Solar Energy 162 (2018) 454–471

Contents lists available at ScienceDirect

Solar Energy

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

Assimilating observations to simulate marine layer stratocumulus for solar forecasting

Dipak K. Sahu*, Handa Yang, Jan Kleissl

Center for Renewable Resource Integration, Center for Energy Research, University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA

ARTICLE INFO

Keywords:
Solar energy forecasting
Marine layer stratocumulus cloud
NWP
Intra-day forecast
Data assimilation

ABSTRACT

Integration of solar energy forecasts into the electric network is becoming essential because of the continually increasing penetration level of solar energy. Three-dimensional numerical weather prediction (NWP) models predict the weather based on the current weather conditions (called initialization) and simulate the ensuing atmospheric processes. The accuracy of forecasts therefore depends, in part, on the accuracy of the model initializations. Data assimilation is recognized as the most widely used technique to improve the initialization into NWP models. In this study, meteorological observations from the surface and upper-air in-situ networks over the southern California coast are assimilated into the advanced research version of the Weather Research and Forecasting (WRF) model using a three dimensional variational data assimilation technique (3DVAR). A single observation test was conducted to tune-up the length scale and variance scale along with the regional domain-dependent background error statistics. A customized version of 3DVAR data assimilation was deployed with two sets of cyclic data assimilation with 6-h and 1-h assimilation windows along with the cold-start mode. The cyclic data assimilation experiments consistently outperformed the cold-start data assimilation and WRF for intra-day Global Horizontal Irradiance (GHI) and Clear Sky Index (CSI) forecast. Hourly cyclic assimilation showed the highest forecast skill score against ground measurements and satellite measurements. Even at the coastal stations with more challenging meteorological conditions, the hourly cyclic assimilation consistently outperformed the 24-h persistence forecast. The average (mean of four case studies) hourly cyclic data assimilation showed the highest forecast skill score in GHI and CSI intra-day forecast with reference to 24-h persistence forecast up to 39.4% and 40.7% respectively at the coastal stations. The spatial distributions of GHI biases estimated against SolarAnywhere satellite measurements showed that the hourly cyclic assimilation consistently improved the stratocumulus cloud coverage, thickness, and life time over the coastal region, but biases are still present further inland.

1. Introduction

The increasing contribution of solar energy to electric power generation motivates accurate forecasts of solar irradiance at the earth's surface. Satellite-based solar irradiance forecasts, sky-imager based solar irradiance forecasts, and statistical forecasts are considered to be the most accurate solar irradiance forecasts for horizons up to 6 h. However, these forecast methods become inferior beyond 6 h ahead. Generally, grid operators aspire to receive forecasts of solar energy for at least a day ahead. State-of-the-art numerical weather prediction (NWP) models can provide improved solar irradiance forecasts for a day ahead period.

Solar irradiance depends on both geographical location and meteorological conditions. The major concern in solar power generation is high variability in the amount of incident solar irradiance due to the presence of clouds. Accurate cloud forecasting within a NWP model is challenging due to the complex multiscale physical processes that influence cloud formation. Forecasting the low clouds in the marine layer over coastal southern California is especially challenging. Marine layer clouds are typically shallow and exist near the inversion of the planetary boundary layer (PBL). Recently, several studies have been conducted to understand the physical processes associated with the formation of marine layer clouds and their representation within NWP models (Ghonima et al., 2016, 2017; Kazil et al., 2016; Sandu et al., 2009; Hu et al., 2010; Jankov et al., 2011).

Most operational meteorological centers use NWP models to forecast solar irradiance, although they are known to over predict solar irradiance. Mathiesen and Kleissl (2011) found the mean bias error (MBE) and root mean square error (RMSE) of hourly averaged surface irradiance forecast exceeded $30\,W\,m^{-2}$ and $110\,W\,m^{-2}$, respectively,

* Corresponding author at: Department of Earth, Ocean, and Atmospheric Science, Florida State University, 1017 Academic Way, 430 LOVE Building, Tallahassee, FL 32306, USA.
E-mail address: dsahu@fsu.edu (D.K. Sahu).

https://doi.org/10.1016/j.solener.2018.01.006
Received 6 May 2017; Received in revised form 30 December 2017; Accepted 4 January 2018

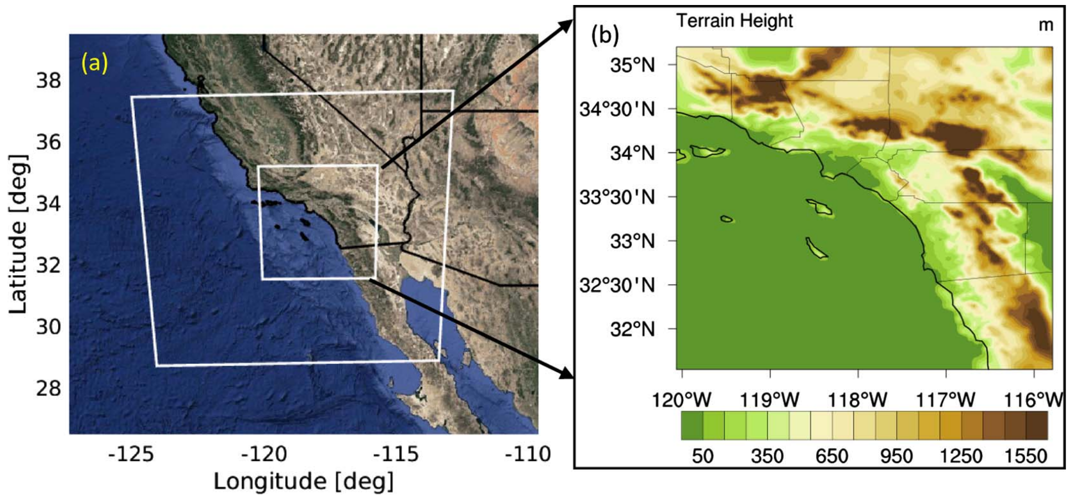

Fig. 1. (a) Map of the two-nested domains used in the numerical model and (b) The terrain height (m) of the inner domain centered over southern California. Outer and inner domains are at 8.1 km (131 × 121 × 75 grid points) and 2.7 km (151 × 151 × 75 grid points) horizontal resolutions, respectively. Satellite image ©2015 Google.

over southern California. Similarly, the multi-model inter-comparison study conducted by Perez et al. (2013) found noticeably positive MBE over the United States. A similar study conducted by Lara-Fanego et al. (2012) over southern Spain found for forecasts with a 24-h lead time that the MBE of forecasted global horizontal irradiance (GHI) was 2% for clear sky conditions and 18% for cloudy conditions. The RMSE of forecasted GHI ranged from 10% under clear sky to 50% for cloudy conditions.

Three-dimensional NWP models predict future weather conditions based on the current weather conditions (called initial conditions) and physical parameterizations to simulate ensuing atmospheric processes. The accuracy of forecasts therefore depends, in part, on the accuracy of the model initializations (Kalnay, 2003). Initialization describes the process of determining the initial state of the atmosphere in terms of various meteorological parameters. Data assimilation is recognized as the most widely used technique to improve the initialization into NWP models.

In the past few decades, the advancement of computational resources, novel data assimilation algorithms and new observation networks including remote sensing and in-situ observations has continually improved the model initialization (or analysis) and subsequent forecasts. Despite the advancements in data assimilation and computational facilities, the use of in-situ observations, especially surface observations, remains a challenging problem. In addition, due to heterogeneous topography, considerable difficulties have been encountered when assimilating surface observations over complex terrain. Currently, only a limited number of surface observations are used in NWP. In the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) 50-year reanalysis (Kalnay et al., 1996), only surface pressure observations were assimilated.

Since NWP models still have significant biases in GHI forecast over different parts of the globe, and forecast accuracy significantly depends on model initialization, in this study we attempted to improve the model initialization by assimilating in-situ observations. The in-situ observations from near-surface and radiosonde networks contain a wealth of information about the lower atmosphere. Notably, near-surface observations are available at high temporal resolutions and dense spatial coverages, compared to the other sources of upper-air or remote sensing observations. However, several recent studies (Sahu et al., 2016; Dash et al., 2013; Hacker et al., 2007; Stensrud et al., 2009; Pu et al., 2013; Reen and Stauffer, 2010) showed significant improvements in short-range forecasts by assimilating near surface observations within a mesoscale NWP model. In this study, three-dimensional variational (3DVAR) data assimilation (Barker et al., 2004) is used within the Weather Research and Forecast (WRF) modeling framework WRF V3.6.1 (Skamarock et al., 2008) to improve the model initialization by assimilating the in-situ observations (temperature, dew point temperature, wind speed, wind direction and surface pressure) from the surface and radiosonde networks over the southern Californian coast. The purpose of this study is to examine the impact of assimilation of near-surface and radiosonde observations on the improvement of an intra-day GHI forecast. The initial and boundary conditions for the model simulation are derived from the North American Mesoscale model (NAM) 6-h analysis. The NAM data assimilation system assimilates most of the in-situ and remote sensing observations over United States to generate the 6-h analysis. Hence, in addition to the first objective, an hourly cyclic data assimilation is also carried out along with the six-hourly cyclic data assimilation for inter-comparison of the model simulations.

The following sections are organized as follows. Detailed descriptions of the experimental design including the WRF and 3DVAR models and marine layer events are given in Section 2. The sensitivity of the tuning parameters such as length scale, variance scale, and background error statistics (BES) on the analysis increment and model forecast are analyzed in Section 3. The validation of the model simulations for liquid water content (LWC), global horizontal irradiance (GHI), stratocumulus cloud base tendency and other meteorological parameters are described in Section 4. Finally, conclusions are provided in Section 5.

2. WRF model and data assimilation technique

2.1. WRF model configuration

The WRF version 3.6.1 model (Skamarock et al., 2008) is configured with two nested domains with horizontal resolutions of 8.1 km and 2.7 km, respectively. The inner domain is centered over the southern Californian coast (as shown in Fig. 1), which is subject to marine layer stratocumulus clouds during the summer months. The total atmospheric column within the model is divided into 75 terrain-following vertical σ-levels, and 50 levels are below 3 km altitude. The model initial and boundary conditions were derived from the 0000 UTC NAM data. Here the 0000 UTC NAM data is used for model initialization instead of the most recent 0600 UTC NAM, because it delivers lower mean absolute errors (MAE) for forecasted GHI. The physics parameterizations used for this study include Morrison 2-moment microphysics scheme (Morrison et al., 2009), Goddard scheme for shortwave and longwave radiation parameterization (Chou and Suarez, 1994), Kain-Fritsch New