Solar Energy 144 (2017) 529–539

Contents lists available at ScienceDirect

Solar Energy

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

Assessing the value of simulated regional weather variability in solar forecasting using numerical weather prediction

Jing Huang [a,*], Marcus Thatcher [b]

[a] CSIRO Oceans and Atmosphere Flagship, Yarralumla, ACT, Australia
[b] CSIRO Oceans and Atmosphere Flagship, Aspendale, Victoria, Australia

ARTICLE INFO

Article history:
Received 17 October 2016
Received in revised form 3 January 2017
Accepted 25 January 2017

Keywords:
Solar forecasting
Numerical weather prediction
Global horizontal irradiance
Post-processing
Bias correction

ABSTRACT

Numerical weather prediction (NWP) is currently the best tool to forecast solar radiation beyond several hours ahead. However, mainly due to the stochastic nature of clouds, spatial resolution used by NWP significantly affects the forecasting accuracy of solar irradiance and power. In this study, the effects of the simulated regional weather variability at a relatively fine spatial resolution on the forecasting accuracy of solar irradiance are systematically investigated using the Conformal Cubic Atmospheric Model (CCAM) and the Global Forecast System (GFS). Nudging from the US National Centers for Environmental Prediction (NCEP) global analysis, CCAM has been run to forecast solar radiation at a resolution of 4 km in horizontal space covering the whole Australia. For the prediction of Global Horizontal Irradiance (GHI), we find that the high-resolution CCAM generally produces more accurate forecasts than the low-resolution GFS for all nine observation stations we investigate, when using the nearest grid point approximation in combination with bias correction. Spatial averaging to a certain scale is able to enhance the performance of both NWP models in solar forecasting as measured by mean errors. However, spatial averaging, which is similar to a low resolution used in NWP models, tends to significantly and unrealistically reduce the extent of solar variability. The optimal scale of spatial averaging, when determined by the minimum of Mean Absolute Error (MAE), relies on the climatic characteristics of the location and ranges from about 100 km to about 400 km for the nine stations.

1. Introduction

Forecasting solar irradiance is receiving more and more attention from solar research community as the penetration of solar power into the electric grid increases rapidly around the world. Numerical Weather Prediction (NWP) is currently the best tool to forecast solar irradiance beyond several hours ahead. Although NWP models capture some aspects of the large scale clouds, NWP models do not accurately predict the stochastic nature of clouds at small spatial and temporal scales. Coupled with other factors, this leads to randomness to some degree in the forecast of solar irradiance near the Earth's surface.

Recently, it has been an important topic to examine and compare the performance of NWP models in forecasting solar irradiance (see e.g. Larson, 2013; Perez et al., 2013; Lorenz et al., 2016). The European Centre for Medium-Range Weather Forecasts (ECMWF) was used to predict regional solar photovoltaic (PV) power output up to three days ahead in Germany in Lorenz et al. (2009). The Global Horizontal Irradiance (GHI) forecasts produced by the North American Model (NAM), Global Forecast System (GFS), and ECMWF were validated for the continental United States (US) against ground measurement data in Mathiesen and Kleissl (2011). Despite its relatively coarse resolution, the bias-corrected GFS was found to provide the most accurate solar forecasts for the continental US followed by ECMWF and then NAM. In addition, the performance of a number of prevailing NWP models was compared by validating their irradiance forecasts for multiple sites in North America and Europe in Perez et al. (2013). It was shown that global NWP models (e.g. ECMWF) tend to forecast GHI more accurately than mesoscale models (e.g. the Weather Research and Forecasting (WRF) model) as evidenced by the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE). This was attributed to the shortcomings in the physical schemes of the relevant mesoscale model as well as the need of handling lateral boundary conditions at initialisation. A recent work under the International Energy Agency (IEA) Solar Heating and Cooling (SHC) Task 46 "Solar Resource Assessment and Forecasting" evaluated GHI forecasts for multiple sites in three European countries based on a variety

* Corresponding author at: CSIRO Oceans and Atmosphere Flagship, Wilf Crane Crescent, Yarralumla, ACT 2601, Australia.

E-mail address: jing.duke@gmail.com (J. Huang).

http://dx.doi.org/10.1016/j.solener.2017.01.058
0038-092X/

of NWP models and two Model Output Statistics (MOS) systems (Lorenz et al., 2016). The effects of spatial and temporal averaging were particularly focused and it was found that spatial and temporal averaging generally decreased the RMSE of irradiance forecasts. Higher resolution in mesoscale NWP models tended to forecast more variability in solar irradiance and cloud cover, which was in better agreement with measurements. However, temporal correlation of variability was found to be low between even the best NWP model used in this work and measurements.

In Australia, the Bureau of Meteorology (BoM) tested the ability of its NWP model to predict solar irradiance (Gregory et al., 2012). While the forecasts are generally in good agreement with satellite-derived values across Australia, differences in solar irradiance forecasts were attributed to incorrect representation of clouds in the tropical and mountainous areas. A following work further validated the performance of the recently upgraded system, the Australian Community Climate Earth-System Simulator (ACCESS) model, in forecasting global, direct and diffuse solar irradiance over Australia (Gregory and Rikus, 2016). While the improvement in global forecasts over the previous operational model was confirmed, it was demonstrated that forecasts of diffuse and direct solar irradiance still suffer from large biases. Furthermore, two versions of ECMWF were assessed in terms of their ability in forecasting direct and global irradiance for Australia (Troccoli and Morcrette, 2014). While the relative MAE for direct irradiance was found to be significantly larger than that for global irradiance, there was a marked dependency on cloudiness conditions as well as the background climatic characteristics of the associated location.

NWP models can be run at various spatial scales depending on the complexity of the model and the computational resource available. However, little is known regarding to which degree the simulated regional variability contributes to the performance of solar forecasting. In this paper, we systematically assess the value of regional variability in forecasting solar irradiance by analysing the output of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Conformal Cubic Atmospheric Model (CCAM) and GFS. The observation data for 9 sites across Australia in 2012 was used to validate the performance of the two NWP models. CCAM was initialised by the same analysis data as GFS. However, the output resolution of GFS is about 50 km for the entire globe while CCAM was run at 4 km across Australia, a resolution beginning to resolve clouds (Larson, 2013). The CCAM-GFS comparison demonstrates that it is possible for a model which mainly focuses on resolving regional activities to improve the prediction of GHI in a statistical sense, compared to a global NWP model. We believe this is a result of CCAM dynamically simulating the regional variability, which does not exist in the coarse output of GFS.

Another way we assess the value of regional variability in solar forecasting is to compare the raw NWP output and the output that has been spatially-averaged (e.g. Pelland et al., 2013; Lorenz et al., 2016). This helps to clarify the spatial scale of any forecast improvements by the regional simulation. To be more specific, we refer to regional variability as weather patterns described by the physical states at the spatial scales between the coarse resolution and the fine resolution, such as wind speed, temperature and humidity, etc. In addition, the role of the bias-correction procedure in the forecasting accuracy of solar irradiance is also investigated for the two NWP models, respectively. While the accuracy of solar forecasting has mostly been evaluated by MAE and RMSE of GHI (e.g. Mathiesen and Kleissl, 2011; Gregory et al., 2012; Troccoli and Morcrette, 2014), we intend to analyse more advanced statistical properties such as the distribution structure of GHI and temporal solar variability as well in this study.

The rest of the paper is structured as follows: Section 2 introduces the two forecast models and the observation data used. Section 3 explains the post-processing techniques used for the NWP models as well as the error metrics of evaluation. The results on the forecasting accuracy of solar irradiance are presented in Section 4. Section 5 then demonstrates the existence of an optimal scale for spatial averaging and how it varies with location. Section 6 further examines the effects of spatial averaging and bias correction on the forecasting accuracy of solar variability. The paper is summarised and concluded in Section 7.

2. Forecast models and observation data

2.1. Global Forecast System (GFS)

GFS is a global NWP model developed and run by the US National Oceanic and Atmospheric Administration (NOAA) through the National Centers for Environmental Prediction (NCEP). Its products are available online at http://nomads.ncep.noaa.gov. For the year of 2012, GFS forecasts have a spatial resolution of $50\ \mathrm{km} \times 50\ \mathrm{km}$ for the global domain. Temporally, the resolution is 3 h up to 180 h ahead, of which we use only the first 8 intraday forecasts, i.e., 3–24 h ahead. Note that the GHI values from GFS represent a 3-hourly temporal average. In Section 3.1 we introduce how the GFS output is temporally interpolated to match the interval of the CCAM output. Although the forecasting results are issued four times a day, viz., 00, 06, 12 and 18 UTC, only those at 18 UTC are used in our study. Thus, we have a continuous time series at each point on the GFS grid. Unfortunately, there are 15 days of data unavailable out of 366 days in 2012.

Among the models used to parameterise atmospheric physical processes, the radiation and the cloud microphysics schemes will affect the forecasting of solar irradiance most directly. In the case of GFS, the radiation is parameterised with the rapid radiative transfer model (Clough et al., 2005), while in CCAM the radiation model is the GFDL exponential-sum-fit and simplified-exchange-approximation by Freidenreich and Ramaswamy (1999) and Schwarzkopf and Ramaswamy (1999), respectively. The shortwave radiation spectra are divided into a comparable number of bands for both models with 14 and 18 bands for GFS and CCAM, respectively. For longwave, both the GFS and CCAM models use 16 bands. Despite differences in the cloud microphysics parameterisations for GFS (Sundqvist et al., 1989; Zhao and Carr, 1997; Moorthi et al., 2001) and CCAM (Rotstayn, 1997), both models support prognostic cloud condensate for water and ice and both models diagnose cloud cover as well as allow partial cloud within a grid-box. Both models also use a max-random approach to represent cloud overlap. In the case of CCAM, it takes 3–4 h to fully spin-up the cloud behaviour in the model.

2.2. Conformal Cubic Atmospheric Model (CCAM)

Like GFS, CCAM is also a global forecasting model using a semi-Lagrangian dynamical core (e.g. McGregor, 2005; McGregor and Dix, 2008; Thatcher and McGregor, 2011). However, CCAM supports a variable-resolution global grid through the use of Schmidt transformation (Schmidt, 1977). Therefore, it is able to focus on a target area with a fine grid spacing while avoiding the need for any special treatment of simulation boundaries. As illustrated in Fig. 1, the CCAM grid used in our study centres on the Australian continent. Although the grid covers the whole globe, the grid spacing increases as the distance from Australia increases. For Australia, the resolution is approximately $4\ \mathrm{km} \times 4\ \mathrm{km}$. CCAM was initialised by the GFS analysis data at 18 UTC each day and was run for 2012 with a forecasting horizon of 24 h. The output interval is half an hour. Similar to GFS, the GHI forecasts of CCAM represent temporal average over each output interval.