Impact of local broadband turbidity estimation on forecasting of clear sky direct normal irradiance

Rich H. Inman, James G. Edson, Carlos F.M. Coimbra*

Department of Mechanical and Aerospace Engineering, Jacobs School of Engineering, Center for Renewable Resource Integration and Center for Energy Research, University of California, San Diego, La Jolla, CA 93093, USA

Received 25 March 2014; received in revised form 13 April 2015; accepted 26 April 2015

Available online 16 May 2015

Communicated by: Associate Editor Christian A. Gueymard

Abstract

Clear-sky modeling is of critical importance for the accurate determination of Direct Normal Irradiance (DNI), which is the relevant component of the solar irradiance for concentrated solar energy applications. Accurate clear-sky modeling of DNI is typically best achieved through the separate consideration of water vapor and aerosol concentrations in the atmosphere. Highly resolved temporal measurements of such quantities is typically not available unless a meteorological station is located in close proximity. When this type of data is not available, attenuating effects on the direct beam are modeled by Linke turbidity-equivalent factors, which can be obtained from broadband observations of DNI under cloudless skies. We present a novel algorithm that allows for a time-resolved estimation of the average daily Linke turbidity factor from ground-based DNI observations under cloudless skies. This requires a method of identifying clear-sky periods in the observational time series (in order to avoid cloud contamination) as well as a broadband turbidity-based clear-sky model for implicit turbidity calculations. While the method can be applied to the correction of historical clear-sky models for a given site, the true value lies in the forecasting of DNI under cloudless skies through the assumption of a persistence of average daily turbidity. This technique is applied at seven stations spread across the states of California, Washington, and Hawaii while using several years of data from 2010 to 2014. Performance of the forecast is evaluated by way of the relative Root Mean Square Error (rRMSE) and relative Mean Bias Error (rMBE), both as a function of solar zenith angle, and benchmarked against monthly climatologies of turbidity information. Results suggest that rRMSE and rMBE of the method are typically smaller than 5% for both historical and forecasted CSMs, which compare favorably against the 10–20% range that is typical for monthly climatologies.

© 2015 Elsevier Ltd. All rights reserved.

Keywords: Direct normal irradiance; Linke turbidity; Solar forecasting; Aerosols

1. Introduction

Knowledge of clear-sky irradiance plays a critical role in several solar engineering applications including the definition of clear-sky indices, the development of smart persistence forecasts, the normalization of information retrieved from satellite data, and the calculation of forecasting skill metrics (Inman et al., 2013). In particular, clear-sky modeling is essential for the accurate determination of Direct Normal Irradiance (DNI) under cloudless skies. DNI is the critical component of the solar irradiance for Concentrated Solar Power (CSP) applications such as the recently completed Ivanpah Solar Electric Generating System located in the Mojave Desert of California, which at the time of writing is the largest solar thermal project in the world (392 MW). Although CSP technologies currently represent only a small fraction of renewable energy
portfolios on a global scale, annual energy generated from such technologies are expected to exceed 30 TW h by 2017 (IEA, 2013).

Increased CSP market share will require further policy action to tackle technical and financing challenges that currently hinder deployment (IEA, 2014). One approach to lower the cost of grid integration is the application of DNI/CSP forecasting, which makes CSP plants more financially attractive to deploy. These forecasts are used to determine optimal operational strategies that maximize profit by minimizing penalty charges resulting from differences between plant output and forecasted output. As a result of CSP plants being driven by DNI, the determination of such optimal operational strategies for CSP plants depends strongly on the accuracy of DNI forecasting. In general, forecast uncertainties are driven by variability in cloud cover. However, it is well known that under cloudless skies the presence of aerosol particles and water vapor become the most important factors influencing the intensity of ground level DNI (Gueymard, 2012a; Eltbaakh et al., 2012).

The term aerosol is used to describe either liquid or solid particles that are suspended in the atmosphere with sizes ranging from 1 to $10^3$ nm in radius (Wen, 1996). Aerosols in the atmosphere may be either natural or anthropogenic in origin and include particles such as fine soil, pollen, and microorganisms lifted by the wind; sea salts escaping from breaking waves; smoke and soot emitted from fires; ash and dust erupting from volcanoes; sulfates created by the burning of coal and oil; and black carbon released during the incomplete combustion of heavy petroleum products. The net effect of aerosols on local microclimates depend on three primary mechanisms: direct radiative forcing as a result of scattering and absorption of visible and infrared radiation in the atmospheric boundary layer, indirect radiative forcing associated with changes in the microphysical and optical properties of cloud fields, and local heating in the cloud formation layer due to highly absorbent aerosols such as black carbon (Stefan et al., 2006). Although it is clear that DNI attenuation under cloudless skies is driven by aerosol variability, the magnitude of these influences is poorly constrained as a result of the highly spatial-temporal variability of aerosol particles in the atmosphere as well as the fragmentary knowledge of the processes which control the physical, chemical, and optical properties of aerosol distributions (Morcrette et al., 2003; Kaskaoutis and Kambezidis, 2008; Chin et al., 2002).

Several methods for the quantification of atmospheric aerosol loading are available in the literature including both ground-based and remote sensing techniques (see Section 2). While current satellites provide daily multi-wavelength AOD data for nearly any location on the planet, their quality is questionable at times as a result of missing pixels in AOD retrievals and cloud contamination. Ground-based pyrhiometers, on the other hand, are typically located at CSP sites and offer a highly resolved temporal signal of DNI, which under cloudless skies is related to atmospheric aerosol loading. Therefore, highly temporally resolved ground-based observations of DNI under clear skies allow for a robust sampling of local turbidity, specifically at locations of interest to CSP plant operators.

Several authors have examined the derivation of atmospheric aerosol loading from broadband irradiance measurements. More specifically, Louche et al. (1987) assigned a fixed value to the Ångström exponent $\alpha$ (see Section 2) and calculated the Ångström turbidity coefficient $\beta$ from DNI observations over Ajaccio (France). Gueymard and Vignola (1998) developed a semi-empirical model that demonstrated the utility of the diffuse component of broadband irradiance for estimating atmospheric turbidity. Canada et al. (1993) also assigned a fixed value to $\alpha$ in order to estimate $\beta$ in Valencia (Spain) and compared the results with those from Ajaccia, Avignon, and Dhahran. Ineichen (2008) presented a conversion function between $T_{Li}$, the atmospheric water vapor and urban aerosol content that also accounts for the altitude of the application site.

More recently, Polo et al. (2009) proposed a method to estimate daily Linke turbidity factor by using global irradiance measurements at solar noon. Gueymard (2013) provided an efficient method to derive Aerosol Optical Depth (AOD) information from broadband DNI measurements and addressed several critical issues including: instrument error, impact of model performance, propagation of errors due to incorrect precipitable water, elimination of cloudy conditions, and evaluation of $\alpha$. Gueymard (2014) also evaluated the impact of on-site atmospheric water vapor estimation methods on the accuracy of local solar irradiance predictions. Bilbao et al. (2014), proposed a method for deriving Ångström’s turbidity coefficient and the AOD at 550 nm from broadband DNI observations over Castilla y León (Spain), from July 2010 to December 2012.

In addition to the implicit calculation of aerosol loading from irradiance observations, detailed algorithms exist in the literature that produce aerosol forecasts for aerosol fields using remote sensing techniques and transport models, see for example Masmoudi et al. (2003). However, this contribution demonstrates the utility of ground-based estimations of average daily turbidity for the day-ahead forecasting of broadband clear-sky DNI at a specific site. Because of CSP’s dependance on broadband DNI, this irradiance component tends to be measured on site resulting in readily available broadband turbidity information at such locations. An endogenous clear-sky detection algorithm for DNI is developed, which is based on the work of Reno et al. (2012), and applied to nearly ten site-years of data from seven stations spread across the states of California, Washington, and Hawaii (see Fig. 3). These data represent a number of widely varying microclimates which are used to speak to the robustness of the algorithm. Observations of clear-sky DNI are subsequently used to calculate the daily average air mass independent Linke
turbidity factor $T_{LI}$. A day ahead forecast of clear-sky DNI is then issued under the persistence assumption of average daily turbidity. The observed DNI under clear skies can then be used to calculate the forecast’s error. In order to demonstrate the value of using daily values of turbidity, the method is compared against the monthly climatologies from Remund et al. (2003).

Section 2 describes quantification of aerosol loading. Section 3 covers the development of the airmass independent Linke turbidity factor, Section 4 details the experimental data, Section 5 covers the endogenous clear-sky detection algorithm, Section 6 describes the methodology, Section 7 describes the metrics used for comparison, and Sections 8 and 9 contain the results and conclusions respectively.

2. Aerosol loading

While many formulations exist in the literature for the quantification of aerosol loading, typically, this is achieved through the use of Ångström’s Law,

$$\tau(\lambda) = \beta \lambda^{-z}$$  \hspace{1cm} (1)

where $\tau(\lambda)$ is the monochromatic aerosol attenuation coefficient, or Aerosol Optical Depth (AOD); $\beta$ is the Ångström turbidity coefficient, which represents the AOD at a wavelength of 1 μm; $\lambda$ denotes the wavelength (μm); and $z$ is the Ångström exponent, which is a function the size distribution of the aerosols. The Ångström turbidity coefficient does not incorporate the optical depth of water vapor or other participating gasses and therefore provides an uncontaminated measure of atmospheric aerosol loading. However, $z$ and $\beta$ are known to exhibit daily fluctuations driven by atmospheric processes such as photochemical activity, local emissions, mesoscale circulation, wind ventilation, and changes in temperature resulting from phase changes in the atmosphere (Bilbao et al., 2014). As a result, capturing the intra-daily variability in AOD requires highly resolved temporal data of AOD, which is typically not available without a ground-based sunphotometer in close proximity (e.g., NASA’s Aeronet network).

Other sources of spectrally resolved AOD and water vapor data exist, specifically those that involve remote sensing (e.g., MODIS onboard NASA’s AQUA/TERRA), however, it is known that such data may contain significant errors especially over regions of low turbidity and high albedo, such as the southwest US (Guemymard, 2013; Ruiz-Arias et al., 2013). Furthermore, due to the nature of the Polar Low-Earth Orbit (PLEO) of the TERRA and AQUA satellites, which descend/ascend across the equator at about 10:30 a.m. and 1:30 p.m. local time respectfully, a strict sampling is imposed. Often clouds are present during a satellite’s pass over a location (see Appendix A), which can result in cloud contamination or missing pixels in AOD retrievals; particularly at high spatial and temporal resolutions. It should be noted that these issues are typically addressed by combining aerosol fields retrieved from both TERRA and AQUA MODIS imagery to produce a more complete data set of sufficiently relaxed spatial resolution (e.g. $1^° \times 1^°$ used by MODIS). Until the launch of new generation instruments on geostationary satellites that provide a constant full-field view, such as GOES-R, the sampling rate of remote sensing techniques may be unsuitable for the accurate determination of highly resolved intra/inter-daily turbidity information. As a result, it may be desirable to use other sources of data with higher sampling rates such as broadband DNI observations at ground level, which are typically co-located with CSP plants.

Ground-based observations of broadband DNI under clear-skies can be used to characterize the behavior of an aerosol and moisture laden atmosphere through broadband turbidity indices, which quantify the atmosphere’s overall level of opacity to broadband solar radiation. While this method offers a less accurate broadband quantification of aerosol and water loading, it is much simpler to implement and its errors are acceptable for the purposes of day ahead clear-sky DNI forecasts. The most commonly used broadband turbidity coefficients are the Unsworth–Monteith turbidity coefficient $\tau_a$, and the Linke turbidity factor $T_L$.

Unsworth and Monteith (1972) defined $\tau_a$ using the following expression

$$D_c(\tau_a) = D_0 \cdot \exp(-\tau_a \cdot m),$$  \hspace{1cm} (2)

which relates the measured clear-sky direct normal irradiance $D_c(\tau_a)$ to the direct irradiance calculated for a dust-free atmosphere $D_0$ where the air mass is $m$. Similarly, Linke (1922, 1929) defined his turbidity factor as the number of clean dry atmospheres required to reproduce the attenuation of extra-terrestrial radiation caused by the real atmosphere as

$$D_c(T_L) = I_0 \cdot \exp(-\delta_{cda} \cdot T_L \cdot m),$$  \hspace{1cm} (3)

where $D_c(T_L)$ is the observed DNI under cloudless skies, $I_0$ is the sun-earth distance corrected extraterrestrial irradiance, $\delta_{cda}$ is the optical thickness of a clear and dry atmosphere, and $m$ is the air mass. It is known that $\tau_a$ and $T_L$ are both marginally dependent on air mass $m$ (Guemymard, 1998). More recently, however, Ineichen and Perez (2002) developed an airmass independent formulation of Linke turbidity factor $T_{LI}$, which provides greater stability throughout the day, (see Fig. 1). Unlike the rigid sampling rate of remote sending techniques, the increased stability of $T_{LI}$ allows for a flexible sampling of atmospheric turbidity at any time that the circum-solar region is not obscured by clouds.

The Linke turbidity factor has also become a key input for a number of Clear Sky Models (CSMs) (Page, 1996; Dumortier, 1995; Rigollier et al., 2000; Perez et al., 2002; Lefèvre et al., 2004; Remund and Page, 2002). On the basis of the widespread use of the Linke turbidity factor for clear-sky modeling, a number of regional and worldwide maps of monthly/seasonal Linke turbidity have become available in the literature (Perez et al., 2002;
Remund et al., 2003), an example of which is shown in Fig. 2. While the monthly/seasonal resolution of these maps are convenient for locations where no ground truth exists and are typically sufficient for the clear-sky modeling of GHI (Reno et al., 2012), they are not suitable for the accurate modeling of DNI, which demonstrates an increased sensitivity to variability in aerosol loading.

In general, the scattering caused by aerosols under cloudless skies increases the local Diffuse Horizontal Irradiance (DHI) and commensurately decreases the DNI. This exchange between DHI and DNI effectively buffers the impact of aerosols on clear-sky Global Horizontal Irradiance (GHI), which can be seen in the closure equation,

\[ G = D \cos \theta_z + d, \]

where the global horizontal irradiance \( G \) is defined as the geometric sum of the direct normal irradiance \( D \) and diffuse horizontal irradiance \( d \), and \( \theta_z \) is the solar zenith angle. This buffering is the reason aerosol loading is known to attenuate DNI in the range of 30–100% as opposed to GHI where the attenuation is significantly lower at approximately 10% (Schroedter-Homscheidt and Oumbe, 2013; Lara-Danego et al., 2012; Marquez and Coimbra, 2011). Furthermore, Gueymard (2012b) showed that DNI exhibits an Aerosol Sensitivity Index (ASI), which relates the magnitude of relative variations in irradiance to absolute variations in aerosol optical depth, that is 2–4 times greater than that for GHI. The difference in intensity reduction between GHI and DNI is one key reason why GHI forecasting techniques, which have been reviewed extensively in the literature (Gueymard, 2012a; Reno et al., 2012;
should not be directly applied to DNI forecasting. Therefore, it is desirable to calculate spatially and temporally local Linke turbidity information through ground-based observations of broadband DNI under clear skies.

The Linke turbidity factor $T_L$ is, among others, a turbidity index which is calculated by inverting a model which estimates broadband DNI under cloudless skies and thus refers to the spectrally integrated attenuation (Louche et al., 1986; Eltbaakh et al., 2012). As a result, $T_L$ incorporates the optical depth of water vapor and is not a pure turbidity coefficient (Gueymard, 1998). Nevertheless, because $T_L$ is empirically derived from DNI, which is measured at many meteorological stations, its use still remains widespread.

The turbidity factor defined by Linke (1922, 1929) can be expressed explicitly by solving Eq. (3) for $T_L$,

$$T_L = \frac{1}{\delta_{\text{dia}} - m} \ln \left( \frac{I_0}{D_c} \right),$$

where it is clear that $T_L$ is dependent upon the theoretical value of $\delta_{\text{dia}}$ employed. Later, Kasten (1980) made use of a series of spectral data tables published by Feusner and Dubois (1930) to developed the widely used pyrheliometric formula:

$$\delta_{\text{dia}} = (9.4 + 0.9 \cdot m)^{-1},$$

which gave rise to the definition of the Linke turbidity factor as corrected by Kasten $T_{LK}$:

$$T_{LK} = 9.4 + 0.9 \cdot m \cdot \ln \left( \frac{I_0}{D_c} \right).$$

More recently, Ineichen and Perez (2002) showed that both $T_L$ and $T_{LK}$ exhibit diurnal variations, and thus, are a function of air mass even when the aerosol optical depth remains relatively constant. As a result, Ineichen and Perez (2002) proposed an air mass-independent Linke turbidity factor $T_{LI}$, where similar to the work of Rigollier et al. (2000), $T_{LI}$ at $m = 2$ was used as a reference. The Linke turbidity factor as corrected by Ineichen $T_{LI}$ is obtained by inverting the following model for DNI under clear skies:

$$D_c = b \cdot I_0 \cdot \exp \left[ -0.09 \cdot m \cdot \left( T_{LK}^{m-2} - 1 \right) \right].$$

where

$$b = 0.664 + \frac{0.163}{e^{-8/8.000}},$$

has been adopted from Kasten (1984) and $h$ is the elevation in meters. Seasonal trends of $T_{LI}$ were evaluated with (7) at $m = 2$ under cloudless skies after which the expression in (8) was inverted to give $T_{LI}$ as:

$$T_{LI} = \frac{1}{m} \ln \left( \frac{b \cdot I_0}{D_c} \right) + 1,$$

which has the advantages of (1) being independent of the air mass and (2) matching $T_{LK}$ at air mass 2.

Fig. 1 shows the evolution of broadband DNI at 30 s resolution measured with a Normal Incidence Pyrheliometer (NIP) and SMT-3 solar tracker (both available from the Eppley Laboratory, Inc) as well as $T_L$, $T_{LK}$, and $T_{LI}$ during the 3rd of April 2011 at the University of California, Merced. It is important to pause here and discuss the significantly different magnitudes for the three determinations of the Linke turbidity factor, which were also shown by Ineichen and Perez (2002). Contrarily to AOD, the Linke turbidity factor does not represent a pure optical characteristic of aerosol extinction, but is also affected by the parasitic effects of air mass and water vapor. In addition, Linke turbidity cannot be measured directly, and must therefore be derived from DNI measurements by inverting an appropriate solar radiation model. As a result, any published value of Linke turbidity is more or less dependent on water vapor column amount and model used. However, a formulation of turbidity that does not vary with relatively uniform aerosol loading is desirable as it can be used to obtain a representative turbidity value under cloudless skies at any air mass. It is clear from Fig. 1 that $T_{LI} = T_{LK} @ AM = 2$ and the stability of $T_{LI}$ throughout the day is much better than $T_L$ and $T_{LK}$. It is on this basis that $T_{LI}$ is employed in this work.

### 3. Experimental data

Data is used from seven stations spread across the states of California, Washington and Hawaii spanning the years 2011–2014. The locations of these stations possess varying latitudes, and microclimates; including:

- **Bellingham, WA**: latitude, 48.8°N; longitude, 122.5°W; altitude, 6 m; temporal resolution, 60 s; instrumentation, Irradiance Inc. RSR2.
- **Berkeley, CA**: latitude 37.9°N; longitude 122.3°W; altitude, 109 m; temporal resolution, 60 s; Yankee Environmental Systems, Inc. MFR-7.
- **Davis, CA**: latitude, 38.5°N; longitude, 121.7°W; altitude, 19 m; temporal resolution, 60 s; Yankee Environmental Systems, Inc. MFR-7.
- **Ewa Beach, HI**: latitude, 21.3°N; longitude, 158.0°W; altitude, 10 m; temporal resolution, 60 s; instrumentation, Irradiance Inc. RSR2.
- **Folsom, CA**: latitude, 38.6°N; longitude, 121.4°W; altitude, 97 m; temporal resolution, 60 s; instrumentation, Irradiance Inc. RSR2.
- **La Jolla, CA**: latitude, 32.9°N; longitude, 117.2°W; altitude, 106 m; temporal resolution, 30 s; Eppley Laboratory, Inc. NIP on SMT tracker and PSP.
- **Merced, CA**: latitude, 37.3°N; longitude, 120.4°W; altitude, 65 m; temporal resolution, 30 s; Eppley Laboratory, Inc. NIP on SMT tracker and PSP.

Each of the databases contains anywhere from 9 to 26 months of measurements of all three irradiance components at either a 30 or 60 s sampling rate. The duration of the data acquisition as well as the instrumentation used to take observations are listed in Fig. 4.
The LI-COR LI-200SZ utilized in the rotating shadow-band radiometer head unit of the RSR2 (Irradiance Inc.) is a relatively simple device that requires software to correct for cosine errors. This sensor has a typical cosine error of less than 5% up to an 80° angle of incidence. For sun plus sky at a sun elevation of 30° (60° angle of incidence), the error is ≈2%. Similarly, the MFR-7 has a novel input optic with excellent cosine response and long term stability. The individual MFR-7 cosine response, supplied with each instrument, is used by YESDAS system software to correct, in real time, for deviations from the ideal cosine response. Similar to the LI-200SZ, the error is typically ≲2% at a 60° angle of incidence. The Eppley NIP pyrheliometer (mounted on SMT tracker) and PSP pyranometer are secondary standard instruments with uncertainties of less than 1.5% and 3.5%, respectively.

At all of the stations, GHI was observed independently from DNI, which allows for enhanced quality control using the summation method to obtain global irradiance (Gueymard and Myers, 2009). Observations where the ratio of calculated GHI, using Eq. (4), and observed GHI differs from unity by more than 20% were excluded as they typically correspond to an error in at least one of the irradiance components.

4. Clear sky detection

Clear-sky detection through cloud filtering is vital to the success of the methodology, such that only DNI observations that represent clear-sky conditions within the field of view of the radiometer are used. Typically, two approaches have been employed: (1) direct filtering of

![Fig. 4. Plot of the observed average daily $T_{LI}$ for days where clear skies were detected, the tabulated monthly average from Remund and Page (2002) and the fraction of measurements which were classified as clear for the entire data set used in this work.](image-url)
cloud conditions in the DNI time-series or (2) back-end filtering where all DNI data is used and observations that provide physically un-reasonable clear-sky conditions are eliminated.

The first approach has been implemented by Long and Ackerman (2000), however, the Fortran files and executables of the algorithm have been found to be slow and cumbersome to use (Gueymard, 2013). The second approach was adopted by AERONET for cloud-screening from sunphotometric observations (Smirnov et al., 2000) and more recently Gueymard (2013) employed a similar technique based on five independent requirements.

In this work, rather than directly implement the Long filter or adopt a back-end filtering, clear-sky detection is performed using an endogenous statistical model which was originally developed by Reno et al. (2012) for GHI observations. This method uses five criteria to compare a time period containing \( N \) observations of GHI to a corresponding CSM for the same period. The time period is deemed “clear” if threshold values (whose values vary with \( N \)) for each criteria are met.

In this work, the criteria include the mean value of irradiance \( I \) (global and direct) during the time period,

\[
\bar{G} = \frac{1}{N} \sum_{t=1}^{N} I(t),
\]

the maximum irradiance value \( M \) in the time series,

\[
M = \max[I(t)] \quad \forall \quad t \in \{1, 2, \ldots, N\},
\]

the length \( L \) of the line connecting the points in the time series that, unlike the \( L \) defined in Reno et al. (2012), does not account for the length of the time-step,

\[
L = \sum_{t=1}^{N} \sqrt{(I(t + \Delta t) - I(t))^2},
\]

the variance of changes in the time series; specifically, the normalized standard deviation \( \sigma \) of the slope between sequential points,

\[
\sigma = \frac{\sqrt{\sum_{t=1}^{N-1} (s(t) - \bar{s})^2}}{\bar{s} \sum_{t=1}^{N} I(t)}
\]

where

\[
s(t) = I(t + \Delta t) - I(t) \quad \forall \quad t \in \{1, 2, \ldots, N\}
\]

\[
\bar{s} = \frac{1}{N - 1} \sum_{t=1}^{N-1} s(t)
\]

and, \( S \) the maximum deviation from the clear-sky slope \( s_c \),

\[
S = \max\{|s(t) - s_c(t)|\} \quad \forall \quad t \in \{1, 2, \ldots, N\}
\]

where

\[
s_c(t) = L_c(t + \Delta t) - I_c(t).
\]

Similar to Reno et al. (2012), in this work a 10-min sliding window is employed to determine if an observation is identified as “clear”. The five clear-sky criteria \( (\bar{G}, M, L, \sigma, S) \) are evaluated for both the CSM and observational time series for each 10-min window. The window then progresses forward in time by one time step each iteration. Accordingly, for a 1 min sampling rate, each irradiance measurement would be evaluated 10 times as the window steps forward in time. A measurement is identified as clear as long as it is within at least one coherent window declared as clear during the evaluations. In other words, the sliding window checks every combination of 10 contiguous observations, and if at least one window meets all five criteria, then all 10 measurements in that window must be clear. In this way, each individual observation may be classified as clear or cloudy (Reno et al., 2012). Fig. 5 shows the observed GHI, the observed CSMs calculated from SoDa turbidity values, as well as the values of the metrics centered on each of the 10 min sliding windows. It is clear that the variability caused by clouds violates at least one of the criteria in all cases and that the clear periods meet all five criteria.

Both DNI and GHI data were available from all stations and, as a result, in this work both are used to determine clear-sky criteria. Each measurement is classified as clear only if threshold values for all five of the clear-sky criteria are met for both the DNI and GHI time series in at least one coherent window. This approach results in a more robust clear-sky determination protocol as it is known that the global irradiance is less sensitive to thin clouds that act as diffusers. The threshold values from Reno et al. (2012) as well as the values used in this work are listed in Table 1. In addition, the fraction of the day time irradiance measurements classified as clear as well as the calculated average daily \( T_{LI} \) are given for each day in Fig. 4. It should be noted that clear-sky DNI observations are required for the calculation of \( T_{LI} \) and, consequently, no turbidity information is available for days which lack clear-sky observations.

5. Method

It is well known that under cloudless skies the presence of aerosol particles and water vapor in the atmosphere are among the most important factors affecting the attenuation of downwelling irradiance; particularly in the case of DNI which is highly sensitive to the optical properties of the atmospheric column. It is the highly variable nature of aerosol particles and water vapor in the atmosphere that serves as motivation for including spatially and temporally local broadband turbidity information for the reduction of DNI forecasting errors. In order to achieve this a number of techniques are employed.

First, it should be noted that, observations corresponding to solar zenith angles greater than 75° were removed from the data set so that only daylight hours with reasonable air mass are included. For each month in the data set, turbidity information for the location is looked up from the monthly SoDa world maps of Linke turbidity (Remund and Page, 2002). This monthly turbidity approximation is then used as an input to calculate the DNI CSM for the entire month using Eq. (8) and assuming that \( T_{LI} = T_{LI}^{\text{mean}} \)
The airmass independent Linke turbidity factor $T_{LI}$ is calculated from DNI measurements, which is equal to Kasten’s Linke turbidity factor $T_{LK}$, see Fig. 6(a). After this initial CSM and $T_{LI}$ are calculated, its general shape is used to perform a binary clear-sky detection as described in Section 4, see Fig. 6(b). For each of the partially clear days an average $T_{LI}$ is calculated which is then used to as an input to Eq. (8) to re-calculate the CSM, see Fig. 6(d). Each of the average daily $T_{LI}$ is shown in Fig. 4 for days where clear-sky observations exist.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>$\text{GHI}_{\text{Reno}}$</th>
<th>$\text{GHI}_{\text{Inman}}$</th>
<th>$\text{DNI}_{\text{Inman}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G$</td>
<td>$\pm 75 \text{ W m}^{-2}$</td>
<td>$\pm 100 \text{ W m}^{-2}$</td>
<td>$\pm 200 \text{ W m}^{-2}$</td>
</tr>
<tr>
<td>$M$</td>
<td>$\pm 75 \text{ W m}^{-2}$</td>
<td>$\pm 100 \text{ W m}^{-2}$</td>
<td>$\pm 200 \text{ W m}^{-2}$</td>
</tr>
<tr>
<td>$L$</td>
<td>$-5 \text{ to } 10$</td>
<td>$\pm 50$</td>
<td>$\pm 100$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$&lt; 0.005$</td>
<td>$&lt; 0.01$</td>
<td>$&lt; 0.015$</td>
</tr>
<tr>
<td>$S$</td>
<td>$&lt; 8 \text{ W m}^{-2}$</td>
<td>$&lt; 10 \text{ W m}^{-2}$</td>
<td>$&lt; 15 \text{ W m}^{-2}$</td>
</tr>
</tbody>
</table>

Fig. 5. Observed GHI and DNI, the CSMs calculated from persistence turbidity values, as well as the value of them metric centered at the 10 min sliding window. It is clear that the variability caused by clouds violates at least one of the criteria in all cases and that the clear periods meet all five criteria.

as well as the GHI CSM from Ineichen and Perez (2002), see Fig. 6(a). The airmass independent Linke turbidity factor $T_{LI}$ is calculated from DNI measurements, which is equal to Kasten’s Linke turbidity factor $T_{LK}$ @ $m = 2$, see Fig. 6(b). After this initial CSM and $T_{LI}$ are calculated, its general shape is used to perform a binary clear-sky detection as described in Section 4, see Fig. 6(c). For each of the partially clear days an average $T_{LI}$ is calculated which is then used to as an input to Eq. (8) to re-calculate the CSM, see Fig. 6(d). Each of the average daily $T_{LI}$ is shown in Fig. 4 for days where clear-sky observations exist.
A day-ahead DNI clear-sky forecast is then issued under the assumption of the persistence of the average daily $T_{LI}$. For days when no clear-sky observations exist, the most recent historical clear-sky observation is used to calculate the forecast allowing for the investigation of errors associated with lagged turbidity data. Because forecasting models are often tuned for specific microclimates, the distribution of parameters such as $T_{LI}$ provides valuable insight into which methods should be employed for such climates. For example, climates with a broader distribution of turbidity values would perform worse under the persistence assumption. As a result, histograms of the clear-sky $T_{LI}$ are also provided.

Finally, the performance of the day ahead forecast is evaluated as a function of zenith angle using two metrics described in the next section in conjunction with the clear-sky observations of DNI at the seven stations.

6. Metrics

In order to evaluate the performance of the CSMs both the Root Mean Square Error (RMSE) and Mean Bias Error (MBE) as a function of solar zenith angle will be calculated. However, in many contexts relative error is more commonly desired than absolute error, especially for users in the utility industry (Inman et al., 2013). In accordance with Hoff et al. (2012), simplified reporting approaches for relative (percent) versions of RMSE and MBE are used in this work. The rRMSE is normalized with the average value of the observed clear-sky irradiance at that zenith angle

$$rRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( \frac{\hat{I}(t) - I(t)}{I(t)} \right)^2}.$$  \hfill (19)

Similarly, the rMBE, which has the advantage of being independent of the number of observations as well as being easy to understand, is also obtained through a normalization of the MBE by the average value of the irradiance:

$$rMBE = \frac{1}{N} \frac{\sum_{i=1}^{N} (\hat{I}(t) - I(t))}{\sum_{i=1}^{N} I(t)} = \left( \frac{1}{\sum_{i=1}^{N} I(t)} \right) \sum_{i=1}^{N} (\hat{I}(t) - I(t)).$$  \hfill (20)

It is important to pause here and note that as a result of the metrics being calculated as a function of zenith angle, the

---

**Fig. 6.** Illustration of methodology: (a) for each month in the data set, turbidity information for the location is looked up from the monthly SoDa world maps of Linke turbidity (Remund and Page, 2002). This monthly turbidity approximation is then used as an input to calculate the DNI CSM for the entire month using Eq. (8), assuming that $T_{LI} = T_{m}$. (b) The airmass independent Linke turbidity factor $T_{LI}$ is the calculated from DNI measurements, which is equal to Kasten’s Linke turbidity factor $T_{LK} @ m = 2$. (c) After the initial CSM is calculated, its general shape is used to perform binary clear-sky detection as described in Section 4. (d) For each of the partially clear days an average $T_{LI}$ is calculated which is then used to as an input to Eq. (8) to estimate the CSM.
average value in (19) and (20) correspond to the average value at a given zenith angle, rather than the average value over the entire time series. Furthermore, as a point of clarification, observations were separated in bins of equal zenith angle rounded to the nearest degree, prior to the calculation of the metrics at each zenith angle. While this comes with an additional computational expense, the impact is insignificant compared to the implementation of the method itself.

7. Results

As mentioned in Section 5, the calculated average daily $T_{LI}$ for days with clear-sky observations, the monthly SoDa $T_{LI}$, and the fraction of each day that was classified as clear using the method described in Section 4 are shown in Fig. 4. Fig. 4 illustrates that, with the exception of Bellingham and La Jolla, the calculated average daily $T_{LI}$ for all stations was typically 15–30% less than the monthly SoDa values for $T_{LI}$ indicating that the SoDa database for $T_{LI}$ tends to underestimate DNI for the particular CSM employed in this work. It is also clear from Fig. 4 that Folsom, Davis, Merced and Berkeley (all located within ~160 km from each other) were characterized by a large fraction of clear days. The remaining sites exhibited significantly fewer clear-sky observations with tropical Ewa Beach, HI, showing the fewest followed by Bellingham,

Table 2
Table listing the total number of days that were included in the study, the total observations that met the closure equation quality control, percentage of the observations that were classified as clear, the average airmass independent Linke turbidity factor $T_{LI}$, and the standard deviation of the distribution of $T_{LI}$ for each site.

<table>
<thead>
<tr>
<th>Site</th>
<th>Days</th>
<th>Total obs.</th>
<th>% Clear</th>
<th>$T_{LI}$</th>
<th>Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oahu</td>
<td>397</td>
<td>215,746</td>
<td>9.78</td>
<td>3.34</td>
<td>0.37</td>
</tr>
<tr>
<td>Bell</td>
<td>760</td>
<td>366,050</td>
<td>19.51</td>
<td>3.43</td>
<td>1.02</td>
</tr>
<tr>
<td>La Jolla</td>
<td>820</td>
<td>629,576</td>
<td>28.02</td>
<td>3.09</td>
<td>0.82</td>
</tr>
<tr>
<td>Berkeley</td>
<td>701</td>
<td>353,210</td>
<td>44.19</td>
<td>2.90</td>
<td>0.62</td>
</tr>
<tr>
<td>Davis</td>
<td>670</td>
<td>366,391</td>
<td>44.60</td>
<td>2.81</td>
<td>0.54</td>
</tr>
<tr>
<td>Merced</td>
<td>608</td>
<td>596,327</td>
<td>46.15</td>
<td>2.94</td>
<td>0.88</td>
</tr>
<tr>
<td>Folsom</td>
<td>516</td>
<td>262,896</td>
<td>52.59</td>
<td>2.79</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Fig. 7. Comparison of clear-sky rRMSE and rMBE as a function of zenith angle for the monthly SoDa turbidity maps, the experimental method and persistence forecast described in Section 5. The extensive decrease of both rRMSE and rMBE of the improved DNI CSM over a wide range of zenith angles suggest that increased accuracy is achievable with minimal resources.
which is consistently overcast, and coastal La Jolla, with its volatile marine layer. Table 2 lists the total number of days that were included in the study, the total observations that met the closure equation quality control, percentage of the observations that were classified as clear, the average air-mass independent Linke turbidity factor, and the standard deviation of the $T_{LI}$ distribution for each site.

The relative (or percent) versions of the RMSE and MBE described in Section 6 were also calculated as a function of solar zenith angle in this work. Again, we compare the CSM obtained from the monthly SoDa maps of $T_{LI}$ with the experimental method and persistence forecast described in Section 5. The statistics are based on days where at least 10% of the observations are clear in order to ensure an accurate sampling of the attenuation caused by the atmospheric column. Overall, this set of validation metrics is presented in Fig. 7. The rRMSE for the method applied to the same day is typically less than 5% and can be as low as 0.5% for zenith angles less than 30°; with the exception of Bellingham, WA whose rRMSE exceeds 5% for zenith angles greater than 60°. The $rRMSE_{Pers}$ of the forecasted persistence $T_{LI}$ CSM had a similar shape to the rRMSE with an additional 1–3% error while the monthly SoDa $T_{LI}$ produced an $rRMSE_{SoDa}$ of a comparable shape with an additional 3–5% error over the $rRMSE_{Pers}$.

Fig. 7 also shows that the rMBE of the monthly SoDa $T_{LI}$ is consistently negative for all zenith angles, which is also in agreement with the results presented in Fig. 4 and again suggests that the SoDa input data in conjunction with the CSM employed in this work consistently underestimates DNI through an overestimation of $T_{LI}$. It should be noted that the $rMBE_{Pers}$ was similar to the rMBE of the improved CSM for all of the sites, including Bellingham, WA. The relatively small (~1–2%) and positive rMBE of the improved and forecasted CSMs also indicate a slight over estimation of DNI rather than the notable underestimation provided by the monthly SoDa $T_{LI}$ data, which is typically desirable of CSMs.

Finally, Fig. (8) shows the histograms of the average daily $T_{LI}$ for each of the seven stations. It is clear that, with the exception of Bellingham and La Jolla, a general trend is apparent in the distribution of $T_{LI}$. Most of the sites exhibit an almost normal distribution centered near $T_{LI} = 3$. Bellingham, however, has a distribution which is shifted to the right and centered near 3.75. This value is characteristic of its marine oceanic climate, which is strongly influenced by the Cascade Range to the east that retains marine temperature influences and the Olympic Mountains to the southwest that provide a strong-rain shadow effect. On the other hand, La Jolla’s distribution is significantly broader with no apparent peak. This is characteristic of the wind dominated weather pattern in La Jolla. Typically, eastwardly winds are cooled as they pass over the Pacific Ocean and advect with them a thick marine layer of fog which dissipates near mid-day. Occasionally, however, the westward Santa Ana winds will
increase aerosol loading as they pass over the deserts to the east and push temperatures to nearly 40 °C, well above the average annual high of approximately 20°. These marine layer climatologies may explain some of the decreased performance at Bellingham apparent in Fig. 7.

Finally, a general trend is apparent in Fig. 7 which should be taken into account. For all of the stations included in this study errors increase in magnitude with the zenith angle suggesting that the DNI CSM and consequently the definition of $T_{LI}$ still has some dependence on airmass.

8. Conclusions

This study focus on the sensitivity of day-ahead DNI forecasts under clear skies to local fluctuations in turbidity. An endogenous clear-sky detection algorithm for DNI was developed based on the recent work of Reno et al. (2012), which was then used to determine average daily air mass independent Linke turbidity $T_{LI}$ information. Average daily $T_{LI}$ factors were then used to correct for temporally and spatially local aerosol loading and water vapor content. When compared to monthly climatologies, the rRMSE and rMBE of CSMs that used daily climatologies were lower by approximately one order of magnitude. This suggests that daily turbidity information should be used for DNI clear-sky modeling, which is in agreement with the findings of Remund et al. (2003).

While remote sensing techniques are available for turbidity calculation (MODIS aboard NASA’s AQUA/TERRA) the strict sampling rate of PLEOs may not be sufficient for accurate DNI clear-sky modeling, see Appendix A. It should be noted that this issue may be solved by new generation instrumentation on future geosynchronous satellite campaigns; e.g. GOES-R scheduled for launch in 2016. In this work a persistence forecast, which employed lagged turbidity information, resulted in an rRMSE which increased by a factor of 2 with respect to daily climatologies. The rMBE remained unchanged which is a result of the persistency turbidity time-series being unbiased in its error distribution. As a result, it is recommended that CSP plants use existing broadband pyrheliometers and endogenous cloud filtering techniques in order to correct for local turbidity fluctuations.

Error analysis of the ground bases sensing CSM as a function of zenith angle suggests that the rRMSE is typically bounded by 5% ($50 \text{ W m}^{-2}$) and can be as low as 0.5% ($5 \text{ W m}^{-2}$) for solar zenith angles less than 30°. In addition, the positive and notably small magnitude (1–3%) of the rMBE of the improved model suggests that the algorithm only slightly overestimates DNI which is typically desirable for CSMs which should provide an outer envelope of available irradiance and can also be used in the detection of cloud enhancement of irradiance (Tapakis and Charalambides, 2014). The substantial reduction of both rRMSE and rMBE of the improved DNI CSM over a wide range of solar zenith angles indicates that increased accuracy is achievable with minimal resources. In addition the improvement appeared to be very weakly dependent on location, as similar improvements were observed in nearly all of the micro-climates. The proposed approach is therefore simple to implement, computationally inexpensive, and geographically robust.

Although the CSM employed in this work is based on an inversion of the definition of $T_{LI}$, the algorithm is applicable to any turbidity-based CSM. They key to the improvement is the observation of the inverted parameter under clear-skies, which is then assumed to persist.

Acknowledgements

The authors gratefully acknowledge funding from the California Energy Commission (CEC) Project EPC-14-008 and DOE SUNRISE project number 0865-1517. Support from Clean Power Research for the SolarAnywhere database is also gratefully acknowledged.

Appendix A. While it is true that remote sensing techniques (e.g. MODIS onboard NASA’s AQUA/TERRA) provide global maps of aerosol optical depth, this does not invalidate the determination of Linke turbidity from ground-based DNI observations. Due to the PLOE of the AQUA and TERRA satellites they only pass over a given location up to four times a day. Typically half of these occur during the night, leaving only two passes per day for a given location. However, during partly cloudy days there is no guarantee that the satellite will pass over the location during a clear portion of the day. For that reason, a case study was performed in order to investigate the probability of a satellite missing an opportunity to sample a clear-sky portion of the day. In other words, the question of how likely was it that the circum-solar region was clear (from the point of view of a ground-based radiometer) before or after a satellite passed over a location and observed cloud cover on the same day. It is important to note here that the extent of the cloud cover is not known. While cloud contamination exist (to some extent) in all of the satellites ‘missed’ passes, final AOD retrievals from MODIS are typically spatially relaxed ($\sim 1° \times 1°$), which may solve the issue in cases when the cloud field had a limited horizontal extent.

In order to provide a lower bound on the probability that a satellite observation would contain cloud cover on a day that was partly clear, data from Folsom was used as a result of the high fraction of clear-sky observations. Information regarding the satellites orbit was used to calculate the times when a satellite passed over Folsom and the endogenous clear-sky detection algorithm was used to determine if the entire pass was clear. The statistics for the AQUA and TERRA satellites from May 2013 to May 2014 are presented in Table A.3. In addition the histogram of the number of clear-sky observations for both the ground-based sensing and remote sensing are shown
in Fig. A.9. It is clear from Table A.3 that the AQUA and TERRA satellites miss the opportunity to sample a clear window in a day nearly 50% of the time that the day was partly cloudy or about 25% of all of its passes. In addition, the number of clear-sky observations the ground-based methodology recorded was larger than the remote sensing techniques by three orders of magnitude, suggesting a much more robust sampling of atmospheric turbidity.

References


