Solar Energy 135 (2016) 854–863

Contents lists available at ScienceDirect

Solar Energy

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

Improving the accuracy of hourly satellite-derived solar irradiance by combining with dynamically downscaled estimates using generalised additive models

Robert J. Davy *, Jing R. Huang, Alberto Troccoli

CSIRO Oceans and Atmosphere, GPO Box 3023, Canberra, ACT 2601, Australia

ARTICLE INFO

Article history:
Received 10 September 2015
Received in revised form 14 June 2016
Accepted 20 June 2016
Available online 1 July 2016

Keywords:
Clear sky index
DNI
GHI
Insolation

ABSTRACT

The gridded hourly solar irradiance derived from satellite imagery by the Australian Bureau of Meteorology represents the current state of the art in quantification of the long-term solar resource for locations where no ground measurements are available. Using nonparametric regression, we test the potential for the satellite-derived global horizontal irradiance and direct normal irradiance to be improved by combining with irradiance that has been dynamically downscaled using a numerical weather prediction (NWP) model. NWP irradiance, together with the satellite irradiance, solar zenith angle and their interaction terms are used as inputs to generalised additive models (GAM) using smoothing splines. The spatial weighting of these empirical models according to distance is also tested. Cross validation with ground measurements indicates that RMSE can be improved by a few percent over the satellite-derived irradiance. The addition of dynamically downscaled irradiance as a GAM predictor further improves RMSE by a few percent, depending on location. When these empirical models are weighted spatially over large distances (hundreds of kilometres), the results are more equivocal. However, spatial weighting of the regression functions should be possible over smaller regions where the atmospheric turbidity properties are similar.

1. Introduction

Presently, there are no very large solar energy generators in Australia; however, rooftop photovoltaic power now exceeds 4500 MW capacity (Clean Energy Regulator, 2015), a 102 MW photovoltaic solar farm was commissioned in 2015 at Nyngan, and a 2 GW plant at Bulli Creek has received planning approval. Solar farms require accurate resource estimation and quantified uncertainty at the project feasibility stage. It is important to estimate the resource accurately while quantifying and minimising the uncertainty.

For solar resource assessment, two main sources of data have traditionally been used: ground measurements and satellite-derived modelling. More recently, it has been found that reanalysis-based irradiance estimates can be useful when the satellite irradiance is not available (Bojanowski et al., 2014). This suggests that there is opportunity to combine all three data products in a way that is better than either the satellite or reanalysis on its own. This is particularly important for Australia, where there is a heavy reliance on the satellite irradiance to provide estimates over vast amounts of terrain where there are no nearby ground stations.

The satellite-derived irradiance has errors that depend on clear sky index and solar zenith angle (Badescu and Dumitrescu, 2013), as well as other variables such as aerosols, cloud properties and satellite viewing angle. Corrections can be performed if ground measurements are available, assuming ground stations represent the truth. Ground stations are the most reliable source of solar irradiance measurements, but these are sparsely located. Successful attempts have been made at fusion of satellite model and ground measurements (Beyer et al., 1997; Zelenka et al., 1992), but this has generally used daily mean irradiance and a relatively dense ground station network, enabling such techniques as co-kriging. Recently, Ruiz-Arias et al. (2015) used an optimal interpolation approach for combining numerical weather model estimates of irradiance with ground measurements for monthly averaged data. These methods rely on a knowledge of the spatial covariance of the errors in the data. In Australia, the sparse monitoring network makes estimating the spatial covariance highly problematic. This suggests that a more empirical approach may be more appropriate. Globally, there is also a trend towards energy assessment using

* Corresponding author.
E-mail address: robert.davy@csiro.au (R.J. Davy).

http://dx.doi.org/10.1016/j.solener.2016.06.052
0038-092X/

time series with higher temporal frequency (hours or minutes), partly due to the nonlinear behaviour of concentrating solar thermal plants (Gueymard and Ruiz-Arias, 2015). Therefore, it is desirable to work with this higher frequency information where possible.

In an Australian context, Blanksby et al. (2013) performed statistical corrections on the Bureau of Meteorology (BoM) hourly satellite solar irradiance product. They used linear models of the satellite irradiance variables, along with interaction and squared terms. Although such models proved to be adequate for the purpose of correcting bias, the full potential of nonlinear and nonparametric models was not explored. Also, the models were apparently not validated on independent data. An interesting feature of that work was to include direct normal irradiance (DNI) estimates as predictors to correct the global horizontal irradiance (GHI) estimates. Work has also been done to improve pre-satellite era estimates of daily irradiance for Australia using sunshine duration and manual cloud cover observations (Zajaczkowski et al., 2013). Although accurate enough for agriculture and climate studies, this procedure is unlikely to be useful for quantifying energy solar resources.

Some random error in gridded satellite irradiance is inevitable due to its spatially averaged nature, in contrast with ground stations (Zelenka et al., 1999). Regarding systematic errors, polynomial regression models have been used for correcting the bias of the satellite irradiance as a function of satellite clear sky index (the ratio of the irradiance to the clear sky value) and cosine of the solar zenith angle. Fourth-order models have been reported in the literature in a forecasting context (Lorenz et al., 2009, 2011). Third-order polynomials have been used by Mieslinger et al. (2014) for hourly data. Although polynomial regression models are simple to implement, there remains scope to improve the accuracy through adoption of other kinds of regression such as generalised additive models (GAM) using cubic smoothing splines (Venables and Ripley, 2002).

The objectives of this work are to investigate and improve the accuracy of hourly satellite-derived irradiance estimates. This is done in four main steps: (i) investigate the use of GAM as an alternative to polynomial regression; (ii) investigate the value of including irradiance derived from a reanalysis weather model as an additional predictor for the hourly measured solar irradiance; (iii) develop an estimate for the error variance as a function of the two irradiance sources and the zenith angle, and thus investigate the extent to which the weather model irradiance contributes knowledge regarding the uncertainty in the combined irradiance estimate; (iv) explore spatial weighting of the regression functions, to the extent that the ground monitoring network allows validation.

2. Methods and data

2.1. Ground stations

The BoM maintains a solar radiation monitoring network – since 1999, 18 stations have been in operation at some time on the Australian mainland (Fig. 1). The number of stations fluctuates according to budget priorities and 11 of these stations had already ceased operation as of 2014. Instrumentation consists of thermopile pyranometers (Kipp & Zonen CM11) and pyrheliometers (Kipp & Zonen CH1 or Middleton DN5) that are regularly calibrated and serviced. Measurement averages and other summary statistics derived from the 1 second readings are logged each minute. The stations are quite comprehensive in that they measure direct beam and diffuse irradiance as well, employing a mechanical solar tracker (Bureau of Meteorology, 2012). Table 1 shows the station locations and their approximate run dates, noting that 2013 was the final year considered for this study.

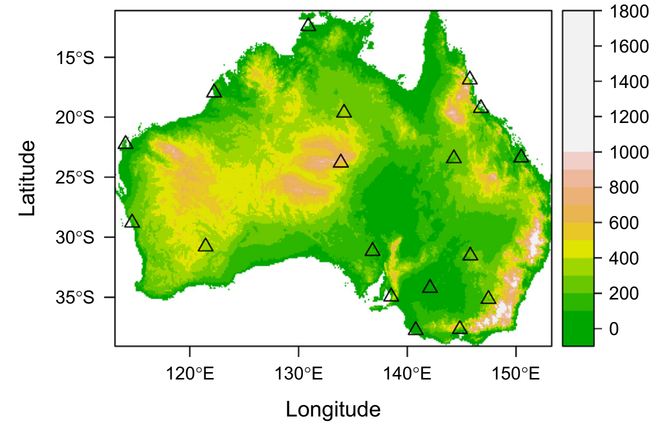

Fig. 1. BoM ground stations and topography contours on the Australian mainland. Legend denotes height in metres above sea level.

Table 1
Names and locations of the ground stations used in this study. Note Mildura was not operational from 2006 to 2011.

Location	Latitude	Longitude	Start	End
Broome Airport	−17.95	122.23	1996	2013
Learmonth Airport	−22.23	114.08	1996	2006
Geraldton Airport Comparison	−28.80	114.70	1996	2006
Kalgoorlie-Boulder Airport	−30.78	121.45	1998	2006
Darwin Airport	−12.42	130.88	1993	2013
Tennant Creek Airport	−19.63	134.18	1996	2006
Alice Springs Airport	−23.80	133.88	1993	2013
Woomera Airport	−31.16	136.81	2012	2013
Adelaide Airport	−34.95	138.52	1994	2013
Mount Gambier Airport	−37.73	140.78	1993	2006
Cairns Airport	−16.88	145.75	1997	2004
Townsville Airport	−19.25	146.77	2012	2013
Longreach Airport	−23.44	144.28	2012	2013
Rockhampton Airport	−23.38	150.48	1996	2013
Cobar Airport	−31.54	145.80	2012	2013
Wagga Wagga AMO	−35.17	147.45	1997	2013
Mildura Airport	−34.23	142.08	1996	2013
Melbourne Airport	−37.67	144.85	1999	2013

2.2. Satellite data

The satellite irradiance product from the BoM forms the basis of our analysis. The product represents hourly GHI energy density in MJ m^{-2}gridded at 0.05° spatial resolution, derived from cloud imagery of resolution 0.01° (Bureau of Meteorology, 2009). The record begins in 1990. The original satellite imagery is from the Geostationary Meteorological Satellite series operated by Japan Meteorological Agency and from GOES-9 operated by the National Oceanic & Atmospheric Administration (NOAA) for the Japan Meteorological Agency.

Instantaneous hourly irradiance values are derived from cloud imagery using the physical model of Weymouth and Le Marshall (2001), which uses the visible and near-infrared spectral bands. The hourly irradiance is then integrated over the day to give daily insolation. The accuracy of this procedure is tabulated extensively in that publication. Although the hourly deviations are not reported, some assessment of the hourly error statistics is presented in Blanksby et al. (2013).

In this work we have disregarded the satellite data during the period 2001-07-01 to 2003-06-30. Visual inspection shows that the data seems inconsistent with that either before or after this period.