Solar Energy 164 (2018) 12–24

Contents lists available at ScienceDirect

Solar Energy

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

Bias correction of a novel European reanalysis data set for solar energy applications

Christopher W. Frank[a,b,*], Sabrina Wahl[a,c], Jan D. Keller[a,d], Bernhard Pospichal[b], Andreas Hense[c], Susanne Crewell[b]

[a] Hans-Ertel-Centre for Weather Research, Climate Monitoring and Diagnostics, Germany
[b] Institute of Geophysics and Meteorology, University of Cologne, Germany
[c] Meteorological Institute, University Bonn, Germany
[d] Deutscher Wetterdienst, Offenbach, Germany

ARTICLE INFO

Keywords:
Global radiation
Global horizontal irradiance
Renewable energy
Solar energy

ABSTRACT

One of the major challenges during the transition phase of the energy system is to maintain the balance between energy supply and demand. Rising questions are often related to site mapping, variability, extremes and compensation effects for example. A fundamental source of information to answer these questions are high quality data sets of renewable energy related variables. As reanalyses provide all relevant data to assess wind and solar power generation over a long period of time (decades) in a gridded consistent way, they exhibit great potential in the field of renewable energy. A new regional reanalysis is COSMO-REA6, which covers the European domain over the years 1995–2014 with a horizontal resolution of about 6 km and a temporal resolution of 15 min. In this paper, we first assess the quality of the Global Horizontal Irradiance (GHI) provided by COSMO-REA6. High quality GHI measurements obtained through the Baseline Surface Radiation Network (BSRN) are used as reference and reveal systematic short comings in the reanalysis: (1) an underestimation of GHI in clear sky situations and (2) an overestimation of GHI in cloudy sky situations. In order to reduce these systematic regime dependent biases, a post-processing is developed. The applied post-processing method is a scaling based on orthogonal distance regressions for two different regimes, i.e., "clear sky" and "cloudy sky". The two regimes are distinguished by the use of a transmissivity threshold. The post-processed GHI shows a significant reduction of the systematic biases and an improvement in representing the marginal distributions. A spatial cross-validation shows the applicability to the whole model domain of COSMO-REA6. Moreover, COSMO-REA6 as well as the post-processed GHI data reveal an added-value when compared to global reanalysis ERA-Interim and MERRA-2. The higher resolution reanalysis exhibits a significantly better performance of representing GHI variability, as well as biases, RMSE and other conventional scores. The post-processed GHI data are freely available for download.

1. Introduction

For a sustainable planning of the transition towards renewable energy production, the assessment of the solar energy potential and its variability has become more and more important (Kleissl, 2013). Due to the high spatial and temporal variability of solar radiation long-term data over large domains are necessary to identify potentials for the production of renewable energy and risks regarding the growing dependency on this form of power generation. In this respect also the covariability of solar and wind energy becomes more important, as its anticorrelation is expected to balance the volatility of the individual sources to some extent (e.g. Bett and Thornton, 2016; Santos-Alamillos et al., 2012; Grams et al., 2017). More extensive studies simulate the electricity network in order to study the electricity grid as a whole system. In this context, realistic meteorological data allow studying for example the future need of storage and/or back-up capacity (e.g. Heide et al., 2010; Mulder, 2014).

Traditionally, solar energy potential has been assessed from measured time series of solar irradiance at ground level. This is limited in its geographical distribution especially if high temporal resolution (< 1 h) and high quality measurements are concerned. Most frequently, the Global Horizontal Irradiance (GHI) also called Surface Solar Irradiance (SSI), is measured within the networks of meteorological services. A spatially extended view is provided by satellite estimates like the

* Corresponding author at: Institute of Geophysics and Meteorology, University of Cologne, Germany.
E-mail address: cfrank@meteo.uni-koeln.de (C.W. Frank).

https://doi.org/10.1016/j.solener.2018.02.012
Received 14 September 2017; Received in revised form 20 December 2017; Accepted 3 February 2018

HelioClim project by MINES ParisTech (Blanc et al., 2011) or the SARAH (Solar surfAce RAdiation Heliosat) data set (Müller et al., 2015) produced by the Satellite Application Facility on Climate Monitoring (CM-SAF). They exploit geostationary satellite measurements to derive GHI for the full disk with up to hourly temporal and 0.05° spatial resolution. Atmospheric reanalyses compiled from observations and numerical weather prediction models provide not only GHI but rather the complete state of the atmosphere including the vertical profiles of wind, temperature, etc. in a physically consistent way. Therefore, these multiyear data sets which continually improve in resolution allow for a joint investigation of renewable energy resources (Bett and Thornton, 2016).

Global reanalyses that come at relatively coarse horizontal resolutions (40–100 km) are frequently used for investigating wind power generation (e.g. Staffell and Pfenninger, 2016; Ritter et al., 2015; Cannon et al., 2014; Kubik et al., 2013; Bett et al., 2013), but rarely for solar energy application (e.g. Boilley and Wald, 2015; Richardson and Andrews, 2014). One of the first applications by Lohmann et al. (2006) revealed large differences among two global reanalyses for monthly mean values at horizontal resolutions of about 200 km. When comparing two state-of-the-art reanalyses and satellite derived (Helio-Clim-1) daily solar irradiance with surface measurements across the globe Boilley and Wald (2015) find that a large part of the variability in surface radiation is not captured by the reanalyses. A reason for the deviation between reanalysis and measurements might arise from the difficulty to parameterize small scale processes related to clouds and aerosols including interaction with solar radiation.

The work of Richardson and Andrews (2014) indicated the potential of reanalyses in PV applications. For Ontario, Canada, Richardson and Andrews (2014) evaluated the use of NASA's global reanalyses MERRA (Rienecker et al., 2011) as input for PV modeling. They found that the modeled PV yields driven by MERRA results in just slightly higher errors than ground-measured driven results, despite relatively larger errors in the MERRA GHI data. Later, Pfenninger and Staffell (2016) showed a comparable performance of PV output simulations based on MERRA and MERRA-2 (Molod et al., 2015) compared to satellite estimates when aggregated to country-level.

One reason for the few studies using radiation from reanalyses for solar energy applications is the availability of the high quality satellite products. Many publications in the past are based on either solar or wind energy which caused the use of different data sources in the two fields. In the field of wind energy reanalyses products are frequently used (Rose and Apt, 2015) while for solar energy satellite products are found to be most accurate (Jia et al., 2013), at least compared to global reanalyses. To our best knowledge up to now high resolution regional reanalyses are not considered in the solar energy community. In recent times the question of co-variability and compensation effects of wind and solar energy become more and more important. Thus, the need of a common data source for both variables increased. Reanalyses provide wind and radiation in a physically consistent way in space and time. This is crucial for studying joint distributions, otherwise results and interpretation might be distorted due to physical inconsistencies. Using both variables from one source causes the question which reanalysis performs best in representing wind speed and radiation? This study addresses this question concerning the radiation part and takes regional reanalyses into account.

While global reanalyses mainly resolve clouds associated with synoptic disturbances, high resolution regional reanalyses have the potential to better describe smaller scale clouds associated with mesoscale processes like thunderstorms or orographic circulations and therefore are more suitable for solar energy applications. This paper investigates the quality of the novel European regional reanalysis COSMO-REA6 (Bollmeyer et al., 2015) available with a horizontal resolution of 6 km over a time period of 20 years and a temporal resolution of 15 min. Evaluation of COSMO-REA6 meteorological variables such as precipitation, temperature and wind speed (Wahl et al., 2017; Bollmeyer et al., 2015; Kaiser-Weiss et al., 2015; Borsche et al., 2016; Henckes et al., 2018) has already shown a superior performance with respect to the European Centre for Medium-range Weather Forecasting (ECMWF) Re-Analysis Interim data set (ERA-Interim, Dee et al., 2011) but the representation of radiation has not been addressed so far.

An even higher resolution data set is available for Central Europe with a horizontal grid spacing of 2 km (COSMO-REA2, Wahl et al., 2017) albeit for a much shorter time period of seven years (2007–2013). However, due to its larger range of applicability, e.g., cross-country energy trading, we focus on the long-term European data set COSMO-REA6.

With this paper we focus on the following questions:

1. How accurate is COSMO-REA6 GHI compared to ground observations?
2. Does COSMO-REA6 GHI improve upon global reanalyses, i.e. ERA-Interim and MERRA-2, in bias and variability metrics?
3. Can the expected biases and deviations be corrected with a post-processing algorithm?

In order to address these questions the paper is structured as follows. Section 2 describes the reanalysis and observational data sets. A quality assessment of GHI from reanalyses is given in Section 3 which reveals some systematic deficits under clear and cloudy conditions. Therefore a post-processing procedure to correct these issues is developed in Section 4. An evaluation, including a cross-validation, of the post-processed radiation fields is presented in Section 5. Section 6 summarizes our findings followed by the conclusions in Section 7.

2. Data sets

2.1. COSMO-REA6

COSMO-REA6 has been developed and produced within the Climate Monitoring Branch of the Hans-Ertel-Centre for Weather Research[1] and is based on the COnsortium for Small-Scale Modelling (COSMO) limited-area model (Schättler and Doms, 2011), which is part of the operational NWP model chain of the German Meteorological Service (DWD). It is a 20-year regional atmospheric reanalysis covering the European CORDEX EUR-11 domain with a horizontal resolution of 0.055° (approximately 6 km, see Fig. 1) and 40 vertical levels in terrain following coordinates. 3D model variables are archived every hour and 2D variables every 15 min. The most important variables, e.g. GHI, wind speed at the six lowest model level, can be downloaded via ftp (http://reanalysis.meteo.uni-bonn.de). In the reanalysis, a continuous nudging scheme is used to assimilate a wealth of observations into the model allowing for a detailed but temporally smooth representation of the prognostic variables (for further information the reader is referred to Bollmeyer et al., 2015).

The COSMO reanalyses uses the radiation scheme by Ritter and Geleyn (1992) based on the δ-two-stream approximation. The scheme is called every 15 min and calculates how solar radiation is modified in the atmosphere due to scattering and absorption by atmospheric gases, aerosol and clouds. The one dimensional radiative transfer is solved separately, once for the clears sky and once for the cloudy column which are subsequently combined according to cloud fraction. As the instantaneous distribution of clouds and water vapor are input to the radiation scheme, GHI reflects the strong dynamic variability of the atmosphere (Fig. 1).

The aerosol input to the radiation scheme is based on the Tanré et al. (1984) climatology and combines the effect of five different types of aerosols: continental, maritim, urban, volcanic and stratospheric background aerosols. The horizontal distribution of the aerosol types is based on the Global Aerosol Data Set (GADS, Koepke et al., 1997).

[1] https://www.herz-tb4.uni-bonn.de.