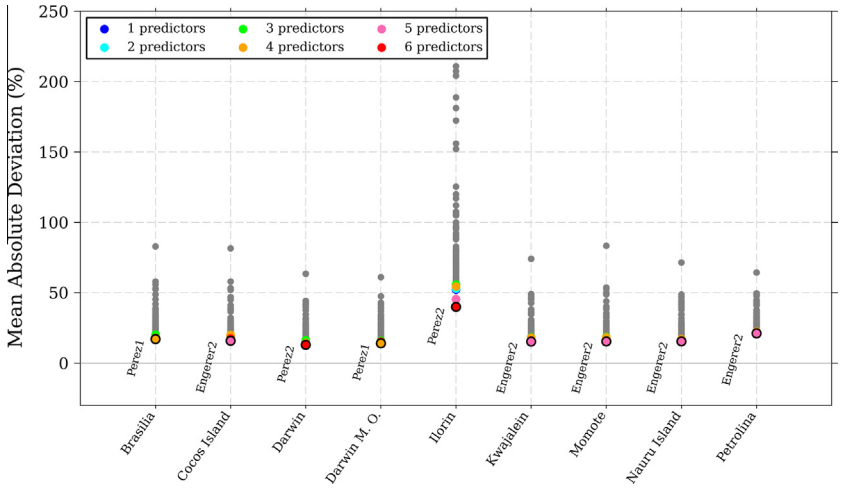

Fig. 12. MAD results for each station of the tropical climate zone. For all models with the same number of predictors, the MAD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mentioned. Their focus on "high-GHI" rather than "high-K_T" events might explain this situation. In any case, the impact of such circumstances on the diffuse fraction, and ultimately DNI, is worth evaluating here.

Fig. 18 provides a comparison of the $K = f(K_T)$ relationship at two sites distant by ≈460 km and at roughly the same latitude: Payerne, Switzerland (low elevation with rare snow accumulation) and Sonnblick, Austria (high elevation and high albedo all winter). Fig. 18a displays 4 regions of interest in the K–K_T plot, corresponding to clear-sky, partly cloudy-sky, overcast-sky, and cloud-enhancement situations. The popular ERBS function is also plotted, showing that it follows the general features of the observations, except under cloud-enhancement conditions. In comparison, Fig. 18b shows more scatter and a significant horizontal spread of the experimental points. Compared to Fig. 18a, two new zones, corresponding to partly-cloudy and overcast conditions over high-albedo ground, are added. Interestingly, there are now many points characterized by unit values of *both* K and K_T—a situation that apparently never occurs under low-albedo conditions, as suggested by Fig. 18a. The figure also displays the HOLLANDS2 functions for both a snow-free ground ($\rho = 0.2$) and a brighter ground ($\rho = 0.6$). The latter curve better represents the high-albedo diffuse fraction's spatial pattern than the former, but still predicts a too low diffuse fraction under overcast and cloud-enhanced conditions. Furthermore, for any value of ρ, the HOLLANDS2 function diverges for K_T beyond ≈0.84 and then returns unphysical values. It would thus require

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

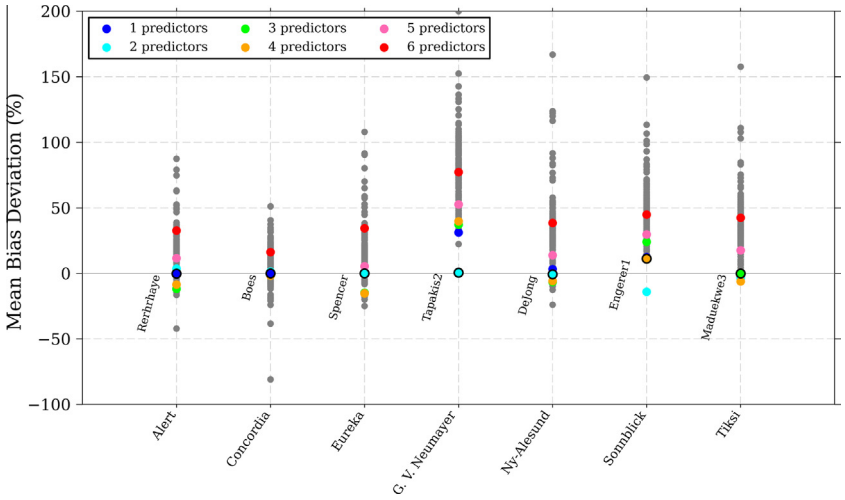

Fig. 13. MBD results for each station of the high-albedo zone. For all models with the same number of predictors, the MBD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

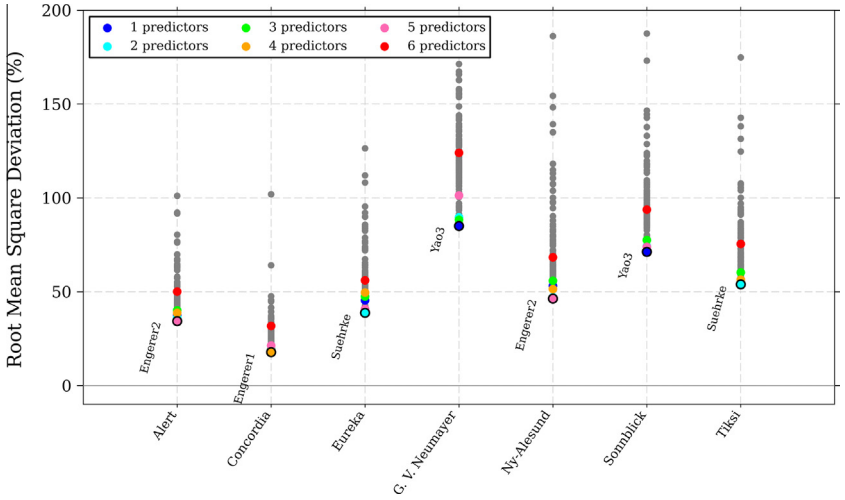

Fig. 14. RMSD results for each station of the high-albedo zone. For all models with the same number of predictors, the RMSD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

important corrections to be used safely with all possible 1-min GHI data. These model issues would not have been guessed by just looking at the bulk statistics of Table 7, for instance. This also confirms previous findings (Gueymard and Ruiz-Arias, 2014) that not all hourly separation models can be used with 1-min data without appropriate modification.

A close inspection of the models' performance for different K_T regimes, number of predictors and climate zone is shown in Fig. 19. To speed up the calculation of these curves, the year with the largest number of observations at each of six stations randomly chosen from each climate zone have been selected. Therefore, a similar number of years (6 stations × 1 year/station) and a similar number of observations (6 years × 0.2 million/year) are used to make the plots for each climate zone. The average MBD value (normalized to the mean DNI value of the 6-year dataset in each climate zone) of all the models with the same number of predictors is stratified into 50 equal-width bins of K_T from 0 to 1.2. The individual results for ENGERER2 and for the combined predictions of the models with the lowest MBD for each station (referred to as "Lowest MBD" in Fig. 19) are also shown for reference. The shaded area goes from the 15th percentile to the 85th percentile of all MBD values from the 140 individual models. Thus, this area does not include the 15% of the models with the lowest or highest MBD values. For each climate zone, the frequency distribution of K_T is also shown to highlight the importance of the bias in any particular K_T interval relatively to the whole zonal dataset.

Please cite this article in press as: Gueymard, C.A., Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

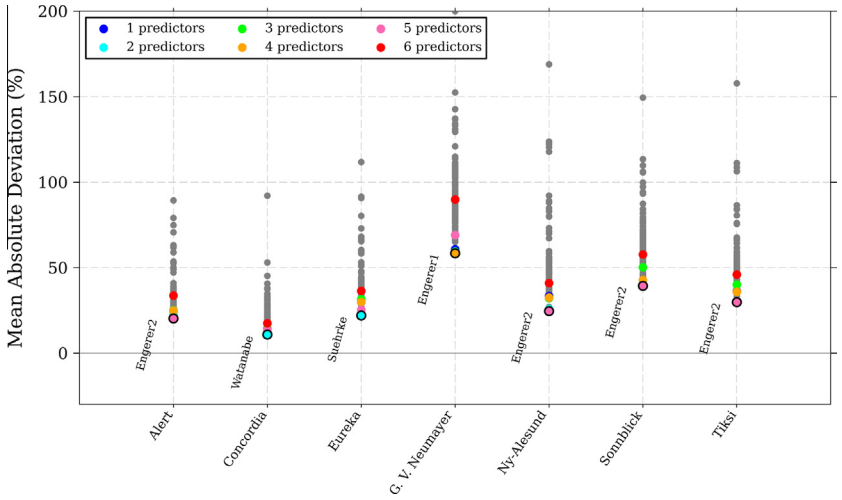

Fig. 15. MAD results for each station of the high-albedo zone. For all models with the same number of predictors, the MAD value of the best performing model at each station is indicated with a colored marker that maps the number of predictors. The best performing model is named and indicated by a marker surrounded with a thick black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

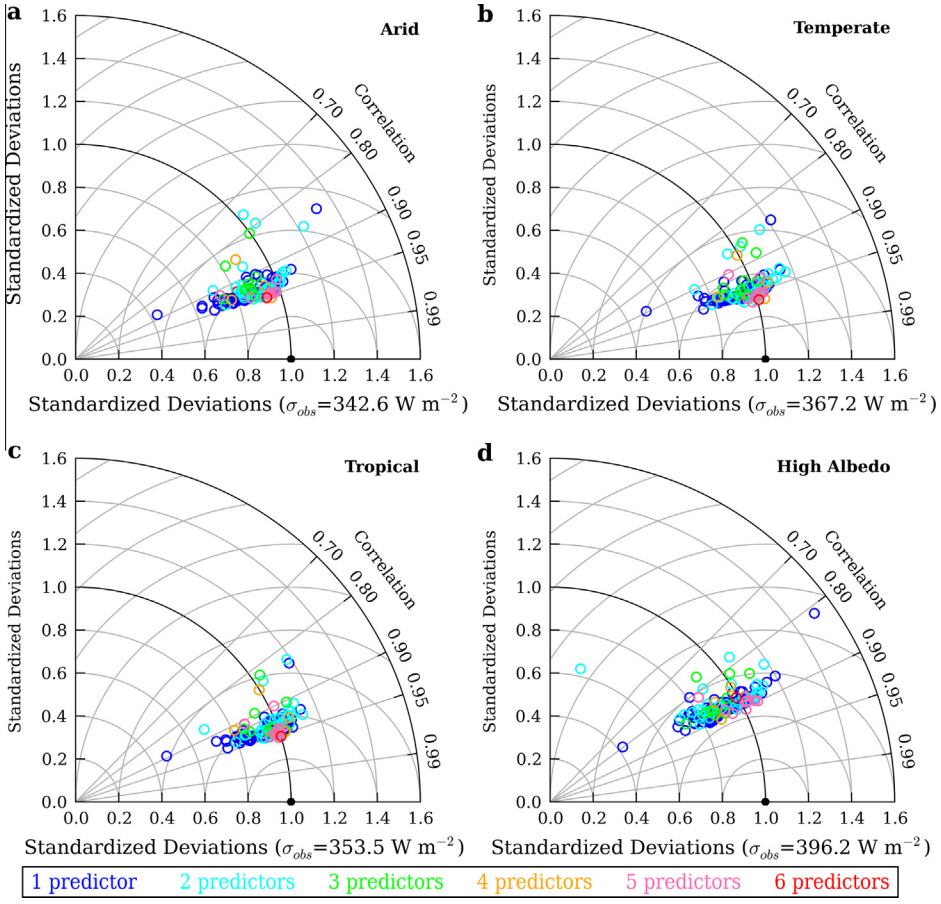

Fig. 16. Taylor diagrams for all models and four climate zones. Note axes show standardized deviations, i.e., the standard deviation of the models is normalized by the standard deviation of the observational dataset.

Within each zone, what is referred to as "Lowest MBD" can be considered the best estimate overall, in terms of bias. Even though the combination of a site-by-site selection of models might outperform the results of a single model, it is stressed that the issue of how to "smartly" select and combine these models remains unknown and is

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

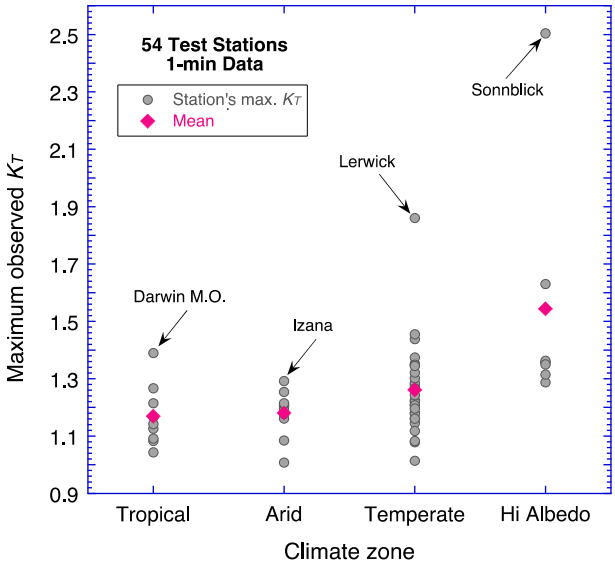

Fig. 17. Maximum observed value of K_T, K_{TM}, at all sites.

beyond the scope of this study. The majority of models overestimate during cloud-enhancement events regardless of climate zone. PEREZ2 and ENGERER2 are notable exceptions to this rule: They tend to underestimate, albeit to a lesser extent. Although the biases in this K_T region (beyond ≈0.8) are quite high, their significance is low, except perhaps for high-albedo stations, since cloud-enhancement events are relatively rare overall. In general terms, using models with more predictors does not guarantee lower MBDs under high-K_T situations. More generally, the latter observation can be extended to all K_T regimes.

The modeling of K as a function of, primarily, K_T and some ancillary variables seems to be a good approach for clear skies, at least in terms of MBD. This is evidenced by the low MBD values that are found when K_T is in the 0.7–0.8 range under arid, temperate or tropical climates. The MBD values and differences between models increase when K_T decreases (i.e., as cloudy situations come into play). As could be expected, this is more obvious at the temperate and tropical stations than at arid sites. The overall situation at the high-albedo sites is singular, as was already comparatively introduced in Fig. 18. In particular, the prevailing high albedo of both the surface and clouds produces a higher DIF than what the vast majority of models predict, thus inducing the large DNI overestimations shown in Fig. 19d for intermediate K_T regimes ($K_T \approx 0.4$–0.6).

Since temperate climates are prevalent worldwide, the Boulder station is used here as an example of the typical conditions that can be expected in this climate zone. It is found that a significant fraction (13.7%) of the data record's K_T is larger than 0.8, a value assumed here to be the lower limit for 1-min cloud-enhanced situations at that site. There is still 0.3% of the data points whose GHI is larger than its extraterrestrial counterpart ($1.0 < K_T \leqslant 1.438$). For these overunity points, Fig. 20 compares the response of three models (ENGERER2, ERBS, and PEREZ2) against the observed DNI. Whereas ERBS strongly overestimates DNI (MBD = 37.2%), both PEREZ2 and, to a lesser extent, ENGERER2 significantly underestimate (MBD = −25.4% and −16.9%, respectively). The marked inability of ERBS to make physically sound predictions under high-K_T situations has already been noted by these authors (Gueymard and Ruiz-Arias, 2014). As an explanation, they pointed to the fixed low value of the diffuse fraction ($K = 0.165$) imposed by the model when $K_T > 0.80$. The discrepancy that results in such cases is obvious when considering plots such as Figs. 1 or 18a: The model then predicts a too low K, and thus a much too high DNI. ENGERER2 performs better than ERBS or PEREZ2 in this case, consistently with the fact that it is a truly minutely model, and that it explicitly includes a correction for cloud-enhanced situations. The underprediction of PEREZ2 results from its tendency to generate very low, or even zero, DNI when its measured value is actually in the range 500–1000 W/m^2. Compared

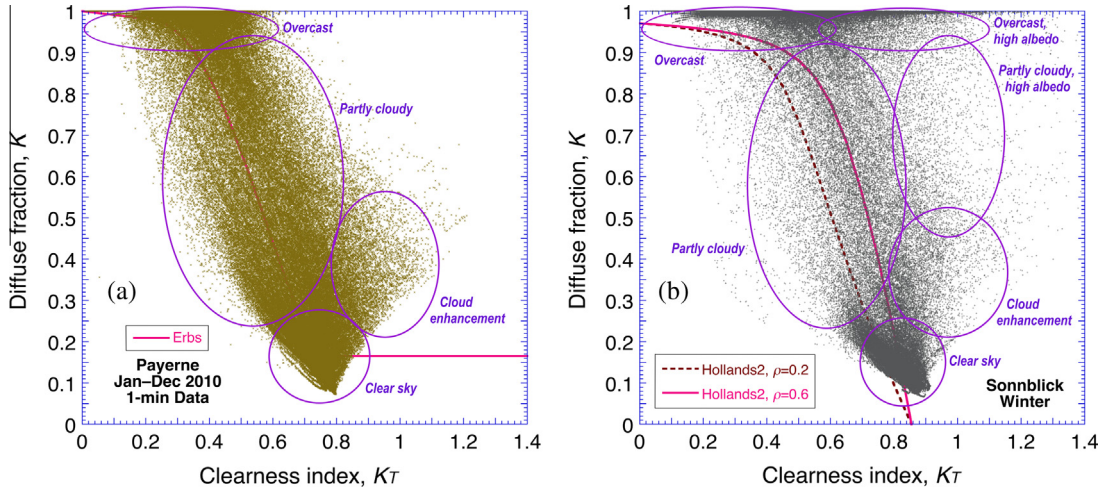

Fig. 18. Observed K–K_T relationship at (a) Payerne during one complete year and (b) Sonnblick during winter, showing various zones corresponding to distinct atmospheric conditions. The ERBS function and the HOLLANDS2 functions for two surface albedo values are also shown.

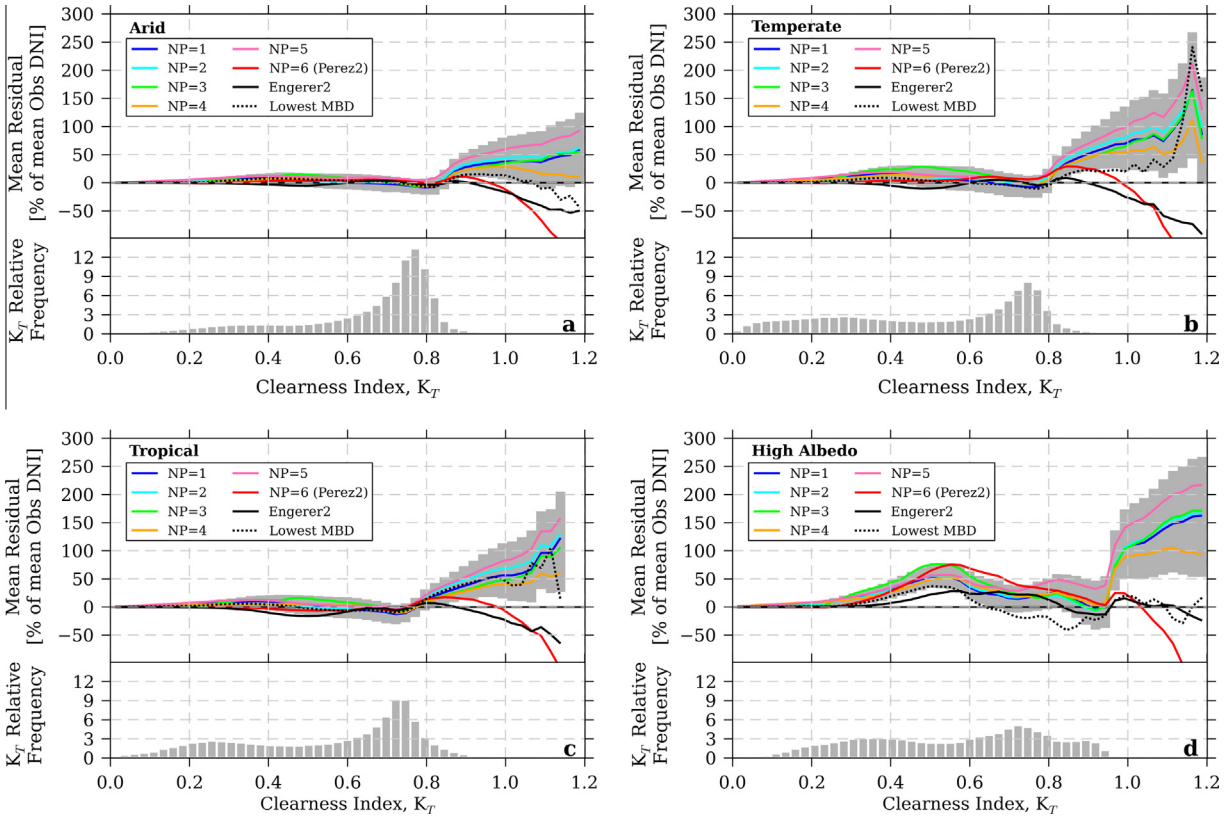

Fig. 19. Average MBD value (normalized to the mean DNI value for each climate zone) of all the models with the same number of predictors stratified into 50 equal-width bins of K_T from 0 to 1.2. The individual results for ENGERER2 and for the combined predictions of the models with the lowest MBD for each station (referred to as "Lowest MBD") are also shown. The shaded area goes from the 15th percentile to the 85th percentile of all MBD values from the 140 individual models. For each climate zone, the frequency distribution of K_T is also shown.

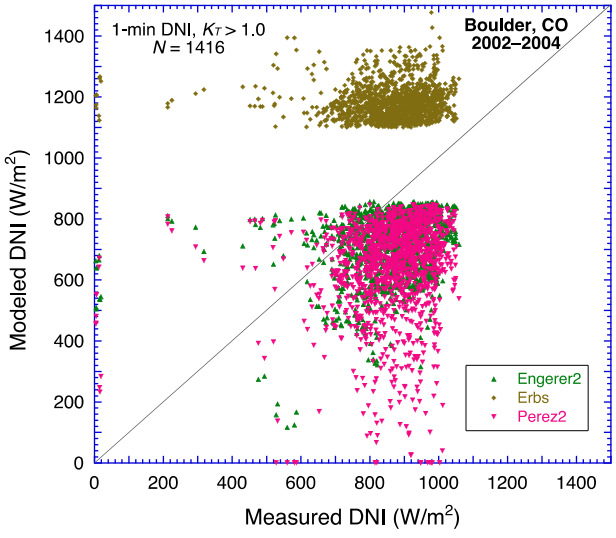

Fig. 20. Predicted vs. measured DNI at Boulder with three models for all specific overunity situations when $K_T > 1.0$.

to ERBS, which unphysically predicts DNI values up to 1681 W/m^2 here, JANJAI and PEREZBURGOS overpredict even more, sometimes reaching values 6 times larger (i.e., greater than 10 kW/m^2). Conversely, 7 models (CIBSE, CUCUMO, MUNEER2, PAGOLA, REINDL2, SKARTVEIT1 and SUEHRKE) behave remarkably well in this special case—at that site at least—by maintaining an MBD within 5% of the mean measured DNI (862.4 W/m^2) and a maximum DNI below 1160 W/m^2 (i.e., 10% more than the measured maximum when $K_T > 1$). The results shown in Fig. 20 for ERBS and PEREZ2 generalize previous findings (Gueymard and Ruiz-Arias, 2014), which were limited to arid sites.

4.6. Recommended models

An examination of Tables 4–7 shows that some models are more frequently cited for their good results than others. Table 8 lists all 9 models that are cited at least 4 times in Tables 4–7, and thus are candidates for providing accurate and/or consistent results over different climate zones. Among these, the dominating model (ENGERER2) appears 17 times, in comparison with 9 times for PEREZ2. Other models cited 4 or 5 times are (in alphabetical order) BOLAND5, ENGERER1, HOLLANDS2, PEREZ1, PEREZ3, SKARTVEIT1, and YAO2. The very presence of ENGERER1 in this list is surprising because it is a *clear-sky* separation model. This explains why it is highly negatively biased in the Arid, Temperate and Tropical climate zones, where it cannot be recommended at all. However, it has the lowest

Table 8
Statistical results for the best models, averaged over each climate zone. The mean value and its standard deviation are given in percent of the mean observed DNI for MBD, RMSD and MAD.

Model	Citations	Predictors	Arid			Temperate			Tropical			High albedo		
			MBD	RMSD	MAD	MBD	RMSD	MAD	MBD	RMSD	MAD	MBD	RMSD	MAD
Engerer2	17	5	−3.2 ± 4.8	16.0 ± 2.8	11.8 ± 2.3	2.1 ± 5.6	28.2 ± 10.9	16.8 ± 4.5	−1.3 ± 3.5	32.0 ± 21.8	19.5 ± 10.0	18.6 ± 17.8	53.2 ± 26.7	32.1 ± 17.9
Perez2	9	6	0.4 ± 5.2	15.3 ± 3.5	10.6 ± 3.3	8.8 ± 7.7	29.7 ± 12.0	17.6 ± 5.5	2.6 ± 4.7	37.8 ± 36.9	19.2 ± 8.1	40.9 ± 18.6	71.4 ± 30.4	46.0 ± 22.9
Hollands2	5	2	−2.0 ± 9.5	21.7 ± 4.1	15.4 ± 4.1	11.3 ± 9.3	39.0 ± 16.7	23.4 ± 7.6	9.4 ± 7.1	42.5 ± 24.3	26.8 ± 13.0	19.3 ± 15.4	62.0 ± 24.4	39.3 ± 13.1
Perez1	5	4	3.5 ± 5.1	15.1 ± 4.0	10.0 ± 3.4	12.7 ± 7.6	31.5 ± 13.7	18.5 ± 6.7	5.7 ± 3.6	36.2 ± 26.6	21.7 ± 12.5	32.7 ± 26.2	66.1 ± 31.2	42.7 ± 22.4
Skartveit1	5	2	−5.7 ± 4.5	17.3 ± 3.0	13.9 ± 2.4	3.8 ± 7.7	29.3 ± 11.2	18.9 ± 4.8	−0.1 ± 4.9	35.3 ± 25.2	23.2 ± 12.4	17.6 ± 24.1	60.8 ± 27.4	43.6 ± 21.6
Boland5	4	5	0.1 ± 5.8	16.5 ± 3.3	11.6 ± 2.8	7.8 ± 7.9	30.9 ± 12.6	19.7 ± 6.3	1.2 ± 4.2	35.4 ± 25.1	23.6 ± 14.1	32.8 ± 22.1	60.9 ± 28.8	39.7 ± 19.9
Engerer1	4	4	−22.8 ± 3.9	30.2 ± 3.9	26.2 ± 3.2	−16.6 ± 7.5	40.3 ± 10.7	28.2 ± 4.3	−23.2 ± 3.4	51.8 ± 29.0	34.6 ± 9.9	1.9 ± 18.5	53.4 ± 21.7	35.5 ± 13.6
Perez3	4	2	−0.4 ± 4.7	15.5 ± 3.5	11.1 ± 3.1	8.9 ± 7.7	30.3 ± 12.8	18.1 ± 5.7	3.8 ± 4.2	36.5 ± 26.8	23.0 ± 13.2	24.8 ± 25.4	63.0 ± 29.3	42.8 ± 23.1
Yao2	4	1	−23.7 ± 4.7	31.0 ± 4.4	27.3 ± 3.5	−16.7 ± 6.7	40.3 ± 9.9	28.8 ± 4.0	−19.8 ± 3.6	45.6 ± 23.5	31.5 ± 9.3	1.4 ± 19.0	56.4 ± 22.0	40.3 ± 15.5

bias in the high-albedo climate zone, and also has a low standard deviation everywhere—hence high site-to-site consistency. Hollands2 obtains reasonable scores everywhere, but should not be used whenever $K_T > 0.8$ for reasons detailed in the previous section. Hence, it cannot be recommended for the prediction of 1-min DNI in general. Yao2 appears in this list in part because of its relatively good performance in the high-albedo zone. However, it does not perform as well in other zones, where it is negatively biased. Perez1 and Perez2 often have low RMSD and MAD, but are also frequently affected by significant bias, as was noted previously (Gueymard, 2010; Gueymard and Ruiz-Arias, 2014), and by inaccuracies under cloud-enhanced situations (Fig. 20). This may in turn impact the absolute accuracy of plane-of-array irradiance calculations aimed at tilted solar collectors, for instance (Lave et al., 2015b). In contrast, Engerer2's bias is generally lower and more consistent. Its success may be due in part to its derivation considering 1-min data, rather than hourly data as with nearly all other existing models, and its generally better response to cloud-enhanced situations. It must be noted that, unexpectedly, popular and frequently cited models, such as DeMiguel, Erbs, Maxwell, Orgill, Reindl2, or Skarveit2, do not appear in this shortlist.

It is also remarkable that 8 of the 9 best models listed in Table 8 use at least one predictor in addition to K_T. (The exception is Yao2, which performs well only at high-albedo stations.) However, relying on a large number of predictors is not in itself a guarantee of success, since most models using 3 or more predictors are actually absent from this list. A possible explanation for this counter-intuitive finding, which corroborates that in Gueymard and Ruiz-Arias (2014), is that additional predictors are usually added multi-linearly, without any physical justification or basis, which likely results in over-fitting and a decline of the model's generalization skill. Thus, such models may work locally for the sites originally used for their empirical derivation, but not necessarily elsewhere.

Even though Engerer2, and to a lesser extent Perez2, can be considered the best models overall based on the compilation in Table 8, they do not perform best at all sites, and most particularly at high-albedo sites. In that sense, no model is actually found to have true "universal" validity, or to offer both accuracy and consistency even over a single climate zone. A better way to derive future separation models, using less empiricism and more physics, could be a way to solve this critical "universality" issue. Until better models are proposed, it is suggested that analysts select their model of choice based on a number of application-specific criteria and on a climate-by-climate basis. For instance, for solar applications located at close proximity to one of the present test sites, the accuracy criterion relative to either that specific site (Supplementary Material) or its climate zone (Tables 4–7) could be used as a selection tool. In contrast, for solar applications pertaining to multiple sites in, e.g., arid zones, one of the most

consistent models of Table 4 should be selected. Finally, for any application outside of snowy areas, ENGERER2 is currently the best candidate to become the "quasi-universal" 1-min model.

5. Conclusion

A comprehensive statistical analysis and validation of 140 separation models of the literature is proposed here. Their performance assessment is obtained by comparing their DNI predictions to high-quality 1-min measured data obtained at 54 research-class stations in four different climate zones of the world, resulting in a remarkably high number of validation data points (≈25 million). The combination of this large number of models and test stations represents an unprecedented level of effort.

The accuracy of the predicted DNI is found highly dependent on the specific separation model used, local specificities (such as surface albedo or atmospheric turbidity), as well as the number of predictors each model uses. Some of those separation models that include a variability index, and some leverage provided by clear-sky irradiance estimates, tend to be more accurate than those that do not. This is not a general conclusion, however, because some of these detailed models have more bias than simpler models and may even generate spurious predictions.

The impact of cloud enhancement—a transient atmospheric condition that can typically increase the 1-min clearness index (K_T) to over-unity values—appears very significant here. Such cases are characterized by infrequent but large K_T values (up to 2.5) and result in highly overestimated predictions of DNI, particularly over the temperate and high-albedo climate zones. This indicates that most separation models that were developed using hourly radiation data are not designed to operate correctly under high-K_T conditions. This finding may have implications for, e.g., the proper design or operation of solar PV systems wherever cloud enhancement situations are frequent.

The impact of high surface albedo on the accuracy of the predicted DNI has also received much attention. Seven stations, where snowy conditions prevail during all or most of the year, provide data for this original study. Curiously, the very few (only three) albedo-savvy models that have been found in the literature do not perform satisfactorily or better than simpler models under such conditions. Actually, no existing model performs consistently well under such conditions. Overall, the largest DNI overestimations (in relative terms) are found at high-albedo sites for K_T values larger than ≈0.95. This means that both high-albedo and cloud enhancement situations need to be better taken into consideration in separation models, so that more accurate DNI predictions can be obtained under any possible condition.

In addition to the usual summary statistics that describe model performance, new criteria are proposed to evaluate the robustness of each model within a given climate zone. Based on these criteria, it is found that two models stand out over the arid, temperate and tropical climate zones: ENGERER2 and PEREZ2. These two models share two important features: (i) They include a variability predictor; and (iii) They leverage clear-sky irradiance estimates. The reason why ENGERER2 performs consistently better than PEREZ2 or other models is most likely because it was actually derived from 1-min data (compared to hourly data for PEREZ2 or most other models tested here). Based on the ensemble of statistical results obtained here, it is concluded that ENGERER2 has the best generalization skill, and can thus be considered a "quasi-universal" 1-min separation model, wherever and whenever low-albedo conditions prevail.

Compared to the very first separation model that was introduced some 55 years ago, most more recent models do not generally offer much improved accuracy, except maybe locally for the area for which they were designed. Considering that this type of radiation model is ubiquitous to produce solar resource data used in essentially all solar applications, and that the current trend in solar engineering is to use irradiance data with short time steps, the present study should bring a new perspective on which area of research is still in need of improvement. It is thus hoped that the present findings will help stimulate the development of more advanced separation models.

Acknowledgments

The authors express their gratitude to the dedicated personnel who maintain the 54 radiometric stations considered here, particularly those from the Baseline Solar Radiation Network, whose high-quality measurements are central to this study. The PSA-DLR and Masdar data were kindly provided by Stefan Wilbert and Peter Armstrong, respectively, whose collaboration is highly appreciated.

Appendix A. Corrections to existing models

Details are provided below on the corrections made to some models, as obtained from their authors through personal communication.

- BOLAND5

  The original publication (Boland et al., 2013) had some errors and omissions. Eq. (4) of the paper needs to be corrected as:

$$I_{DN} = 0.02628/[0.006 + 4.374\ \exp(-7.75K_T - 1.185K_{Td} - 1.05\psi - 0.004\mathrm{AST} + 0.003\alpha)]$$

  where K_{Td} is the mean daily value of K_T, and all other variables are as defined in the original publication. The implied unit of α is degree, and that of I_{DN} is MJ/m^2, which thus needs to be multiplied by 1000/3.6 to obtain W/m^2 as used here.

- CHENDO3
  Coefficient −0.633 in Eq. (3c) of Chendo and Maduekwe (1994) should read +0.633.

- POSADILLO models

  In Posadillo and Lopez Luque (2009), coefficient 1.77 of Eq. (7) should read 1.17, coefficient 1.4139 of Eq. (9) should read −1.4139, and coefficient 5.839 of Eq. (11) should read –5.839.

- RUIZARIAS2

  In Ruiz-Arias et al. (2010), there is a typo in Eq. (34) that defines Model G2, but coefficients in the paper's Table 4 are correct.

- TAPAKIS models

  Tables 5 and 7 of Tapakis et al. (2015) contain errors. In Table 5, columns p1 to p6 should read in reverse order (p6 to p1). In Table 7, columns p5, p6 and p7 should read p6, p7 and p5, respectively.

- TORRES1

  Coefficient a8 in Eq. (1) and Table 1 of Model 1 in Torres et al. (2010) should read 0.1923.

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.solener.2015.10.010. Consistent with the units used in the text, all irradiances are in W/m^2 and all statistical results are in percent.

References

Al-Riahi, M., Al-Hamdani, N., Tahir, K., 1992. An empirical method for estimation of hourly diffuse fraction of global radiation. Renew. Energy 2, 451–456.

Almeida, M.P., Zilles, R., Lorenzo, E., 2014. Extreme overirradiance events in Sao Paulo, Brazil. Sol. Energy 110, 168–173.

Bakhsh, H., Srinivasan, R., Bahel, V., 1985. Correlation between hourly diffuse and global radiation for Dhahran, Saudi Arabia. Sol. Wind Technol. 2, 59–61.

Battles, F.J., Riubio, M.A., Tovar-Pescador, J., Olmo, F.J., Alados-Arboledas, L., 2000. Empirical modeling of hourly direct irradiance by means of hourly global irradiance. Energy 25, 675–688.

Bertrand, C., Vanderveken, G., Journée, M., 2015. Evaluation of decomposition models of various complexity to estimate the direct solar irradiance over Belgium. Renew. Energy 74, 618–626.

Blanco-Muriel, M., Alarcon-Padilla, D.C., Lopez-Moratalla, T., Lara-Coira, M., 2001. Computing the solar vector. Sol. Energy 70, 431–441.

Boes, E.C., 1975. Estimating the direct component of solar radiation. Rep. SAND75-0565, Sandia Lab., Albuquerque, NM.

Boland, J., Huang, J., Ridley, B., 2013. Decomposing global solar radiation into its direct and diffuse components. Renew. Sustain. Energy Rev. 28, 749–756.

Boland, J., Ridley, B., 2008. Models of diffuse solar fraction. In: Badescu, V. (Ed.), Modeling Solar Radiation at the Earth Surface. Springer, pp. 193–219.

Boland, J., Ridley, B., Brown, B., 2008. Models of diffuse solar radiation. Renew. Energy 33, 575–584.

Boland, J., Scott, L., Luther, M., 2001. Modelling the diffuse fraction of global solar radiation on a horizontal surface. Environmetrics 12, 103–116.

Bourges, B., 1992. Climatic data handbook for Europe. Kluwer, Dordrecht.

Bruno, R., 1978. A correction procedure for separating direct and diffuse insolation on a horizontal surface. Sol. Energy 20, 97–100.

Bugler, J.W., 1977. The determination of hourly insolation on an inclined plane using a diffuse irradiance model based on hourly measured global horizontal insolation. Sol. Energy 19, 477–491.

Cebecauer, T., Suri, M., Gueymard, C.A., 2011. Uncertainty sources in satellite-derived direct normal irradiance: how can prediction accuracy be improved globally? In: Proc. SolarPACES Conf., Granada, Spain.

Chandrasekaran, J., Kumar, S., 1994. Hourly diffuse fraction correlation at a tropical location. Sol. Energy 53, 505–510.

Chendo, M.A.C., Maduekwe, A.A.L., 1994. Hourly global and diffuse radiation of Lagos, Nigeria—correlation with some atmospheric parameters. Sol. Energy 52, 247–251.

Chikh, M., Mahrane, A., Haddadi, M., 2012. Modeling the diffuse part of the global solar radiation in Algeria. Energy Proc. 18, 1068–1075.

CIBSE, 2002. Weather, Solar and Illuminance Data. The Chartered Institution of Building Services Engineers, London, UK.

Clarke, P., Munawwar, S., Davidson, A., Muneer, T., Kubie, J., 2007. An investigation of possible improvements in accuracy of regressions between diffuse and global solar irradiation. Build. Serv. Eng. Res. Technol. 28, 189–197.

Cronin, A.D., Richardson, W., Brooks, A.E., DellaGiustina, D.N., 2013. Comparing ramp rates from large and small PV systems, and selection of batteries for ramp rate control. In: Proc. IEEE Photovoltaic Specialists Conference, Tampa, FL.

Cucumo, M., De Rosa, A., Ferraro, V., Kaliakatsos, D., Marinelli, V., 2006. Experimental validation of correlations for the estimation of beam and diffuse components of hourly radiation. In: Proc. 61st ATI National Congress Solar Heating and Cooling, Associazione Termotecnica Italiana, Perugia, Italy.

De Jong, J.B.R.M., 1980. Een karakterisering van de zonnestraling in Nederland. Doctor-aalverslag Vakgroep Fysiche Aspecten van de Gebouwde Omgeving afd. Bouwkunde en Vakgroep Warmte-en Stromingstechnieken afd. Werktuigbouwkunde. Technische Hogeschool, Eindhoven, Netherlands.

De Miguel, A., Bilbao, J., Aguiar, R., Kambezidis, H.D., Negro, E., 2001. Diffuse solar irradiation model evaluation in the north Mediterranean belt area. Sol. Energy 70, 143–153.

Dervishi, S., Mahdavi, A., 2012. Computing diffuse fraction of global horizontal solar radiation: a model comparison. Sol. Energy 86, 1796–1802.

Elminir, H.K., Azzam, Y.A., Younes, F.I., 2007. Prediction of hourly and daily diffuse fraction using neural network, as compared to linear regression models. Energy 32, 1513–1523.

Engerer, N.A., 2015. Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia. Sol. Energy 116, 215–237.

Erbs, D., Klein, S.A., Duffie, J.A., 1982. Estimation of the diffuse radiation fraction for hourly, daily and monthly average global radiation. Sol. Energy 28, 293–302.

Erusiafe, N.E., Chendo, M.A.C., 2014. Estimating diffuse solar radiation from global solar radiation. In: Proc. EuroSun 2014 Conf., ISES, Aix-les-Bains, France.

Furlan, C., Oliveira, A.P., 2008. Hourly Diffuse Solar Radiation in the Presence of Clouds and Other Environmental Parameters: The City of São Paulo. Univ. of Padua, Italy.

González, J.A., Calbó, J., 1999. Influence of the global radiation variability on the hourly diffuse fraction correlations. Sol. Energy 65, 119–131.

Gueymard, C.A., 2008. REST2: high performance solar radiation model for cloudless-sky irradiance, illuminance and photosynthetically active radiation—validation with a benchmark dataset. Sol. Energy 82, 272–285.

Gueymard, C.A., 2009. Direct and indirect uncertainties in the prediction of tilted irradiance for solar engineering applications. Sol. Energy 83, 432–444.

Gueymard, C.A., 2010. Progress in direct irradiance modeling and validation. In: Proc. Solar 2010 Conf., American Solar Energy Soc., Phoenix, AZ.

Gueymard, C.A., 2012a. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: improved validation methodology and detailed performance analysis of 18 broadband radiative models. Sol. Energy 86, 2145–2169.

Gueymard, C.A., 2012b. An introduction to solar radiation. In: Meyers, R.A. (Ed.), Encyclopedia of Sustainability Science and Technology. Springer, pp. 608–633.

Gueymard, C.A., 2014. A review of validation methodologies and statistical performance indicators for modeled solar radiation data: towards a better bankability of solar projects. Renew. Sustain. Energy Rev. 39, 1024–1034.

Gueymard, C.A., Myers, D.R., 2008. Validation and ranking methodologies for solar radiation models. In: Badescu, V. (Ed.), Modeling Solar Radiation at the Earth Surface. Springer.

Gueymard, C.A., Myers, D.R., 2009. Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling. Sol. Energy 83, 171–185.

Gueymard, C.A., Renné, D., Vignola, F., Goswami, Y., 2009. Editorial: journal's performance and publication criteria. Sol. Energy 83, 1.

Gueymard, C.A., Ruiz-Arias, J.A., 2014. Performance of separation models to predict direct irradiance at high frequency: validation over arid areas. In: Proc. EuroSun 2014, ISES, Aix-les-Bains, France.

Gueymard, C.A., Ruiz-Arias, J.A., 2015. Validation of direct normal irradiance predictions under arid conditions: a review of radiative models and their turbidity-dependent performance. Renew. Sustain. Energy Rev. 45, 379–396.

Hawlader, M.N.A., 1984. Diffuse, global and extraterrestrial solar radiation for Singapore. Int. J. Ambient Energy 5, 31–38.

Hay, J.E., 1976. A revised method for determining the direct and diffuse components of the total shortwave radiation. Atmosphere 14, 278–287.

Hay, J.E., Davies, J.A., 1980. Calculation of the solar radiation incident on an inclined surface. In: Hay, J.E., Won, T.K. (Eds.), First Canadian Solar Radiation Data Workshop, Canadian Atmospheric Environment Service, Toronto, Canada.

Helbig, N., Löwe, H., Mayer, B., Lehning, M., 2010. Explicit validation of a surface shortwave radiation balance model over snow-covered complex terrain. J. Geophys. Res. 115D. http://dx.doi.org/10.1029/2010JD013970.

Heller, A., Dahm, J., 1999. The Marstal central solar heating plant: design and evaluation. Proc. ISES Solar World Congress. Elsevier, Jerusalem, Israel.

Hijazin, M.I., 1998. The diffuse fraction of hourly solar radiation for Amman, Jordan. Renew. Energy 13, 249–253.

Hirsch, T., Schenk, H., Schmidt, N., Meyer, R., 2010. Dynamics of oil-based parabolic trough plants—impact of transient behaviour on energy yields. In: Proc. SolarPACES Conf., Perpignan, France.

Hollands, K.G.T., 1985. A derivation of the diffuse fraction's dependence on the clearness index. Sol. Energy 35, 131–136.

Hollands, K.G.T., Crha, S.J., 1987. An improved model for diffuse radiation: correction for atmospheric back-scattering. Sol. Energy 38, 233–236.

Hottel, H.C., 1976. A simple model for estimating the transmittance of direct solar radiation through clear atmospheres. Sol. Energy 18, 129–134.

Ineichen, P., 2008. Comparison and validation of three global-to-beam irradiance models against ground measurements. Sol. Energy 82, 501–512.

Ineichen, P., Guisan, O., Razafindraibe, A., 1984. Indice de clarté, Mesures d'ensoleillement à Genève, vol. 9. Available from <http://www.cuepe.ch/html/biblio/pdf/ineichen1984-indicedeclarte(9).pdf>, University of Geneva, Switzerland.

Jacovides, C.P., Boland, J., Asimakopoulos, D.N., Kaltsounides, N.A., 2010. Comparing diffuse radiation models with one predictor for partitioning incident PAR radiation into its diffuse component in the eastern Mediterranean basin. Renew. Energy 35, 1820–1827.

Jacovides, C.P., Tymvios, F.S., Assimakopulos, V.D., Kaltsounides, N.A., 2006. Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation. Renew. Energy 31, 2492–2504.

Janjai, S., Phaprom, P., Wattan, R., Masiri, I., 2010. Statistical models for estimating hourly diffuse solar radiation in different regions of Thailand. In: Proc. PEA-AIT International Conference on Energy and Sustainable Development: Issues and Strategies (ESD 2010), Asian Institute of Technology, Chiang Mai, Thailand.

Jeter, S.M., Balaras, C.A., 1986. A regression model for the beam transmittance of the atmosphere based on data for Shenandoah, Georgia, USA. Sol. Energy 37, 7–14.

Karatasou, S., Santamouris, M., Geros, V., 2003. Analysis of experimental data on diffuse solar radiation in Athens, Greece, for building applications. Int. J. Sustain. Energy 23, 1–11.

Kasten, F., Czeplak, G., 1980. Solar and terrestrial radiation dependent on the amount and type of cloud. Sol. Energy 24, 177–189.

Kasten, F., Duffie, J.A., 1993. Editorial. Sol. Energy 50, 383.

Katz, M., Baille, A., Mermier, M., 1982. Atmospheric turbidity in a semi-rural site—I. Evaluation and comparison of different atmospheric turbidity coefficients. Sol. Energy 28, 323–327.

Klein, S.A., Cooper, P.I., Freeman, T.L., Beekman, D.M., Beckman, W.A., Duffie, J.A., 1975. A method of simulation of solar processes and its application. Sol. Energy 17, 29–37.

Kuo, C.-W., Chang, W.-C., Chang, K.-C., 2014. Modeling the hourly solar diffuse fraction in Taiwan. Renew. Energy 66, 56–61.

Lam, J.C., Li, D.H.W., 1996. Correlation between global solar radiation and its direct and diffuse components. Build. Environ. 31, 527–535.

Lauret, P., Boland, J., Ridley, B., 2013. Bayesian statistical analysis applied to solar radiation modelling. Renew. Energy 49, 124–127.

Lave, M., Reno, M.J., Broderick, R.J., 2015a. Characterizing local high-frequency solar variability and its impact to distribution studies. Sol. Energy 118, 327–337.

Lave, M., Hayes, W., Pohl, A., Hansen, C.W., 2015b. Evaluation of global horizontal irradiance to plane of array irradiance models at locations across the United States. IEEE J. Photovolt. 5, 597–606.

Lee, K., Yoo, H., Levermore, G.J., 2013. Quality control and estimation hourly solar irradiation on inclined surfaces in South Korea. Renew. Energy 57, 190–199.

Li, D.H.W., Lam, J.C., 2001. Analysis of solar heat gain factors using sky clearness index and energy implications. Energy Convers. Manag. 42, 555–571.

Liu, B.Y.H., Jordan, R.C., 1960. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Sol. Energy 4, 1–19.

Long, C.N., Shi, Y., 2008. An automated quality assessment and control algorithm for surface radiation measurements. Open Atmos. Sci. J. 2, 23–37.

Lopez, G., Rubio, M.A., Battles, F.J., 2000. Estimation of hourly direct normal from measured global solar irradiance in Spain. Renew. Energy 21, 175–186.

Louche, A., Notton, G., Poggi, P., Simonnot, G., 1991. Correlations for direct normal and global horizontal irradiation on a French Mediterranean site. Sol. Energy 46, 261–266.

Luoma, J., Kleissl, J., Murray, K., 2012. Optimal inverter sizing considering cloud enhancement. Sol. Energy 86, 421–429.

Macagnan, M., Lorenzo, E., Jimenez, C., 1994. Solar radiation in Madrid. Int. J. Sol. Energy 16, 1–14.

Maduekwe, A.A.L., Chendo, M.A.C., 1997. Atmospheric turbidity and the diffuse irradiance in Lagos, Nigeria. Sol. Energy 61, 241–249.

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

Magarreiro, C., Brito, M.C., Soares, P.M.M., 2014. Assessment of diffuse radiation models for cloudy atmospheric conditions in the Azores region. Sol. Energy 108, 538–547.

Maxwell, E.L., 1987. A quasi-physical model for converting hourly global horizontal to direct normal insolation. Tech. Rep. SERI/TR-215-3087, National Renewable Energy Lab., Golden, CO. <http://rredc.nrel.gov/solar/pubs/PDFs/TR-215-3087.pdf>.

Michalsky, J., Dutton, E., Nelson, D., Wendell, J., Wilcox, S., Andreas, A., Gotseff, P., Myers, D., Reda, I., Stoffel, T., Behrens, K., Carlund, T., Finsterle, W., Halliwell, D., 2011. An extensive comparison of commercial pyrheliometers under a wide range of routine observing conditions. J. Atmos. Ocean Technol. 28, 752–766.

Michalsky, J.J., 1988. The astronomical almanac's algorithm for approximate solar position (1950–2050). Sol. Energy 40, 227–235, Errata: Sol. Energy, 241, 113 (1988).

Michalsky, J.J., Gueymard, C.A., Kiedron, P., McArthur, L.J.B., Philipona, R., Stoffel, T., 2007. A proposed working standard for the measurement of diffuse horizontal shortwave irradiance. J. Geophys. Res. 112D. http://dx.doi.org/10.1029/2007JD008651.

Mondol, J.D., Yohanis, Y.G., Norton, B., 2008. Solar radiation modelling for the simulation of photovoltaic systems. Renew. Energy 33, 1109–1120.

Mondol, J.D., Yohanis, Y.G., Smyth, M., Norton, B., 2005. Long-term validated simulation of a building integrated photovoltaic system. Sol. Energy 78, 163–176.

Moreno, S., Silva, M., Fernandez-Peruchena, C.M., Pagola, I., 2009. Comparison of methodologies to estimate direct normal irradiation from daily values of global horizontal irradiation. In: Proc. SolarPACES Conf., Berlin, Germany.

Muneer, T., Hawas, M.M., Sahili, K., 1984. Correlation between hourly diffuse and global radiation for New Delhi. Energy Convers. Manag. 24, 265–267.

Muneer, T., Kambezidis, H., Tregenza, P., 1997. Solar Radiation and Daylight Models for the Energy Efficient Design of Buildings. Architectural Press.

Muneer, T., Saluja, G.S., 1986. Correlation between hourly diffuse and global solar irradiation for the UK. Build. Serv. Eng. Res. Technol. 7, 37–43.

Muneer, T., Younes, S., Munawwar, S., 2007. Discourses on solar radiation modeling. Renew. Sustain. Energy Rev. 11, 551–602.

Ohmura, A., Dutton, E.G., Forgan, B., Fröhlich, C., Gilgen, H., Hegner, H., Heimo, A., König-Langlo, G., McArthur, B., Müller, G., Philipona, R., Pinker, R., Whitlock, C.H., Dehne, K., Wild, M., 1998. Baseline Surface Radiation Network (BSRN/WCRP): new precision radiometry for climate research. Bull. Am. Meteorol. Soc. 79, 2115–2136.

Oliveira, A.P., Escobedo, J.F., Machado, A.J., Soares, J., 2002. Correlation models of diffuse solar-radiation applied to the city of Sao Paulo, Brazil. Appl. Energy 71, 59–73.

Orgill, J.F., Hollands, K.G.T., 1977. Correlation equation for hourly diffuse radiation on a horizzontal surface. Sol. Energy 19, 357–359.

Oumbe, A., Qu, Z., Blanc, P., Bru, H., Wald, L., 2012. An adaptive method to derive direct irradiance from global irradiance. In: Proc. SolarPACES Conf., International Energy Agency, Marrakech, Morocco.

Pagola, I., Gaston, M., C, F., Torres, J.L., Silva, M., Ramirez, L., 2009. Comparison and fitting of several global-to-beam irradiance models in Spain. In: Proc. SolarPACES Conf., International Energy Agency, Berlin, Germany.

Perez, P., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., Vignola, F., 2002. A new operational model for satellite-derived irradiances: description and validation. Sol. Energy 73, 307–317.

Perez, R., Ineichen, P., Maxwell, E.L., Seals, R., Zelenka, A., 1992. Dynamic global-to-direct irradiance conversion models. ASHRAE Trans. 98 (1), 354–369.

Perez, R., Ineichen, P., Seals, R., Zelenka, A., 1990a. Making full use of the clearness index for parameterizing hourly insolation conditions. Sol. Energy 45, 111–114.

Perez, R., Seals, R., Zelenka, A., Ineichen, P., 1990b. Climatic evaluation of models that predict hourly direct irradiance from hourly global irradiance: prospects for performance improvements. Sol. Energy 44, 99–108.

Perez-Burgos, A., Bilbao, J., De Miguel, A., Roman, R., 2014. Analysis of solar direct irradiance in Spain. Energy Proc. 57, 1010–1076.

Piacentini, R.D., Salum, G.M., Fraidenraich, N., Tiba, C., 2011. Extreme total solar irradiance due to cloud enhancement at sea level of the NE Atlantic coast of Brazil. Renew. Energy 36, 409–412.

Piedehierro, A.A., Anton, M., Cazorla, A., Alados-Arboledas, L., Olmo, F.J., 2014. Evaluation of enhancement events of total solar irradiance during cloudy conditions at Granada (Southeastern Spain). Atmos. Res. 135, 1–7.

Polo, J., Antonanzas-Torres, F., Vindel, J.M., Ramirez, L., 2014. Sensitivity of satellite-based methods for deriving solar radiation to different choice of aerosol input and models. Renew. Energy 68, 785–792.

Posadillo, R., Lopez Luque, R., 2009. Hourly distributions of the diffuse fraction of global solar irradiation in Córdoba (Spain). Energy Convers. Manag. 50, 223–231.

Posadillo, R., Lopez Luque, R., 2010. The generation of hourly diffuse irradiation: a model from the analysis of the fluctuation of global irradiance series. Energy Convers. Manag. 51, 627–635.

Reindl, D.T., Beckman, W.A., Duffie, J.A., 1990. Diffuse fraction correlations. Sol. Energy 45, 1–7.

Remund, J., Salvisberg, E., Kunz, S., 1998. On the generation of hourly shortwave radiation data on tilted surfaces. Sol. Energy 62, 331–344.

Remund, J., Wald, L., Lefèvre, M., Ranchin, T., Page, J., 2003. Worldwide Linke turbidity information. Proc. ISES Conf.. International Solar Energy Society, Stockholm, Sweden.

Rerhrhaye, A., Zehaf, M., Flechon, J., 1995. Estimation of the direct beam from seasonal correlations. Renew. Energy 6, 779–785.

Ridley, B., Boland, J., Lauret, P., 2010. Modelling of diffuse solar fraction with multiple predictors. Renew. Energy 35, 478–483.

Ridley, B., Boland, J., Luther, M., 2004. Quality control of climate data sets. In: Proc. Solar 2004: Life, the Universe and Renewables, Murdoch University, ANZSES, Australia.

Roesch, A., Wild, M., Ohmura, A., Dutton, E.G., Long, C.N., Zhang, T., 2011. Assessment of BSRN radiation records for the computation of monthly means. Atmos. Meas. Tech. 4, 339–354, Corrigendum, <http://www.atmos-meas-tech.net/334/973/2011/>.

Ruiz-Arias, J.A., Alsamamra, H., Tovar-Pescador, J., Pozo-Vazquez, D., 2010. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. Energy Convers. Manag. 51, 881–893.

Sanchez, G., Cancillo, M.L., Serrano, A., 2012. Adapting the Spencer model for diffuse solar radiation in Badajoz (Spain). Opt. Pura Apl. 45, 5–9.

Sengupta, M., Keller, J.L., 2012. PV ramping in a distributed generation environment: a study using solar measurements. Proc. IEEE Photovoltaic Specialists Conference. IEEE, Austin, TX.

Shen, C., He, Y.-L., Liu, Y.-W., Tao, W.-Q., 2008. Modelling and simulation of solar radiation data processing with Simulink. Simul. Modell. Pract. Theory 16, 721–735.

Skartveit, A., Olseth, J.A., 1987. A model for the diffuse fraction of hourly global radiation. Sol. Energy 38, 271–274.

Skartveit, A., Olseth, J.A., Tuft, M., 1998. An hourly diffuse fraction model with correction for variability and surface albedo. Sol. Energy 63, 173–183.

Soares, J., Oliveira, A.P., Boznar, M.Z., Mlakar, P., Escobedo, J.F., Machado, A.J., 2004. Modeling hourly diffuse solar-radiation in the city of Sao Paulo using a neural-network technique. Appl. Energy 79, 201–214.

SOLMET, 1977. Rehabilitation of hourly solar radiation data, vol. 2. National Climatic Center, Asheville, NC.

Spencer, J.W., 1982. A comparison of methods for estimating hourly diffuse solar radiation from global solar radiation. Sol. Energy 29, 19–32.

Please cite this article in press as: Gueymard, C.A., Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010

Stauter, R., Klein, S.A., 1980. Unpublished work. In: Duffie, J.A., Beckman, W. (Eds.), Solar Engineering of Thermal Processes, Wiley, New York.

Suehrcke, H., McCormick, P.G., 1988. The diffuse fraction of instantaneous solar radiation. Sol. Energy 40, 423–430.

Tamura, J., Kurokawa, K., Otani, K., 2003. Estimation of hourly in-plane irradiation by using minutely horizontal data. Sol. Energy Mater. Sol. Cells 75, 585–595.

Tapakis, R., Charalambides, A.G., 2014. Enhanced values of global irradiance due to the presence of clouds in Eastern Mediterranean. Renew. Energy 62, 459–467.

Tapakis, R., Michaelides, S., Charalambides, A.G., 2015. Computations of diffuse fraction of global irradiance: Part 1 – analytical modelling. Sol. Energy (in press), http://dx.doi.org/10.1016/j.solener.2014.10.005.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. 106D, 7183–7192.

Torres, J.L., de Blas, M., Garcia, B., de Francisco, A., 2010. Comparative study of various models in estimating hourly diffuse solar irradiance. Renew. Energy 35, 1325–1332.

Tsubo, M., Walker, S., 2003. Relationships between diffuse and global solar radiation in southern Africa. South Afr. J. Sci. 99, 360–362.

Tuomiranta, A., Ghedira, H., 2015. Evaluation of decomposition and transposition models for irradiance data conversion under a hot desert climate. In: Proc. 3rd Int. Conf. Energy & Meteorology (ICEM), Boulder, USA.

Turner, W.D., Salim, M., 1984. Comparison of two diffuse sky radiation models. Sol. Energy 32, 677–679.

Udagawa, M., Kimura, K., 1978. Estimation of direct radiation from observations of horizontal solar radiation. Trans. AIJ 267, 83–89.

Ulgen, K., Hepbasli, A., 2002. Prediction of solar radiation parameters through clearness index for Izmir, Turkey. Energy Sources 24, 773–785.

Vick, B.D., Myers, D.R., Boyson, W.E., 2012. Using direct normal irradiance models and utility electrical loading to assess benefit of a concentrating solar power plant. Sol. Energy 86, 3519–3530.

Vindel, J.M., Polo, J., Antonanzas-Torres, F., 2014. Improving daily output of global to direct solar irradiance models with ground measurements. J. Renew. Sustain. Energy 5, 063123.

Watanabe, T., Urano, Y., Hayashi, T., 1983. Procedures for separating direct and diffuse insolation on a horizontal surface and prediction of insolation on tilted surfaces. Trans. AIJ 330.

WCRP, 2001. Report of the 6th BSRN Science and Review Workshop (Melbourne, Australia, 1–5 May 2000). WCRP Informal Report No. 17/2001.

Yao, W., Li, Z., Lu, Y., Jiang, F., Li, C., 2013. New models for separating hourly diffuse and direct components of global solar radiation. In: Proc. 8th International Symposium on Heating, Ventilation and Air Conditioning, Springer, Xian, China.

Yordanov, G.H., Saetre, T.O., Midtgard, O.M., 2015. Extreme overirradiance events in Norway: 1.6 suns measured close to 60N. Sol. Energy 115, 68–73.

Younes, S., Claywell, R., Muneer, T., 2005. Quality control of solar radiation data: present status and proposed new approaches. Energy 30, 1533–1549.

Zhang, Q.Y., Lou, C., Yang, H., 2004. A new method to separate horizontal solar radiation into direct and diffuse components. In: Proc. International Solar Energy Society, Asia-Pacific Conf., ISES, Gwangju, Korea.

Please cite this article in press as: Gueymard, C.A.,, Ruiz-Arias, J.A. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy (2015), http://dx.doi.org/10.1016/j.solener.2015.10.010