Using GEOS-5 Forecast Products to Represent Aerosol Optical Depth in Operational Day-ahead Solar Irradiance Forecasts for the Southwest United States.

Patrick T.W. Bunn, William F. Holmgren, Michael Leuthold, Christopher L. Castro

Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ

Submitted to *Journal of Sustainable and Renewable Energy* on July 3rd, 2020

Corresponding author address: Patrick Bunn, Department of Hydrology and Atmospheric Sciences, 1133 E. James E. Rogers Way, J W Harshbarger Bldg (#11), Room 322D, PO Box 210011, Tucson AZ 85721-0011

E-mail: ptwbunn@email.arizona.edu

Abstract

This study aims to improve operational day-ahead direct normal irradiance (DNI) forecasts in clear-sky conditions using the Weather and Research Forecasting (WRF) model. To create three different forecasting methods targeting the direct effect of aerosols on radiation, we use three different types of aerosol optical depth (AOD) data: (1) the Tegen aerosol climatology, (2) persistence of measured AERONET AOD, and (3) GEOS-5 gridded forecasts of AOD. We evaluate each method at the Solana Generating Station, a concentrating solar power plant near Gila Bend, Arizona, and the University of Arizona, Tucson. We perform a retrospective DNI forecast analysis and find that including GEOS-5 forecast AOD improved the DNI forecast compared to using an aerosol climatology at both locations. At Tucson, where AOD is measured, we find that persistence of measured AOD gives the best DNI forecast. However, the accuracy of that measured AOD reduces when translating it 225 km to Solana to forecast DNI 48 hours later. We then include the GEOS-5 AOD forecasts in one member of an operational forecast system and evaluate it against the other ensemble members that use the aerosol climatology. In clear-sky conditions, including GEOS-5 forecast AOD instead of the Tegen aerosol climatology, the DNI forecast root mean square error reduced by 27% at Solana. We found no significant differences during all-sky conditions because the relatively poor performance during cloudy conditions outweighs the improvements made in clear-sky conditions.

1. Introduction

Utility companies benefit from accurate power forecasts to manage different sources of generation. Solar power forecasts primarily rely on solar irradiance forecasts; therefore, those irradiance forecasts need to be accurate. Energy companies can trade energy based on accurate power forecasts. Load balancing, dispatching reserves, curtailing production, and operating energy storage are all management decisions that are informed in part by solar power forecasts (Kleissl, 2013; Tuohy *et al.*, 2015; Antonanzas *et al.*, 2016). These management decisions help energy companies with day-ahead energy scheduling (Brancucci Martinez-Anido *et al.*, 2016).

Concentrating solar power (CSP) systems use an array of mirrors or lenses to heat a fluid or illuminate specialized photovoltaic cells. These optics can only concentrate beams of direct sunlight. Direct normal irradiance (DNI) is downward shortwave radiation received at ground level in a plane normal to the Sun vector from an acceptance angle of ± 2.5° around the Sun. Diffuse radiation (DIF) is solar radiation from the sky, excluding DNI, which has been scattered by the clouds, aerosols, and the other atmospheric constituents. The mirrors cannot concentrate DIF; therefore, the amount of energy produced by CSP systems is maximized during clear-sky conditions and falls off sharply with cloud cover. To predict the energy input to CSP systems, we must accurately forecast DNI in clear-sky and cloudy conditions.

There are different methods to forecast DNI tailored for different timescales. For day-ahead forecasting, numerical weather prediction (NWP) is most appropriate (Jimenez *et al.*, 2016) and is the focus of this study. During cloud-free conditions, the representation of aerosol optical depth (AOD) is the most important factor governing the performance of day-ahead DNI forecasts for solar applications (Ruiz-Arias *et al.*, 2013); this is due to the *direct effect* of aerosols on surface radiation. The usefulness of DNI as a quantity is limited for forecast applications outside solar energy; therefore, many NWP forecasts do not represent AOD. Clear-sky DNI forecast error comprises radiation scheme error, measurement error, and AOD error. However, the portion of the DNI error from the radiation scheme and observations (from well-maintained instruments) is known to be smaller than the portion due to the AOD error (Holben *et al.*, 1998; Ruiz-Arias, 2013). The second most important factor determining the performance of day-ahead DNI forecasts for solar applications is precipitable water (PW). Accurate forecasts of PW have been made using a NWP model (WRF) with readily available forcing data (GFS), without the need for additional PW data (González *et al.*, 2013).

Ground-based observations of AOD, while being the most accurate measurement of AOD, lack spatial coverage. Satellite observations have more coverage than ground-based observations, but alone lack the accuracy required to accurately forecast DNI (Ruiz-Arias *et al.*, 2015). The best representation of aerosol optical properties lies in data that combines observations to coupled atmospheric chemistry and numerical weather prediction models (ACNWP). For an operational forecast system using these types of data, the problem shifts to computational expense and latency (the time from data initialization to availability). The Goddard Earth Observing System model version 5 (GEOS-5) is one of the few ACNWP models that combines satellite and ground-based measurements and meets the criteria for operational day-ahead forecasts because it has a latency of only 8 hours. Section 3.3 describes the GEOS-5 system in more detail.

Solar energy stakeholders want to know which NWP configuration and AOD data will produce the most accurate operational day-ahead DNI forecast for their solar power system? This study will evaluate different methods of incorporating AOD into operational day-ahead forecasts for solar energy applications using the Weather and Research Forecasting (WRF) model. We will compare DNI forecasts made using: no aerosol, an aerosol climatology, a persistence of measured AOD, and GEOS-5 forecast AOD. We also construct a clear-sky DNI persistence forecast, a non-NWP forecast that uses no additional aerosol data, for further comparison. We first perform a retrospective forecast analysis to test these different methods.

Then, we implement the best performing method in our operational forecast system and perform an operational forecast analysis. Unlike a retrospective forecast, an operational forecast is subject to computational, consistency, and time constraints. Considering such factors is essential when formulating a robust configuration.

While previous work (Jimenez *et al.*, 2016) has used GEOS-5 *analysis* AOD (+0 hours forecast) to improve retrospective DNI predictions, this study is the first to use GEOS-5 *forecast* AOD (≥ +24 hours forecast) in an operational forecast system. Contrasting to the controlled experiment in Jimenez *et al.* (2016), we exclusively use real-time data for our operational forecasts. To evaluate the GEOS-5 forecast accuracy in our forecasting periods, we will compare errors from GEOS-5 forecast AOD to analysis AOD and an AOD climatology.

The predominantly clear skies of the United States Desert Southwest mean that it is an ideal location for solar energy production, especially from concentrating solar power plants (Sengupta *et al.*, 2018). We will study two different sites in Arizona: the Solana Generating Station, near Gila Bend, and the University of Arizona, Tucson. However, we expect our results to be similar in other locations with similar climate conditions. At those sites we analyze DNI forecasts made for 105 predominantly clear-sky days, which is over five times the number of clear-sky days evaluated in Jimenez *et al.* (2016).

Atlantica Yield operates the Solana Generating Station, and Arizona Public Service purchases its power. Atlantica Yield and Arizona Public Service are the primary stakeholders motivating this research. Section 2 provides background information on aerosol optical properties, principally AOD, and describes their influence on radiation. Section 3 discusses the representation of AOD in operational forecasts. Section 4 explains configurations of the different forecasting methods that we use, and the observations used to evaluate them. Section 5 presents the forecast analysis and discussion. Section 6 concludes this study.

2. Background on aerosol optical properties

Extinction of radiation from a beam of sunlight, the *direct effect* of aerosols on radiation, is the primary source of error in clear-sky DNI forecasts (Ruiz-Arias, 2013; Jimenez *et al.*, 2016). Aerosol optical depth (AOD) describes the opacity of the cloud-free atmosphere in the visible portion of the solar radiation spectrum. AOD is calculated from the cumulative extinction of radiation from a direct-beam at each wavelength over the atmospheric path length (Holben *et al.*, 1998). The Beer-Lambert Law gives the equation:

$$I_\lambda(\mathrm{SFC}) = I_\lambda(\mathrm{TOA})\, e^{-\tau_\lambda/\mu} \quad [1]$$

where I_λ is irradiance at the surface (SFC) or top of the atmosphere (TOA) at a specific wavelength λ, μ is the atmospheric path length and τ_λ is the AOD at wavelength λ. Because AOD is a spectral quantity, it is measured at specific wavelengths. The Ångström law describes the dependence of AOD on wavelength, and allows for the conversion of AOD from one wavelength to another:

$$\tau_{\lambda,1} = \tau_{\lambda,0}\left(\frac{\lambda_1}{\lambda_0}\right)^{-\alpha} \quad [2]$$

where λ_0 and λ_1 are wavelengths, $\tau_{\lambda,0}$ is AOD measured at the specific wavelength λ_0, and α is the 470-870 nm Ångström exponent (Ångström, 1961). Typically, AOD at 550 nm is used in atmospheric radiative transfer problems because it is approximately in the middle of the visible region of the radiation spectrum and near the wavelength of peak solar emission. However, AOD is not necessarily measured at 550 nm, so the Ångström exponent is used for conversion. The Ångström exponent can be directly calculated from multiple AOD measurements at different wavelengths.

Other spectral parameters that directly influence the transmission of radiation through the atmosphere are: (1) the single scattering albedo (SSA) - which is a ratio of scattering to extinction of radiation within a beam of sunlight, and (2) the asymmetry factor (ASY) - the preferred direction of scattering radiation (ASY = 1 meaning forward, ASY = -1 meaning backward). Greater SSA values will result in more DIF and less absorption. ASY = 1 means more DNI compared to ASY ≤ 0, which results in less DNI and more DIF. AOD has the most dominant effect on DNI, whereas the impact of SSA and ASY are small. The treatment of these variables in NWP is described further in Section 3.1.

The *indirect effect* of aerosols on radiation stems from cloud-aerosol interactions and cloud-radiation feedbacks (Quaas *et al.*, 2009). Aerosols are needed to provide surfaces for cloud particles to form, cloud condensation nuclei, which can have varying effects on cloud droplet number concentration and, therefore, cloud optical thickness and cloud lifetime. Sophisticated parametrization of the *indirect effect* quickly becomes complex because cloud-aerosol interaction and feedbacks introduce large uncertainties. Improving the representation of the *indirect effect* of aerosols on radiation in NWP is a different research question.

3. Representing AOD in an operational forecast system

3.1 Incorporating AOD data in a NWP model

The operational forecasting system at the University of Arizona uses the Weather and Research Forecasting (WRF) model (Skamarock *et al.*, 2019). Representing AOD in NWP models, like WRF, requires a balance of realism, accuracy, and computational expense. Models that include full chemistry simulations can produce more realistic output; for example, WRF-Chem adds simulations of chemical interactions to WRF. However, performing operational WRF-Chem simulations is uncommon because initialization data sets are not readily available, and simulations are computationally expensive (Sessions *et al.*, 2015; Skamarock *et al.*, 2019).

Default NWP configurations are designed for general weather prediction, not solar forecasting, as they do not represent changes in AOD. For solar forecasting, we must activate specific radiation parameterization options that can utilize additional data. Which options and what data to use requires specific analysis for the application in question. The RRTMG scheme (Rapid Radiative Transfer Model for climate and weather models) is commonly used for parameterizing radiative transfer in day-ahead forecasts for solar energy applications. RRTMG uses a spectral range of 0.2 to 12.2 μm and, in clear skies, the expected accuracy compared to line-by-line calculations is about 4 Wm^{-2} for direct fluxes (Iacono *et al.*, 2008; Ruiz-Arias, 2013; Ruiz-Arias, Dudhia and Gueymard, 2014; Gueymard and Ruiz-Arias, 2015). Since the incorporation of most features from WRF-Solar version 1.2 (Jimenez *et al.*, 2016) into the WRF model (version 3.8), it has been possible to represent the *direct effect* of aerosols on radiation in simulations using the RRTMG scheme with two different options:

The first WRF/RRTMG option (namelist option aer_opt=1) uses the Tegen global aerosol climatology data (Tegen and Fung, 1994; Tegen *et al.*, 1997) as an input to RRTMG. This climatology is comprised of monthly values of five species of aerosol (organic carbon, black carbon, sulfate, sea salt, and dust) at each model-level, aggregated into a total column AOD. The data is on a spectral grid, which is equivalent to 5°x4° (625 km) grid-spacing at the equator. The climatology uses a global 3D transport model described in Tegen and Fung (1994) to create a 15-year simulation that is evaluated using ground- and satellite-based observations in Tegen *et al.* (1997). Zubler *et al.* (2011) showed that in areas with complex dust emissions, aerosol climatologies have substantial difficulties reproducing observed AODs. This is relevant to the United States Desert Southwest, where dust emissions can vary on inter-day timescales.

The second WRF/RRTMG option (aer_opt=2) allows 3D (x, y, t) fields of aerosol optical properties (e.g., AOD at 550 nm) to be incorporated into radiation calculations via the WRF auxiliary inputs. These 2D static fields are user-defined and can be either uniform values, different aerosol climatologies, or aerosol analysis/forecasts products. The user can also specify other properties such as the Ångström exponent, the single-scattering albedo (SSA), and the asymmetry factor (ASY) in a similar fashion. Of these optical properties, AOD has the greatest influence on incoming solar radiation; therefore, it should be input as a 3D (x, y, t) field. The dependence of AOD on wavelength, the Ångström exponent, has the next largest influence, and we will examine the impact of using a 3D field versus a single climatology value for DNI forecasts, in Section 5.1.1. For the remaining optical properties, we use the Ruiz-Arias, Dudhia, and Gueymard (2014) parametrization in this study. The rural aerosol type, in this parameterization, prescribes that the aerosol load is a mixture of 70% water-soluble and 30% dust. In contrast, the urban aerosol type is 56% water-soluble, 24% dust, and 20% soot-like particles. We choose rural instead of urban because our validation points are typically not in urban or industrial areas, and dust is a key constituent of the aerosol load in the United States Desert Southwest. Gueymard and Ruiz-Arias, (2015) and Jimenez *et al.*, (2016) show this configuration of the RRTMG radiation scheme to improve the representation of the *direct effect* of aerosols on radiation; therefore, we can expect improvements to clear-sky DNI forecasts compared to a configuration using an aerosol climatology.

3.2 Metrics to evaluate data

We use root mean square error (RMSE), mean bias error (MBE), and RMSE skill score (SS) to quantify analysis and forecast errors;

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_{i,f} - x_{i,o})^2} \qquad [3]$$

$$\mathrm{MBE} = \frac{1}{N}\sum_{i=1}^{N}(x_{i,f} - x_{i,o}) \qquad [4]$$

where $x_{i,f}$, $x_{i,o}$ are the i^{th} entry of the forecast (f) and observation (o) time series (length = N), and:

$$\mathrm{SS_{RMSE}} = 1 - \frac{\mathrm{RMSE_{fx}}}{\mathrm{RMSE_{a\ different\ fx}}} \qquad [5]$$

where a forecast RMSE is used as a benchmark to compare a different forecast. We calculate root mean squared difference (RMSD) and mean bias difference (MBD) using Equations 3 and 4 but with two forecast or observations values instead of one of each. Full derivations are available in Levine and Wilks (2006).

3.3 Observing and Forecasting AOD

An accurate forecast requires proper model initialization. In our case, a good model initialization represents the current state of AOD; this requires observations. The Aerosol Robotic Network (AERONET) measures optical properties from the ground at several sites, one of which is at the University of Arizona in Tucson, AZ (32.23N, 110.95W). A sun-photometer measures AOD and the Ångström exponent at eight wavelengths from 340nm to 1640nm. From observations at 500nm and 675nm, AOD at 550 nm can be calculated as described in Section 2. Holben *et al.* (1998) report an uncertainty of < ±0.01 for AOD measurements from AERONET sites, justifying its use as a benchmark. However, AERONET is a sparse network, with no sites at solar power systems. When using measured AOD to forecast DNI, frequent data gaps in the AERONET network present a forecasting challenge that needs to be considered for an operational configuration.

GEOS-5 (Goddard Earth Observing System, version 5.16) is an Earth-system model that produces operational forecasts (Suarez *et al.*, 2008). Gridded forecast AOD and Ångström exponent are available on a 0.3125°x0.25° global grid every 3 hours. These forecasts come from a prognostic aerosol module that is based on the Goddard Chemistry, Aerosol, Radiation, and Transport Model (GOCART) (Chin *et al.*, 2000, 2002; Colarco *et al.*, 2010). GEOS-5 has a data assimilation system where satellite observations of aerosols are calibrated with ground-based observations (AERONET) and input to GOCART. The GOCART model traces dominant aerosol species and couples them to atmospheric variables at each time-step. Aerosol optical properties are then calculated across gridded horizontal areas for each vertical layer from the aerosol number concentration and aerosol type. The aerosol optical properties are quality controlled using neighboring values and assimilated using a local displacement ensemble methodology (Reale, Lau, and da Silva, 2011; Randles, Colarco, and Da Silva, 2013). Finally, a forecast AOD at 550 nm is calculated based on the modeled aerosol type and distribution. Jimenez *et al.*, (2016) compared DNI forecasts made using WRF/RRTMG with the Tegen aerosol climatology to WRF/RRTMG with GEOS-5 analysis AOD. For the 20 clear-sky days evaluated at 7 Surface Radiation Network sites, Jimenez *et al.*, (2016) reported a reduction in forecast DNI RMSE from 66 to 41 Wm^{-2}.

Schroedter-Homscheidt *et al.* (2013) used a gridded forecast aerosol product, the European Centre for Medium-range Weather Forecasting (ECMWF) Monitoring Atmospheric Composition and Climate project (MACC, 2013), to examine the sensitivity of DNI to differences in AOD versus ground-based measurements. They found that MACC AOD forecasts performed better or equal to a persistence forecast based on ground-based measurements. Schroedter-Homscheidt *et al.*, (2013) concludes that the effect of intra-day variability of AOD on DNI is small.

Direct AOD observations, like those from AERONET, can be used to evaluate gridded AOD data. Figure 1 shows the daily values of AOD at 550 nm from four data sources (GEOS-5 analysis, ECMWF-MACC analysis, the Tegen aerosol climatology, and AERONET observations), at two locations (Tucson and Yuma). We see some seasonal variability, with observed AERONET AOD on average higher during summer months. Also, we see an inter-day variability of AOD, which the climatology fails to represent. The misrepresentation of AOD can cause significant DNI forecast errors because of the sensitivity of DNI to AOD, resulting from their exponential relationship (see Equation 1). At Tucson, using AERONET as a benchmark, the GEOS-5 analysis AOD RMSE is 0.061, and MBE is 0.045 (see Table I). . At Tucson and Yuma, the GEOS-5 analysis AOD has lower RMSE and MBE compared to

ECMWF-MACC analysis AOD. At Tucson, the differences in error from GEOS-5 analysis AOD and the Tegen climatology are within AERONET measurement uncertainty (±0.01). At Yuma, these differences in error are greater than 0.01, and there are more frequent data gaps at the Yuma AERONET site compared to Tucson; the first three months of 2017 are missing in Figure 1. At the time of this study, there are no research-quality measurement sites for DNI or GHI in Yuma, so the evaluation of DNI forecasts is not possible. Therefore, we will only continue to study the Tucson site.

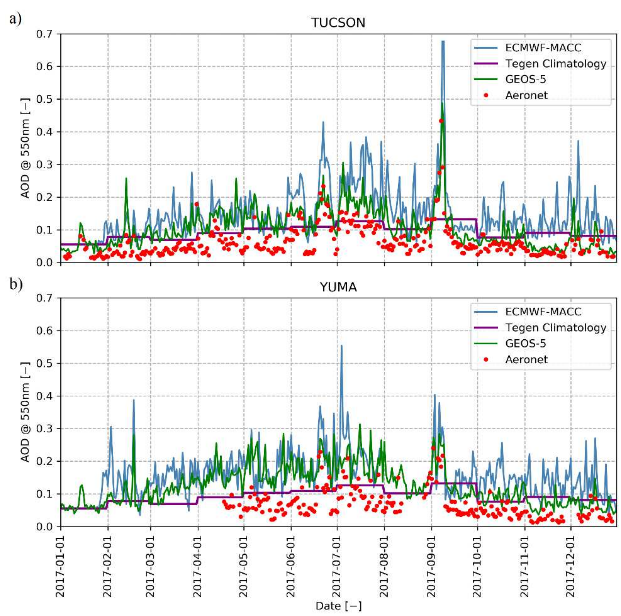


Figure 1: Time series of daily AERONET observations (red points) of Aerosol Optical Depth measured at 550 nm at two locations Tucson (a), and Yuma (b) for 2017. Daily GEOS-5 and ECMWF-MACC analysis AOD are shown alongside the Tegen monthly aerosol climatology (see legend).

Table I: Statistics comparing daily GEOS-5 analysis, ECMWF-MACC analysis and, the Tegen monthly aerosol climatology, to AERONET observations at Tucson and Yuma for 2017. MBE and RMSE are shown. AERONET measurement uncertainty is ±0.01.

	Tucson N=365			**Yuma** N=195		
	GEOS-5	**MACC**	**Tegen**	**GEOS-5**	**MACC**	**Tegen**
RMSE [-]	0.061	0.108	0.053	0.086	0.120	0.054
MBE [-]	0.045	0.090	0.031	0.073	0.105	0.036

The discussion above characterized the GEOS-5 analysis AOD accuracy, but we are most concerned with the GEOS-5 forecast AOD accuracy. Ideally, we would calculate AOD forecast performance for several years, similar to what we have done for the analysis, however, only the most recent three months of GEOS-5 forecast data are available at a given time due to their storage limitations (the analysis AOD, +0 hours forecast, is stored long-term). Here we instead establish that the GEOS-5 forecast accuracy is similar to the GEOS-5 analysis accuracy during the retrospective and operational periods studied (see Section 4.1), and we assume that the correspondence between forecast and analysis accuracy remains similar throughout the year. Table II compares errors at Tucson from the GEOS-5 forecast to analysis AOD. In our study periods, the differences in RMSE and MBE between the forecast and analysis AOD are < 0.01, from Table II. The uncertainty of AERONET observations is ±0.01. Therefore, the forecast AOD used is representative of the analysis data during our forecast periods. We use

Table II: Statistics for our forecast periods (see Section 4.1) comparing errors from GEOS-5 analysis and forecast AOD using AERONET observations. Errors from the Tegen climatology are shown for comparison. MBE and RMSE are shown. Also, MBD and RMSD are shown to compare GEOS-5 AOD at Solana to Tucson.

Tucson (AERONET obs.)		**Retrospective Fx Period N=57**		**Operational Fx period N=74**	
		Analysis AOD	**24-hour Fx AOD**	**Analysis AOD**	**48-hour Fx AOD**
Tegen Climatology AOD	**RMSE [-]**	0.073	-	0.051	-
	MBE [-]	0.027	-	0.047	-

GEOS-5 AOD	**RMSE [-]**	0.056	0.036	0.041	0.042
	MBE [-]	0.028	0.006	0.038	0.036
GEOS-5 AOD (Solana - Tucson)	**RMSD [-]**	0.038	0.021	0.025	0.022
	MBD [-]	0.013	0.006	0.019	0.014

+24 hour AOD forecasts for the retrospective period and +48 hour forecasts in the operational period, due to the latency of GEOS-5 forecast AOD (see Section 4.1). However, Table II shows similar errors for both forecast hours.

At Tucson, the GEOS-5 AOD and the Tegen climatology have similar magnitude errors suggesting a similar representation of AOD from using either data. A key question at this point is: are AERONET measurements from Tucson sufficiently representative of the next day's AOD at the Solana Generating Station in Gila Bend 225 km away? (Figure 2 shows a photograph of the Solana Generating Station and Figure 3 shows its location on a map). Solana is where AOD needs to be better represented, as this is where DNI is to be forecast for the solar power system. Furthermore, a follow-up question is: is the error introduced by translating AERONET observations to a different location/time less than the GEOS-5 forecast error? We cannot directly answer these questions because AOD is not measured at the power plant. However, we can study the relative accuracy of the DNI forecast obtained from either AOD data source. To do this, we must first understand the temporal and spatial variability of GEOS-5 forecast AOD.

Figure 2: Photograph of the Solana CSP system near Gila Bend, AZ, USA (Bunn, 2019)

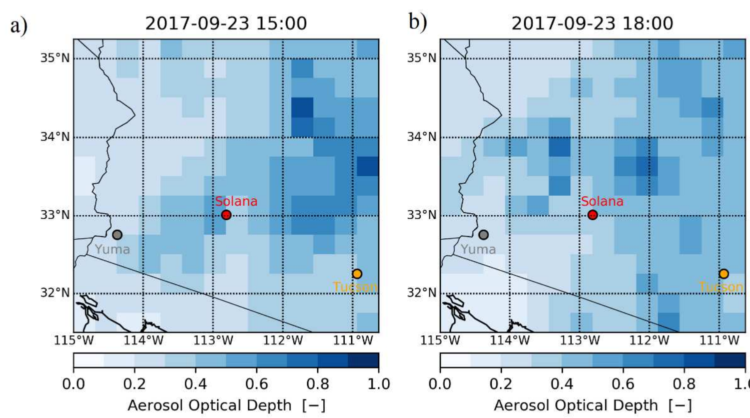


Figure 3: Gridded GEOS-5 Aerosol Optical Depth forecasts in the United States Desert Southwest at (a) 1500Z and (b) 1800Z on September 23rd, 2017. The Solana Generating Station is marked in red. The AERONET observations of AOD performed at the University of Arizona, Tucson (orange), and Yuma (gray) are also marked (see the Supplementary Material for an animation of 29 days of GEOS-5 forecast AOD).

The two snapshots of GEOS-5 forecast AOD in Figure 3, show different spatial patterns 3 hours apart. In Table II, the RMSDs comparing GEOS-5 AOD at Solana to Tucson show consistent differences ($\geq$ 0.02), that are greater than AOD measurement uncertainty, demonstrating the independence of GEOS-5 AOD at these locations. While the AERONET observations capture the temporal variation in AOD at a single point, they miss the spatial characteristics seen in Figure 3. The aerosol climatology has spatial and temporal representation; however, it will miss the inter-day variability in those dimensions. The GEOS-5 products represent the short-term spatial distribution of AOD but are likely not as accurate as ground-based observations. We could, therefore, expect smaller errors from DNI forecasts made using AERONET observations near to the measurement site, with errors increasing the further from the measurement site. Away from an AOD measurement site, we are restricted to inferring better AOD representation from a better DNI forecast.

4. Experiment Details

4.1 Forecast Data

4.1.1 NWP Forecast Configuration

First, we will test three different methods in a retrospective forecasting period, and then implement the best performing method in an operational forecasting period. For the retrospective period, September through October 2017, we use the WRF model version 4.0 (Skamarock *et al.*, 2019) with a domain of 100x100 cells with a horizontal spacing of 5.4 km and 33 vertical levels. The 0.25° National Centers for Environmental Prediction Global Forecast System (GFS) data is used to force the simulations every 3 hours (NCEP, 2015a). The RRTMG radiation scheme is used for short- and longwave radiation. Other parameterization schemes used are the Thompson microphysics scheme (Thompson *et al.*, 2008) and the Asymmetric Convection Model 2 planetary boundary layer scheme (Pleim, 2007). The time-step is 30 seconds with RRTMG called every time-step.

Table III summarizes the three forecast methods tested retrospectively. The 'No Aerosol' experiment serves as a control experiment with WRF/RRTMG in its default aerosol configuration. The 'Climo Fx' uses the Tegen *et al.* (1997) monthly aerosol climatology, which was the configuration for our operational forecasting system at the outset of this study. This climatology has a grid-spacing of 625km at the equator, resulting in approximately one AOD value for Arizona per month. 'Aeronet Fx' is a persistence AOD forecast where the previous day's AOD and Ångström exponent values measured at the Tucson AERONET site are used to forecast DNI for the next day. 3-hourly averaged AOD values are translated onto a grid, yielding a 2D data set of uniform AOD values. This enables the same namelist option to be used for this and the 'GEOS-5 Fx' configuration (aer_opt=2, aer_aod550_opt=2). The 'GEOS-5 Fx' uses day-ahead GEOS-5 forecasts of AOD and Ångström exponent. All 2D aerosol data is extracted and interpolated using a 4-point bi-linear method at the pre-processing stage (WPS), so the AOD values are static between input times.

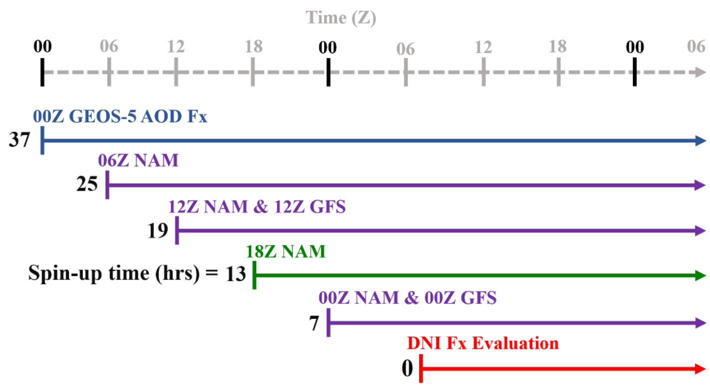
Table III: Table describing the forecast methods implemented in a retrospective manner from September through October 2017.

Forecast Name	Description
0: No Aerosol	No aerosol data used
1: Climo Fx	AOD at 550nm and Ångström exponent are calculated from the Tegen climatology data set.
2: Aeronet Fx	The previous day's observations of AERONET AOD at 550 nm and 470-870 nm Ångström exponent (t, x, y) are input every 3hrs as a uniform value over the forecast domain.
3: GEOS-5 Fx	GEOS-5 forecast AOD at 550 nm and 470-870 nm Ångström exponent are input every 3hrs to the forecast domain.

For the operational forecasting period, April through June 2019, we use WRF (version 3.9.1.1) for a domain of 456x599 cells with a horizontal spacing of 5.4 km and 38 vertical levels. Operational forecasts are initialized daily at 00Z, 06Z, 12Z, and 18Z using NAM forcing (NCEP, 2015b), and GFS forcing at 00Z and 12Z. This is an operational forecast system where numerous ensemble members are run daily for various applications, one of which is solar energy forecasting. The operational forecast system is, therefore, subject to computational, consistency, and time constraints. The current configuration and latest regional forecast products are available at (http://www.atmo.arizona.edu/?section=weather&id=wrf).

We incorporate the GEOS-5 00Z AOD forecast into the NAM 18Z DNI operational forecasts. We use the GEOS-5 00Z AOD forecast, rather than the 12Z AOD forecast, due to the 8-hour latency of the GEOS-5 product (see Figure 4 for the operational forecasting timeline). We will compare the NAM 18Z configuration, with the GEOS-5 forecast AOD, to the other ensemble members that use the Tegen aerosol climatology. Despite the differing initializations, this will provide a fair comparison because the effect of the differing AOD representation on radiation during clear-sky conditions will outweigh any differences in these neighboring initializations; see Section 5.2.2 for supporting evidence. Minor initialization differences could change modeled cloud location and timing; however, both observation and forecast need to be determined as clear-sky for an evaluation to take place, see Section 4.2.3 for filtering methods.

In the operational configuration five more model vertical levels are used. This will not affect DNI forecasts during clear-sky conditions but could improve forecasts of clouds and thus DNI in cloudy conditions. The domain is smaller for the retrospective forecast period. However, the two evaluation sites are > 20 grid points from the boundary in each forecast domain, so the errors from boundary conditions will not influence them. Of critical importance is the radiation parameterization scheme, RRTMG, which is used consistently for each method in both forecasting periods.



<u>**Figure 4:**</u> Schematic showing forecast spin-up times for the GEOS-5 AOD forecast (blue), WRF DNI forecasts (green: with GEOS-5 Fx, purple: with Climo Fx) with various forcing data (NAM/GFS). Spin-up times (in hours) are shown in black and are calculated from the beginning of initialization data to the beginning of evaluation (red). Note that the time from 00Z GEOS-5 AOD Fx to 07Z DNI Fx Evaluation is +37 hours.

4.1.2 DNI Persistence Forecast Configuration

An important benchmark in solar irradiance forecasting is the persistence model. For clear-sky conditions, we construct a DNI persistence forecast from hourly clear-sky observations. For a given hour of the day, we take the most recent clear-sky DNI observation for that hour from a previous day with a limit of -7 days. For all-sky conditions, we construct a strict 24-hour persistence of DNI, where we use yesterday's DNI observations to forecast for today.

4.2 Forecast Evaluation Data

We evaluate DNI forecasts at the Solana Generating Station system and the University of Arizona, Tucson. For the retrospective forecast period, instantaneous forecast values are evaluated every 15-minutes against 1-minute instantaneous observations. Instantaneous forecast values are used every hour in the operational forecasting period due to data archive limitations, but again they are evaluated against 1-minute instantaneous observations.

4.2.1 Solana Generating Station

The Solana Generating Station operates an Eppley normal incidence pyrheliometer that has a spectral range of $\lambda = 0.25$ to $3\ \mu\text{m}$ and measures DNI with an estimated uncertainty of 2%. It is appropriate to use this instrument to evaluate the DNI output from the RRTMG radiation scheme. For global horizontal irradiance (GHI), an unshaded Kipp & Zonen pyranometer is used (CMP22) with an estimated uncertainty of 2%. These are regarded as industry standards for radiation measurements and are maintained regularly.

4.2.2 OASIS NREL

The University of Arizona maintains a research class sun-photometric station (OASIS: 32.23N, 110.95W), which is part of a network of high-performance stations under the supervision of the National Renewable Energy Laboratory (NREL). The data can be accessed through NREL's data portal (Andreas, A. Wilcox, 2010). DNI is observed using a Kipp & Zonen CHP1 pyrheliometer instrument mounted on an automatic sun-following tracker. The CHP1 has a spectral range of $\lambda = 0.2 - 4\ \mu\text{m}$ and an estimated uncertainty of 3 - 4%, making it also an appropriate instrument to evaluate DNI output from RRTMG. More specific information about the instruments and maintenance is available on the portal, but it is reasonable to attribute confidence to these observations relative to DNI forecasting errors.

4.2.3 Filtering Methods

We analyzed forecast errors in clear- and all-sky conditions. Observations were screened for clear-sky conditions using a clear-sky filter on measurements of GHI (Reno and Hansen, 2016). Forecasts were filtered using the clear-sky variable (SWDDNIC) in WRF output; if forecast irradiance deviates from the clear-sky variable by more than $1\ \text{Wm}^{-2}$ then it is flagged as cloudy. For a given time to be considered clear-sky conditions, *both* observation and forecast must be determined as clear-sky by these filters. All DNI forecasts were passed through a zenith angle filter ($\theta_s < 70°$) to restrict evaluations to peak sun hours. This is done because in times when $\theta_s > 70°$ solar energy power production is insignificant in comparison to peak sun hours.

5. Results

5.1 Retrospective forecast analysis

5.1.1 Retrospective clear-sky conditions

Figure 5 shows a time series of the 470-870 nm Ångström exponent, AOD at 550nm, and daily forecast RMSE for DNI during the retrospective forecast period, at both Solana and Tucson. We calculate a RMSE for each day that has more than five clear-sky data points. Gaps in the time series show cloudy days.

The forecast period begins with 12 days of high AOD (>0.2) and relatively constant Ångström exponent, suggesting a uniform type of aerosol. The high AOD was caused by smoke that originated from California and Pacific Northwest wildfires in late August and early September. A high-pressure system over the western United States advected the plume to the Southwest (see Supplementary Material for a synoptic sea-level pressure map and link to satellite imagery archive). The days impacted by the smoke event at the beginning of the retrospective forecast period have much higher RMSEs than when the smoke has passed, September 15th onwards in Figure 5. The final 11 days of this forecast period also have higher RMSEs. This is a period of relatively low AOD, which likely causes the highly variable Ångström exponent seen in both the GEOS-5 and AERONET values. Small errors in one of the low AOD measurements used to compute 470-870 nm Ångström exponent is the likely cause of this variability (Kato *et al.*, 2000). Despite the differences in Ångström exponent values going into each forecast, the Aeronet Fx and GEOS-5 Fx have similarly large errors. This demonstrates the weaker influence of the Ångström exponent compared to AOD if we are to assume that AERONET observations are closer to the true value.

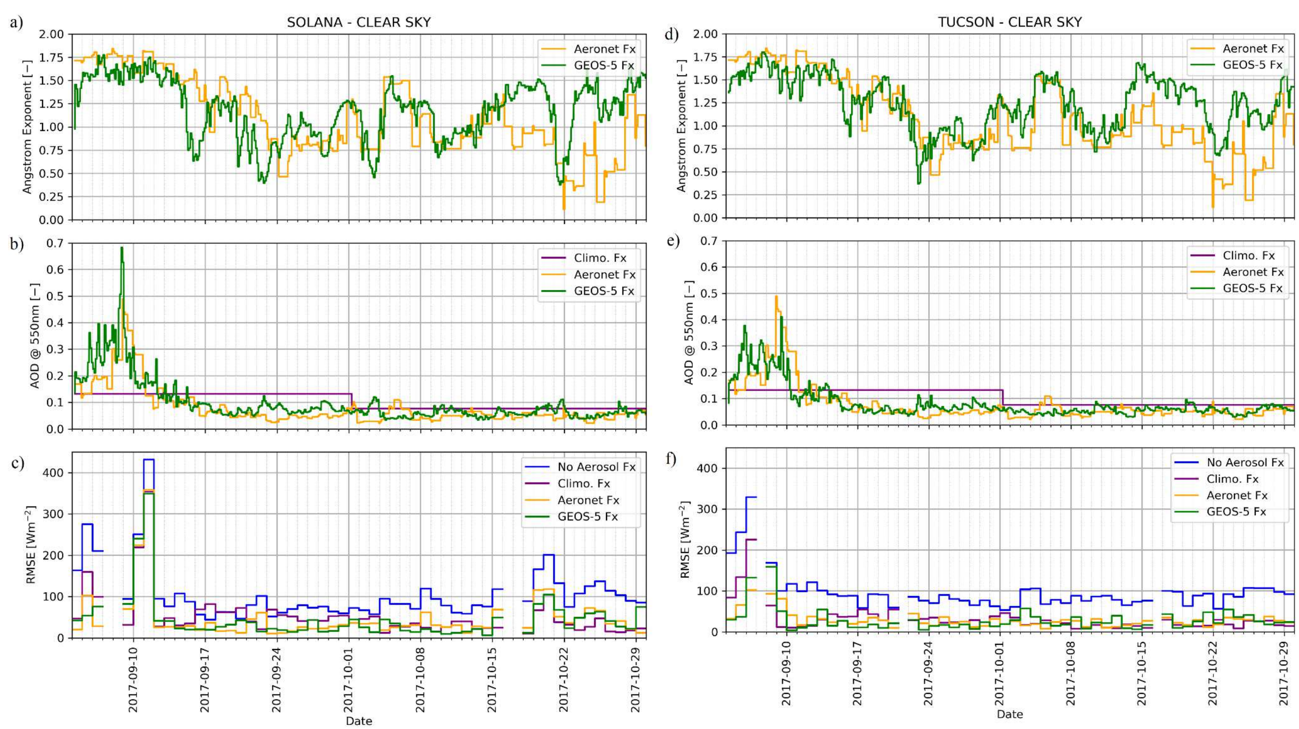


Figure 5: Time series of forecast 470-870 nm Ångström exponent (a, d) and 550nm Aerosol Optical Depth (b, e) from two different data sets, AERONET (yellow) and GEOS-5 (green) for the retrospective forecast period. Panels c) and f) show the daily RMSE value of DNI forecasts during clear-sky conditions for each of the forecasts (see legend) described in Table III, evaluated against observations at the Solana Generating Station (a, b, c) and Tucson (d, e, f). Gaps in this time series indicate cloudy conditions. Note Aeronet Fx Ångström exponent and AOD are the same at both locations by construction.

There is a four-week period (September 13th – October 15th) where DNI forecast errors are lowest. During this period, the forecasts at both locations tend to be similar for each methodology. However, at Solana, there are groups of days that show noticeably better performance from the GEOS-5 Fx versus the Aeronet Fx (for example, Sept 28th – 30th, Oct 2nd-3rd, and Oct 6th-14th). These are days when the Aeronet Fx persistence method has AOD values about 0.05 lower than the GEOS-5 values. The weaker performance of the Aeronet Fx is due to the spatial variability of AOD because those referenced days have relatively constant inter-day AOD values. This data suggests that the temporal characteristics of AOD at Solana are captured better by the GEOS-5 forecast AOD than the Aeronet Fx generated with two days prior observations from Tucson. This is also supported by analyzing the Tucson site; the Aeronet Fx performs better at Tucson than at Solana because AOD is measured in Tucson. Comparing the Aeronet Fx and GEOS-5 Fx to the Climo Fx, we see that on average, both outperform Climo Fx at both Tucson and Solana. During the four weeks (September 13th – October 15th), where errors are lowest for all forecasts, the daily RMSEs for the Climo Fx are typically less than No Aerosol but greater than GEOS-5 Fx and Aeronet Fx.

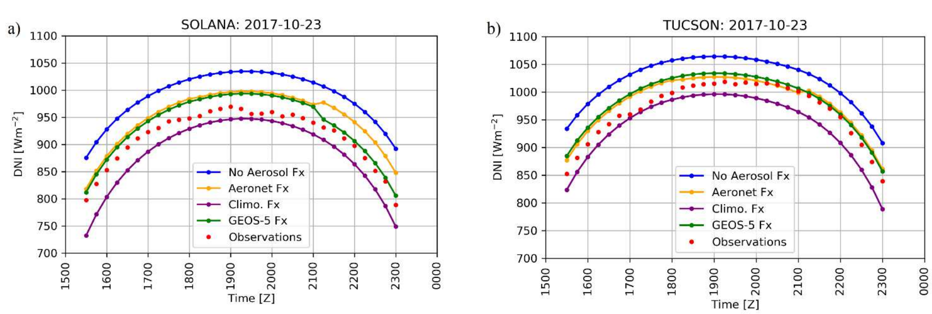


Figure 6: Example time series of DNI forecasts at Solana (a) and Tucson (b) for October 23rd, 2017. Forecast DNI from each forecast method described in Table III is shown against observations.

Figure 6 shows the DNI forecast errors for October 23rd, 2017. The intra-day variability of forecast AOD causes steps in the DNI forecast. At Solana, the jumps in forecast DNI at

2115Z are a product of higher and lower forecast AOD values being introduced at 2100Z for GEOS-5 Fx and Aeronet Fx, respectively. The similar magnitude of the jumps is coincidental, but the opposite sign shows better performance for GEOS-5 Fx. Comparing forecasts at the two evaluation sites for this single day in Figure 6, we can again see the Aeronet Fx method performs better at Tucson and worse at Solana. In contrast, GEOS-5 Fx performs well at both locations. GEOS-5 forecast AOD is likely relatively accurate at both locations; however, Tucson AERONET AOD is less representative at Solana.

An additional forecast method, 'GEOS-5 Fx const. Ang Exp', is provided in the Supplementary Material. The difference between the forecast with a constant climatological Ångström exponent (GEOS-5 Fx const. Ang Exp α=1.3) and 2D gridded data varying in time (GEOS-5 Fx) is minimal, with RMSE typically $\leq 5\ \mathrm{Wm}^{-2}$ different. There are examples where the over- and under-estimation of Ångström exponent relative to this climatological value result in marginally different performance. Including the GEOS-5 Ångström exponent forecast could be considered superfluous to improving DNI forecasts. However, it does not degrade forecasts, and it is not significantly more effort than including only the GEOS-5 AOD forecast. Also, we note there are no dust events with moderate AOD during the retrospective forecast period. If there were, then the Angstrom exponent and it's forecast accuracy could have more impact on the DNI forecast.

Table IV shows the statistical metrics for the forecasts at Solana (left) and Tucson (right). At Solana, including any AOD data in the forecast decreases the DNI RMSE values for clear-sky conditions by about 50 Wm^{-2}. The difference between each of the 'Aeronet Fx' and 'GEOS-5 Fx' forecasting methodologies is < 10 Wm^{-2} in RMSE. The mean bias error (MBE) is positive because, without tropospheric AOD at 550nm represented in the model, radiation can pass through the atmosphere with less scattering and absorption; therefore, overestimating DNI. The MBE decreases in the GEOS-5 Fx to 2 Wm^{-2}; however, the Aeronet Fx is still positively biased at 23 Wm^{-2}. The Climo Fx is negatively biased at -11 Wm^{-2} suggesting an overestimation of AOD. Clear-sky DNI persistence performs worse than Climo Fx at Solana and similarly at Tucson.

The RMSE skill score (SS_{Climo} = 0.29) at Tucson shows Aeronet Fx to be superior to GEOS-5 Fx. This is surprising given the simplicity of the AOD persistence method but unsurprising because the AOD used is measured at this site. The superior Aeronet Fx at Tucson reinforces the point that using direct AOD observations can produce the best DNI forecast at that location. However, with a SS_{Climo} = 0.06 at Solana, the accuracy of that observed AOD reduces when translating it 225 km to Solana to forecast DNI 48-hours later. At Tucson, the smaller DNI persistence skill score compared to Aeronet Fx is due to the number and quality of clear-sky observations during the smoke event. Thus, the DNI persistence forecast has a longer lead time compared to Aeronet Fx for this section of the forecasting period. The decrease in RMSE and increase in SS_{Climo} for all forecasting methods at Tucson, compared to the Solana site, can be attributed to relatively better AOD forecasts at Tucson during the week of highest AOD, 4th-13thSept. When considering the four weeks of lowest errors, there is more consistent performance of each configuration at both locations with RMSEs between 30 and 40 Wm^{-2}.

Table IV: Statistics comparing each forecast method described in Table III to observations during clear-sky conditions for the retrospective forecast period. Clear-sky DNI persistence is also shown for comparison. RMSE, MBE, and an RMSE-based skill score (SS_{Climo}) relative to the 'Climo Fx' forecast is shown on each row.

Solana N=1625	**No Aerosol**	**Climo Fx**	**Aeronet Fx**	**GEOS-5 Fx**	**DNI Pers**	**Tucson** N=1385	**No Aerosol**	**Climo Fx**	**Aeronet Fx**	**GEOS-5 Fx**	**DNI Pers**
RMSE [Wm^{-2}]	125	78	73	71	85	**RMSE [Wm^{-2}]**	106	46	33	33	45
MBE [Wm^{-2}]	98	-11	23	2	2	**MBE [Wm^{-2}]**	88	1	19	15	-2
SS_{Climo} [-]	-0.6	0	0.06	0.08	-0.09	**SS_{Climo} [-]**	-1.31	0	0.29	0.28	0.04

Table V: Same as Table IV but for all-sky conditions and now with a 24-hour DNI persistence forecast for comparison.

Solana N=1986	**No Aerosol**	**Climo Fx**	**Aeronet Fx**	**GEOS-5 Fx**	**24 hr DNI Pers**	**Tucson** N=1986	**No Aerosol**	**Climo Fx**	**Aeronet Fx**	**GEOS-5 Fx**	**24 hr DNI Pers**
RMSE [Wm^{-2}]	265	219	204	200	241	**RMSE [Wm^{-2}]**	284	238	239	237	310
MBE [Wm^{-2}]	152	46	71	53	-7	**MBE [Wm^{-2}]**	152	62	72	69	-9
SS_{Climo} [-]	-0.21	0	0.07	0.07	-0.08	**SS_{Climo} [-]**	-0.19	0	-0.01	0.01	-0.35

5.1.2 Retrospective all-sky conditions

Though clear-sky conditions are best for solar power generation and demonstrate the effect of AOD on DNI most, we also evaluate DNI forecasts in all-sky conditions for completeness. Table V shows the statistical differences in each forecast configuration again but for all-sky conditions. Differences are difficult to distinguish when comparing each of the AOD-aware methodologies in all-sky conditions, and the 24-hour DNI persistence forecast performs worse than AOD-aware forecasts. Since the cloudy forecast performance is relatively weak, the large errors in cloudy conditions dominate the statistical metrics. There is no discernable difference between the all-sky forecast at Solana and Tucson. The large positive MBEs and large RMSEs show that instances of observed cloud but forecast clear-sky are common.

5.2 Operational forecast analysis

5.2.1 Transitioning from retrospective to operational forecasting

The retrospective forecast analysis informed a decision to introduce GEOS-5 AOD forecasts into the operational forecasting system at the University of Arizona. GEOS-5 AOD forecasts are initially tested in one ensemble member (NAM 18Z) of the operational forecast system from April through June 2019 with a plan to introduce the GEOS-5 AOD to all ensemble members after a successful testing period. While the season is different for the forecasting periods, the day-ahead DNI forecast error in clear-sky conditions is primarily driven by the inter-day variability of AOD, not the seasonal variability. During the operational forecasting period, the other ensemble members use the Tegen *et al.* (1997) climatological AOD. This section of the results will focus on the differences between these two methodologies for incorporating AOD into DNI forecasts. The Ångström exponent was set to the climatological value (α=1.3) for all operational forecasts because the retrospective forecast analysis demonstrated the Ångström exponent has an insignificant effect on the DNI forecast accuracy.

5.2.2 Operational clear-sky conditions

Figure 7 shows a time series of forecast AOD at 550 nm and daily RMSE for day-ahead DNI predictions during the operational forecasting period. As with the retrospective forecasting period, gaps in the time series show cloudy days. Panel c) of Figure 7 shows the GEOS-5 forecast AOD at Tucson and the climatological AOD from the Tegen *et al.* (1997) compared to the observed AOD at Tucson. The climatological AOD cannot represent the inter-day variability, which, therefore, negatively impacts the DNI forecast.

We see two similar patterns in the time series of DNI errors between the operational and retrospective periods. First, the daily RMSE values for each forecasting configuration are of similar magnitude (~30 Wm^{-2}) for most days in both periods. Second, there are distinct groups of days where the forecast using GEOS-5 AOD outperforms the other forecasts. The NAM 18Z (with GEOS-5 forecast AOD) performs better than NAM 00Z (with climatology AOD), for example, 22nd-28th May and 17th-24th June at Solana and Tucson.

Table VI shows the statistics for the operational period during clear-sky conditions. The magnitude of the RMSEs at Solana decrease compared to the retrospective period and is of the same order as RMSEs reported at Tucson. The MBE in the operational period is negative but positive in the retrospective period, likely due to minor differences in the forecasting set up or the differing seasons. However, the MBE for each configuration relative to the others is the

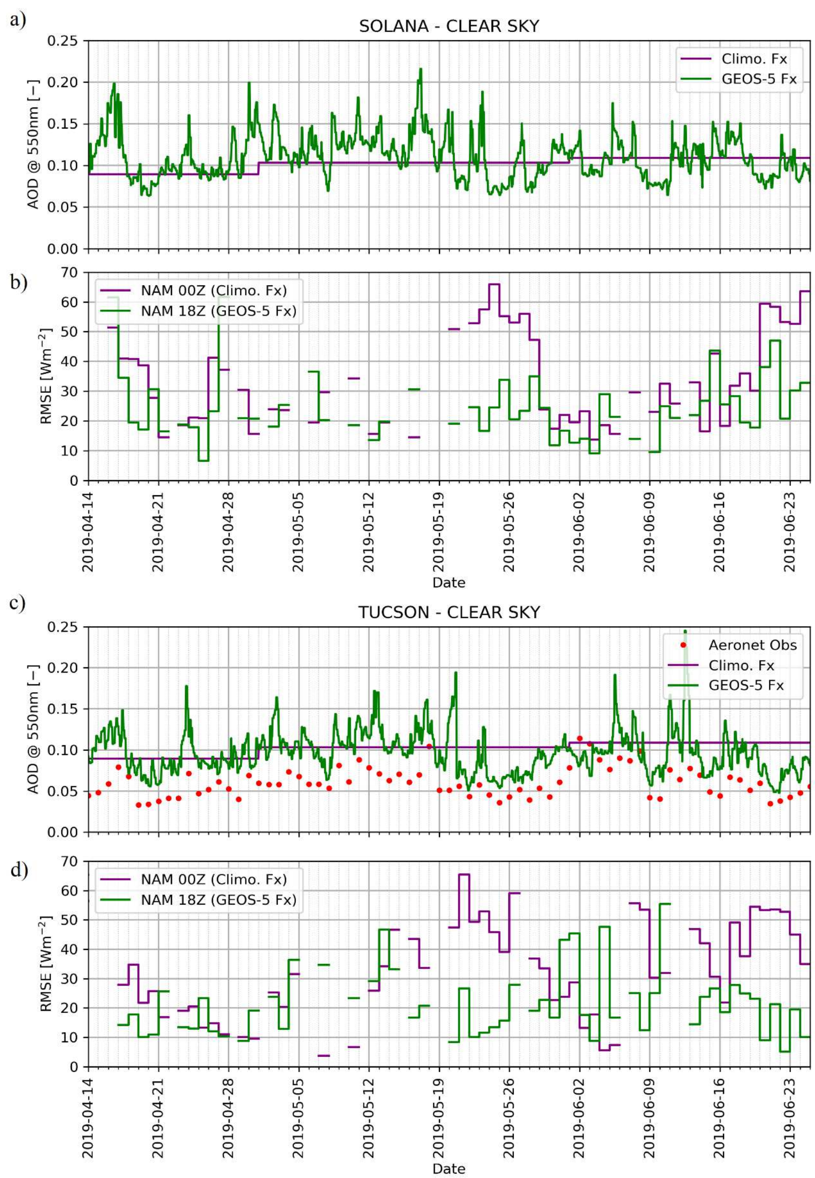


Figure 7: Time series of Aerosol Optical Depth (a, c) from AERONET observations (red), GEOS-5 forecasts (green), and Tegen aerosol climatology (purple) during the operational forecasting period. Data for Solana (a,b) and Tucson (c,d) are shown. Panels b) and d) show the daily RMSE value during clear-sky conditions for NAM 00Z (purple) and NAM 18Z (green), evaluated against observations at the Solana Generating Station and Tucson.

same in both forecasting periods. The negative bias in DNI forecasts using AOD climatology suggests it overestimates AOD, while the MBE for forecasts using GEOS-5 AOD is closer to zero. We can confirm this is true at Tucson by looking at Table II; GEOS-5 forecast AOD bias is lower in the operational forecast period compared to the Tegen AOD climatology.

Differences in performance are indistinguishable among the ensemble members using climatological AOD (NAM 00Z, 06Z, 12Z, and GFS 00Z, 12Z). The RMSEs are about 35 Wm^{-2} with negative biases of about 20 Wm^{-2}. The NAM 18Z member with GEOS-5 AOD forecasts shows an improvement to DNI forecast performance at Solana and Tucson. The RMSE and MBE are both reduced by about 10 Wm^{-2}. The clear-sky persistence DNI forecast performs better in the operational forecast period compared to the retrospective, with RMSEs less than ensemble members using climatological AOD. However, at both locations, DNI forecasts made with GEOS-5 forecast AOD perform better than the DNI persistence forecast.

For this section of the study, we calculate the skill score metric (SS_{NAM00Z}) with the NAM 00Z as the benchmark because it has the shortest forecast lead time. However, similar errors among all benchmark forecasts show this choice is not critical to the presented skill scores. The relative SS_{NAM00Z} for NAM 18Z at both locations is at least 0.27. This reduction in error is comparable to the findings in Jimenez *et al.*, (2016) (where they use GEOS-5 analysis AOD) provided we recalculate the Jimenez *et al.* skill score with respect to the Tegen climatology (SS_{Climo} = 1 – (41/66) = 0.38), instead of their reported skill score using 'No Aerosol'. We report no significant differences in statistical metrics for the NAM 18Z forecast at Solana versus Tucson. This is consistent with the retrospective period outside of its high AOD event. The remaining error (27 Wm^{-2}) approaches the limits of the combined radiation scheme error (4 Wm^{-2}) and observational error (20 Wm^{-2}), mentioned in Section 3.1 and Section 4.2.1.

Table VI: Statistics comparing different ensemble members from the operational forecasting system at the University of Arizona. Forecasts from members with different forcing models (GFS or NAM) at different times (00Z, 06Z, 12Z, 18Z) are evaluated with observations performed at Solana (left) and Tucson (right) during clear-sky conditions. Clear-sky DNI persistence is also shown for comparison. RMSE, MBE, and an RMSE-based skill score (SS_{NAM00Z}) relative to the NAM 00Z forecast is shown on each row.

Solana N=472	GFS 00Z	GFS 12Z	NAM 00Z	NAM 06Z	NAM 12Z	NAM 18Z	DNI Pers	Tucson N=447	GFS 00Z	GFS 12Z	NAM 00Z	NAM 06Z	NAM 12Z	NAM 18Z	DNI Pers
	Climo Fx					GEOS-5			Climo Fx					GEOS-5	
RMSE [Wm^{-2}]	36	37	37	36	37	27	31	RMSE [Wm^{-2}]	34	34	35	33	36	25	32
MBE [Wm^{-2}]	-22	-21	-21	-21	-22	-13	0	MBE [Wm^{-2}]	-22	-19	-21	-19	-21	-8	0
SS_{NAM00Z}	0.01	0.01	0	0.02	0	0.27	0.17	SS_{NAM00Z}	0.02	0.01	0	0.05	-0.03	0.29	0.07

Table VII: Same as Table V but for all-sky conditions and now with a 24-hour DNI persistence forecast for comparison.

Solana N=750	GFS 00Z	GFS 12Z	NAM 00Z	NAM 06Z	NAM 12Z	NAM 18Z	24 hr DNI Pers	Tucson N=750	GFS 00Z	GFS 12Z	NAM 00Z	NAM 06Z	NAM 12Z	NAM 18Z	24 hr DNI Pers
	Climo Fx					GEOS-5			Climo Fx					GEOS-5	
RMSE [Wm^{-2}]	219	240	214	189	223	216	349	RMSE [Wm^{-2}]	277	282	266	236	273	260	380
MBE [Wm^{-2}]	-22	-53	-17	-16	-32	-23	-7	MBE [Wm^{-2}]	-21	-32	-2	1	-38	0	-11
SS_{NAM00Z}	-0.02	-0.12	0	0.12	-0.04	-0.01	-0.63	SS_{NAM00Z}	-0.04	-0.06	0	0.11	-0.03	0.02	-0.43

5.2.3 Operational all-sky conditions

Table VII shows the same statistical metrics for each ensemble member but in all-sky AOD data (24-hour DNI persistence). With GEOS-5 AOD only influencing the radiation scheme of the model, we do not expect to improve cloudy-sky forecasts. The improved clear-sky performance is outweighed in these all-sky metrics by the relatively poor forecast performance during cloudy conditions.

6. Conclusions

In this study, we evaluate three different methods to include Aerosol Optical Depth (AOD) in operational direct normal irradiance (DNI) forecasts. One method (1: Climo Fx) uses the Tegen *et al.* (1997) global aerosol climatological data set. Another (2: Aeronet Fx) uses ground-based AOD observations from an AERONET site and implements those values as a 48-hour persistence of AOD uniformly over a forecast domain in the United States Desert Southwest. The last (3: GEOS-5 Fx) uses gridded GEOS-5 forecasts of AOD. We evaluate all methods at the Solana Generating Station, Gila Bend, AZ, and the University of Arizona, Tucson, AZ.

We perform a retrospective forecast analysis to assess the differences between these methodologies and a control forecast (0: No Aerosol). Including GEOS-5 forecast AOD reduced forecast DNI error during clear skies by at least 10% compared to when using the Tegen aerosol climatology. Negative biases in the DNI forecast result from using the Tegen *et al.* (1997) aerosol climatology data, and we see an average overestimation of AOD with respect to Tucson AERONET observations. Despite the simplicity of a 48-hour persistence of measured AOD (2: Aeronet Fx), this method yielded DNI forecast errors in Tucson (where AOD was measured) that were indistinguishable from DNI forecast errors when using the more complex GEOS-5 data. This result did not extend to the Solana Generating Station.

By contrasting forecasts made for the Solana Generating Station and Tucson, lower DNI errors suggest that GEOS-5 forecast AOD better captures the inter-day variability of AOD at Solana compared to using AERONET observations as a persistence AOD forecast. This inference is robust due to the *direct effect* of aerosols on radiation and the relatively simple relationship between total column AOD at 550 nm and DNI. These results suggest that GEOS-5 AOD forecasts are more representative of the AOD at Solana compared to using previous-day AERONET AOD from the Tucson site 225 km away.

Based on the results of the retrospective forecast period, we incorporated GEOS-5 AOD forecasts into an operational forecast system at the University of Arizona. We evaluated the existing operational configuration (1: Climo Fx) against this new configuration (3: GEOS-5 Fx). In clear-sky conditions, using GEOS-5 forecast AOD reduced DNI forecast RMSE by at least 27%. This reduction in error is comparable to the value reported in Jiménez *et al.* (2016) of 38%, where they use GEOS-5 analysis AOD. We recalculate the Jiménez *et al.* (2016) skill score to a more appropriate one, that uses RMSE from DNI forecasts using the Tegen climatology as the benchmark. The remaining DNI forecast error, $27\ \mathrm{Wm^{-2}}$, approaches the limits of the combined radiation scheme error and observational error, $24\ \mathrm{Wm^{-2}}$. No significant differences were found during all-sky conditions as the relatively poor performance during cloudy conditions outweighs the improvements made in clear-sky conditions.

Acknowledgments

This research was conducted at the University of Arizona and supported by Atlantica Yield and Arizona Public Service. We thank David Ovens at the University of Washington for guidance through processing GEOS-5 data for WRF, Antonio Lorenzo for maintaining archived operational University of Arizona WRF forecast data, and Hsin-I Chang for help with WRF troubleshooting. We extend gratitude Ave Arellano and Armin Sorooshain, for their comments on this manuscript. Direct questions and comments to the lead author at ptwbunn@email.arizona.edu.

Data Availability

The data from the Solana Generating Station that support the findings of this study were used with permission from Atlantica Yield. Restrictions apply to the availability of these data. Data from the Tucson site are available from the corresponding author upon reasonable request.

Supplementary Material

We include an animation of approximately 30 days of 2D gridded GEOS-5 AOD data, from which two snapshots are presented in Figure 3.

A sea-level pressure map for September 6th, 2017 (Storm Prediction Center, 2019) is included to give some synoptic context to the retrospective forecast period during the smoke events. NASA's EOSDIS world view has archived satellite detections of fire, available at https://worldview.earthdata.nasa.gov/.

A different version of Figure 5 shows the minimal differences in DNI forecasts when the Ångström exponent is held at a climatological value (α=1.3) versus using GEOS-5 gridded forecasts of Ångström exponent.

Finally, we show a bootstrap randomization analysis examine the statistical significance of the difference between the NAM 00Z (Climo Fx) DNI and the NAM 18Z (GEOS-5 Fx) DNI forecasts.

References

Andreas, A. Wilcox, S. (2010) 'Observed Atmospheric and Solar Information System (OASIS); Tucson, Arizona (Data)';, *NREL Report No. DA-5500-56494.* Available at: http://dx.doi.org/10.5439/1052226 .

ÅngstrÖm, A. (1961) 'Techniques of Determinig the Turbidity of the Atmosphere', *Tellus*, 13(2), pp. 214–223. doi: 10.3402/tellusa.v13i2.9493.

Antonanzas, J. *et al.* (2016) 'Review of photovoltaic power forecasting', *Solar Energy*. Elsevier Ltd, pp. 78–111. doi: 10.1016/j.solener.2016.06.069.

Brancucci Martinez-Anido, C. *et al.* (2016) 'ScienceDirect The value of day-ahead solar power forecasting improvement', *Solar Energy*. Elsevier Ltd, 129, pp. 192–203. doi: 10.1016/j.solener.2016.01.049.

Bunn, P. T. W. (2019) *Photograph of Solana Power Plant*, *Personal Photograph*. doi: 10.1104/pp.900114.

Chin, M. *et al.* (2000) 'Atmospheric sulfur cycle simulated in the global model GOCART ' Model description and global properties', 105.

Chin, M. *et al.* (2002) 'Tropospheric Aerosol Optical Thickness from the GOCART Model and Comparisons with Satellite and Sun Photometer Measurements', pp. 461–483.

Colarco, P. *et al.* (2010) 'Online simulations of global aerosol distributions in the NASA GEOS - 4 model and comparisons to satellite and ground - based aerosol optical depth', 115. doi: 10.1029/2009JD012820.

González, A. *et al.* (2013) 'Verification of precipitable water vapour in high-resolution WRF simulations over a mountainous archipelago', (October), pp. 2119–2133. doi:

10.1002/qj.2092.

Gueymard, C. A. and Ruiz-Arias, J. A. (2015) 'Validation of direct normal irradiance predictions under arid conditions: A review of radiative models and their turbidity-dependent performance', *Renewable and Sustainable Energy Reviews*. Elsevier, 45, pp. 379–396. doi: 10.1016/j.rser.2015.01.065.

Holben, B. N. *et al.* (1998) 'AERONET-A Federated Instrument Network and Data Archive for Aerosol Characterization', 16(February), pp. 1–16.

Iacono, M. J. *et al.* (2008) 'Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models', *Journal of Geophysical Research Atmospheres*, 113(13), pp. 2–9. doi: 10.1029/2008JD009944.

Jimenez, P. A. *et al.* (2016) 'WRF-SOLAR: Description and clear-sky assessment of an augmented NWP model for solar power prediction', *Bulletin of the American Meteorological Society*, 97(7), pp. 1249–1264. doi: 10.1175/BAMS-D-14-00279.1.

Kato, S. *et al.* (2000) 'A comparison of the aerosol thickness derived from ground-based and airborne measurements', 105.

Kleissl, J. (2013) *Solar energy forecasting and resource assessment*. Academic Press.

Levine, R. A. and Wilks, D. S. (2006) *Statistical Methods in the Atmospheric Sciences*, *Journal of the American Statistical Association*. doi: 10.2307/2669579.

NCEP (2015a) *NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive.*, *Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory.* Available at: https://doi.org/10.5065/D65D8PWK. (Accessed: 3 September 2019).

NCEP (2015b) *NCEP North American Mesoscale (NAM) 12 km Forecast & Analysis*, *Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO.* Available at: http://rda.ucar.edu/datasets/ds609.0/ (Accessed: 3 September 2019).

Pleim, J. E. (2007) 'A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing', *Journal of Applied Meteorology and Climatology*, 46(9), pp. 1383–1395. doi: 10.1175/jam2539.1.

Quaas, J. *et al.* (2009) 'Aerosol indirect effects â€" general circulation model intercomparison and evaluation with satellite data', *Atmospheric Chemistry and Physics*, 9(22), pp. 8697–8717. doi: 10.5194/acp-9-8697-2009.

Randles, C. A., Colarco, P. R. and Da Silva, A. (2013) 'Direct and semi-direct aerosol effects in the NASA GEOS-5 AGCM: Aerosol-climate interactions due to prognostic versus prescribed aerosols', *Journal of Geophysical Research Atmospheres*. Blackwell Publishing Ltd, 118(1), pp. 149–169. doi: 10.1029/2012JD018388.

Reale, O., Lau, K. M. and da Silva, A. (2011) 'Impact of Interactive Aerosol on the African Easterly Jet in the NASA GEOS-5 Global Forecasting System', *Weather and Forecasting*. American Meteorological Society, 26(4), pp. 504–519. doi: 10.1175/waf-d-10-05025.1.

Reno, M. J. and Hansen, C. W. (2016) 'Identification of Periods of Clear Sky Irradiance in Time Series of GHI Measurements', *Renewable Energy*, pp. 1–18.

Ruiz-Arias, J. A. *et al.* (2013) 'Assessment of the Level-3 MODIS daily aerosol optical depth in the context of surface solar radiation and numerical weather modeling', *Atmospheric Chemistry and Physics*, 13(2), pp. 675–692. doi: 10.5194/acp-13-675-2013.

Ruiz-Arias, J. A. (2013) ‘Surface clear-sky shortwave radiative closure intercomparisons in the Weather Research and Forecasting model’, *Journal of Geophysical Research Atmospheres*. Blackwell Publishing Ltd, 118(17), pp. 9901–9913. doi: 10.1002/jgrd.50778.

Ruiz-Arias, J. A. *et al.* (2015) ‘Do spaceborne aerosol observations limit the accuracy of modeled surface solar irradiance?’, *Geophysical Research Letters*. Blackwell Publishing Ltd, 42(2), pp. 605–612. doi: 10.1002/2014GL062309.

Ruiz-Arias, J. A., Dudhia, J. and Gueymard, C. A. (2014) ‘A simple parameterization of the short-wave aerosol optical properties for surface direct and diffuse irradiances assessment in a numerical weather model’, *Geoscientific Model Development*. Copernicus GmbH, 7(3), pp. 1159–1174. doi: 10.5194/gmd-7-1159-2014.

Schroedter-Homscheidt, M. *et al.* (2013) ‘Aerosols for Concentrating Solar Electricity Production Forecasts: Requirement Quantification and ECMWF/MACC Aerosol Forecast Assessment’, *Bulletin of the American Meteorological Society*, 94(6), pp. 903–914. doi: 10.1175/bams-d-11-00259.1.

Sengupta, M. *et al.* (2018) ‘The National Solar Radiation Data Base (NSRDB)’, *Renewable and Sustainable Energy Reviews*. Elsevier Ltd, 89(March 2018), pp. 51–60. doi: 10.1016/j.rser.2018.03.003.

Sessions, W. R. *et al.* (2015) ‘Development towards a global operational aerosol consensus: Basic climatological characteristics of the International Cooperative for Aerosol Prediction Multi-Model Ensemble (ICAP-MME)’, *Atmospheric Chemistry and Physics*. Copernicus GmbH, 15(1), pp. 335–362. doi: 10.5194/acp-15-335-2015.

Skamarock, W. C. *et al.* (2019) ‘A Description of the Advanced Research WRF Model Version 4 NCAR Technical Note’, *National Center for Atmospheric Research*, p. 145. doi:

10.5065/D6DZ069T.

Storm Prediction Center (2019) *Surface and Upper Air Maps*. Available at: https://www.spc.noaa.gov/obswx/maps/ (Accessed: 30 September 2019).

Suarez, Max J *et al.* (no date) *Technical Report Series on Global Modeling and Data Assimilation, Volume 27 The GEOS-5 Data Assimilation System-Documentation of Versions 5.0.1, 5.1.0, and 5.2.0*. Available at: http://www.sti.nasa.gov/STI-homepage.html.

Tegen, I. *et al.* (1997) 'Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results', *Journal of Geophysical Research Atmospheres*, 102(20), pp. 23895–23915.

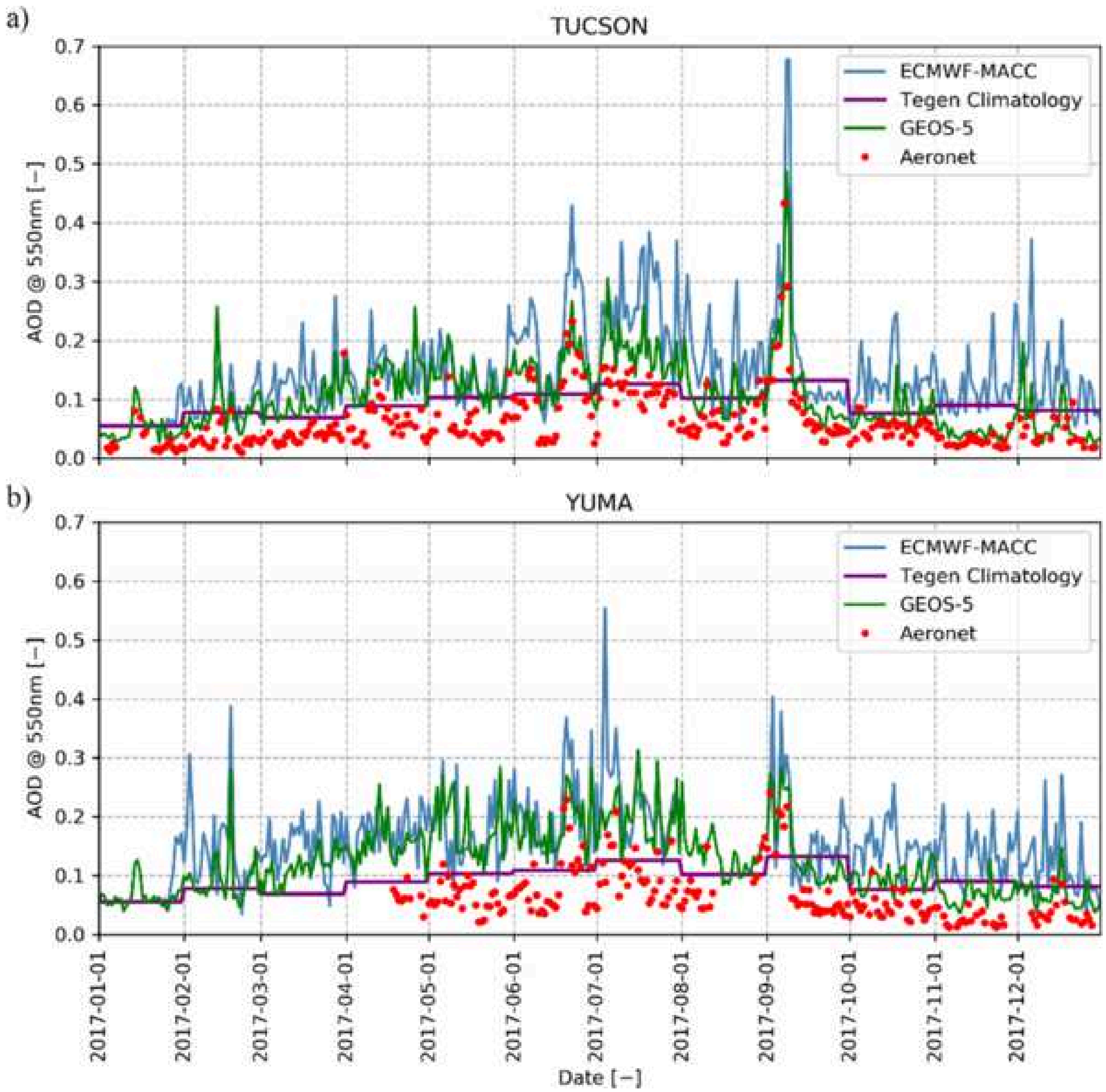
Tegen, I. and Fung, I. (1994) 'Modeling of mineral dust in the atmosphere: sources, transport, and optical thickness', *Journal of Geophysical Research*, 99(D11).

Thompson, G. *et al.* (2008) 'Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization', *Monthly Weather Review*, 136(12), pp. 5095–5115. doi: 10.1175/2008mwr2387.1.

Tuohy, A. *et al.* (2015) 'Solar Forecasting: Methods, Challenges, and Performance', *IEEE Power and Energy Magazine*. Institute of Electrical and Electronics Engineers Inc., 13(6), pp. 50–59. doi: 10.1109/MPE.2015.2461351.

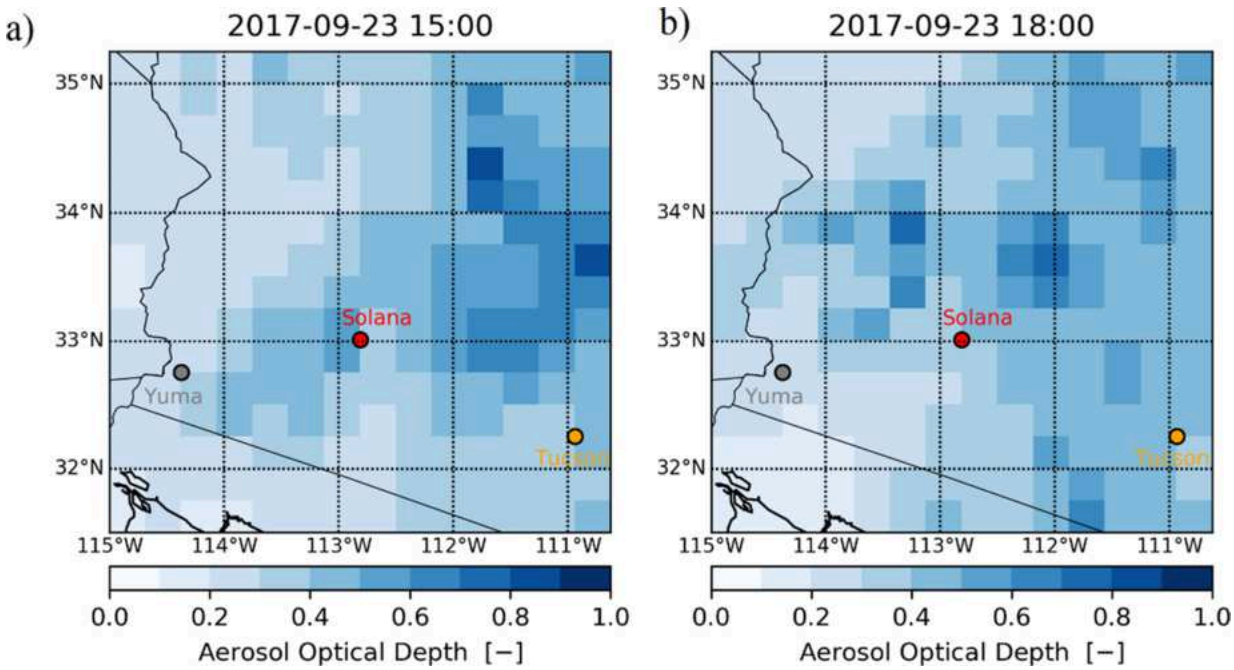
Zubler, E. M. *et al.* (2011) 'Intercomparison of aerosol climatologies for use in a regional climate model over Europe', *Geophysical Research Letters*, 38(15), pp. 1–5. doi: 10.1029/2011GL048081.

Background

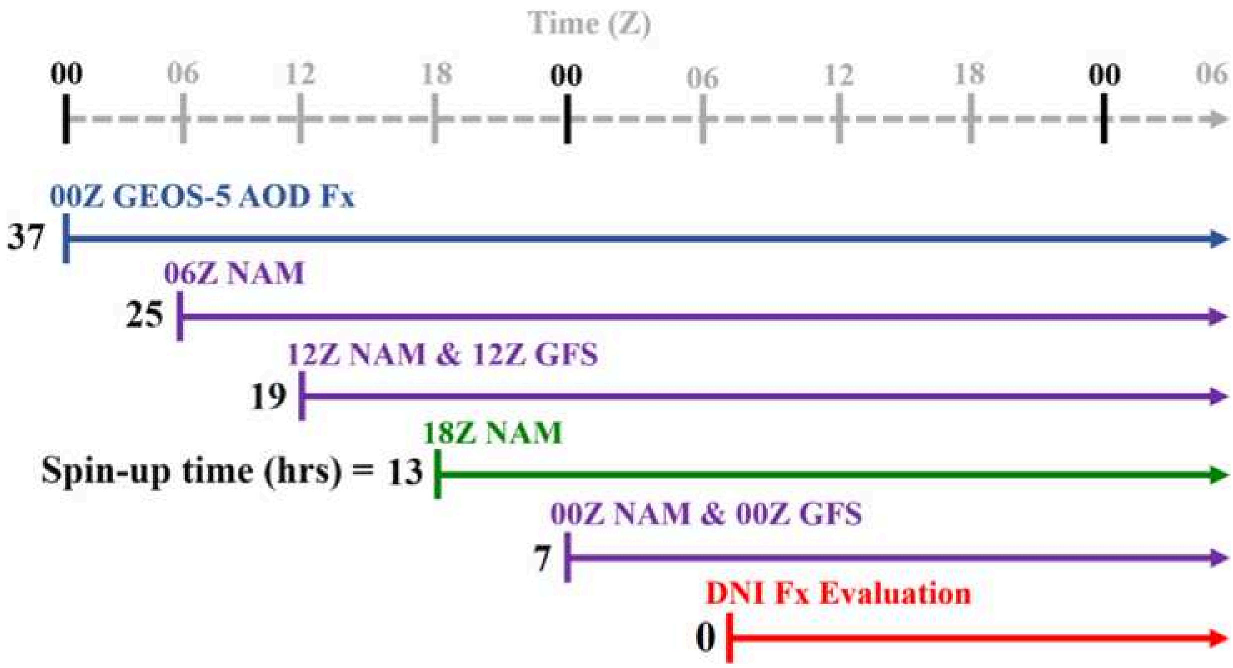

a)
TUCSON
ECMWF-MACC
Tegen Climatology
GEOS-5
Aeronet
AOD @ 550nm [−]
b)
YUMA
Date [−]
2017-01-01
2017-02-01
2017-03-01
2017-04-01
2017-05-01
2017-06-01
2017-07-01
2017-08-01
2017-09-01
2017-10-01
2017-11-01
2017-12-01

Background

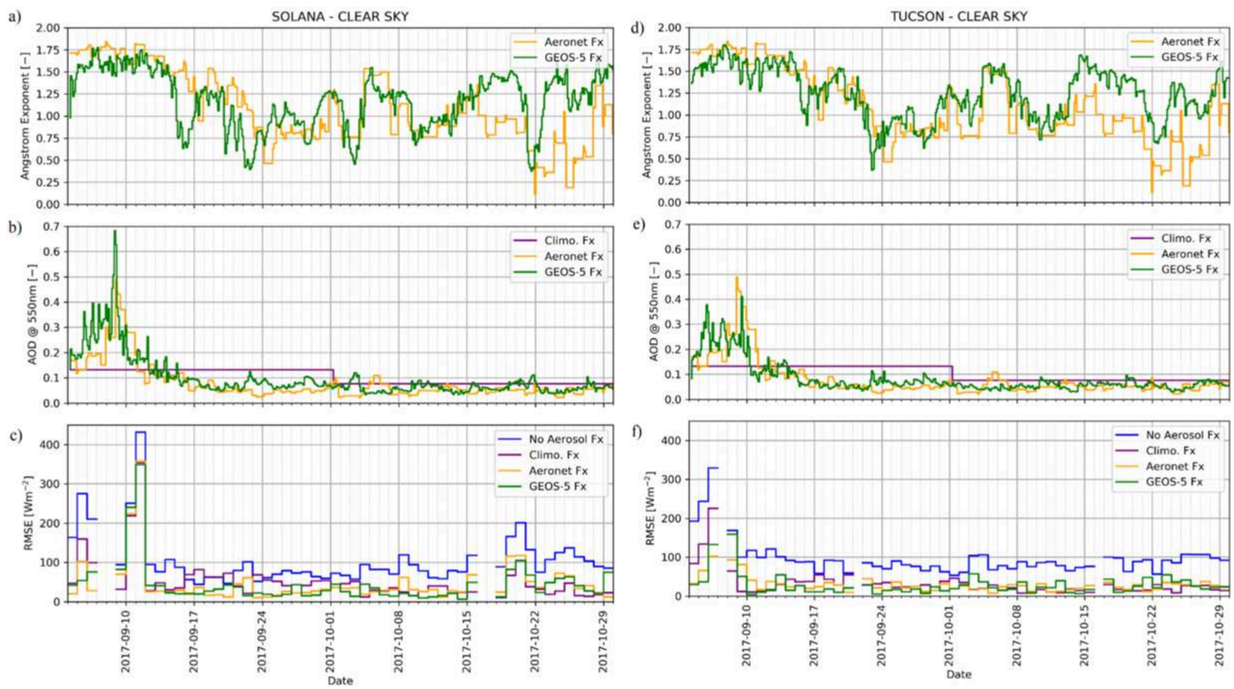
Background


a)
2017-09-23 15:00
b)
2017-09-23 18:00
35°N
34°N
33°N
32°N
115°W
114°W
113°W
112°W
111°W
Solana
Yuma
Tucson
0.0
0.2
0.4
0.6
0.8
1.0
Aerosol Optical Depth [−]

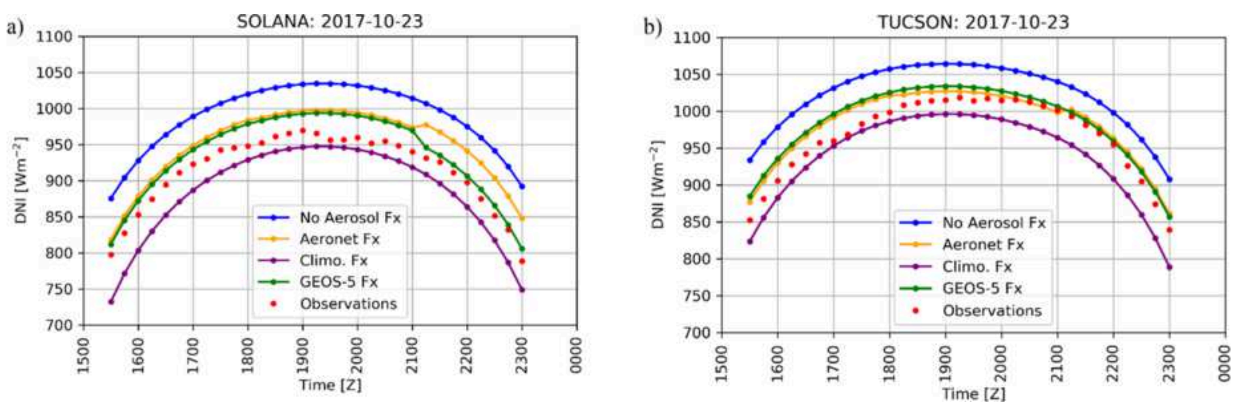
Background


Time (Z)
00
06
12
18
00
06
12
18
00
06
00Z GEOS-5 AOD Fx
37
06Z NAM
25
12Z NAM & 12Z GFS
19
18Z NAM
Spin-up time (hrs) = 13
00Z NAM & 00Z GFS
7
DNI Fx Evaluation
0

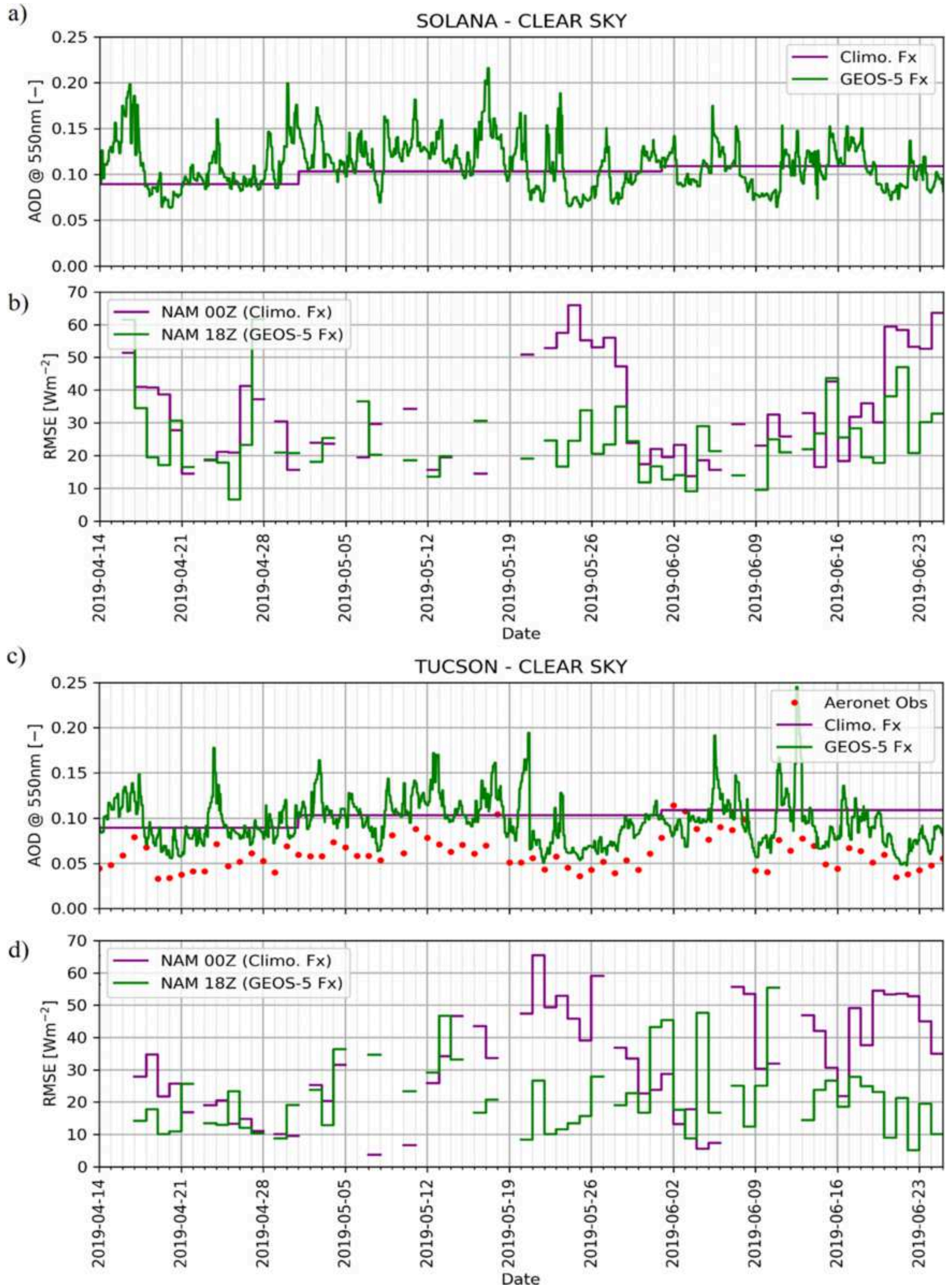
Background


a)
SOLANA - CLEAR SKY
Angstrom Exponent [−]
Aeronet Fx
GEOS-5 Fx
b)
AOD @ 550nm [−]
Climo. Fx
Aeronet Fx
GEOS-5 Fx
c)
RMSE [Wm⁻²]
No Aerosol Fx
Climo. Fx
Aeronet Fx
GEOS-5 Fx
Date
d)
TUCSON - CLEAR SKY
Angstrom Exponent [−]
Aeronet Fx
GEOS-5 Fx
e)
AOD @ 550nm [−]
Climo. Fx
Aeronet Fx
GEOS-5 Fx
f)
RMSE [Wm⁻²]
No Aerosol Fx
Climo. Fx
Aeronet Fx
GEOS-5 Fx
Date
2017-09-10
2017-09-17
2017-09-24
2017-10-01
2017-10-08
2017-10-15
2017-10-22
2017-10-29

Background


a)
SOLANA: 2017-10-23
DNI [Wm^{-2}]
Time [Z]
No Aerosol Fx
Aeronet Fx
Climo. Fx
GEOS-5 Fx
Observations
b)
TUCSON: 2017-10-23
DNI [Wm^{-2}]
Time [Z]
No Aerosol Fx
Aeronet Fx
Climo. Fx
GEOS-5 Fx
Observations

Background


a)
SOLANA - CLEAR SKY
AOD @ 550nm [−]
Climo. Fx
GEOS-5 Fx
b)
RMSE [Wm⁻²]
NAM 00Z (Climo. Fx)
NAM 18Z (GEOS-5 Fx)
2019-04-14
2019-04-21
2019-04-28
2019-05-05
2019-05-12
2019-05-19
2019-05-26
2019-06-02
2019-06-09
2019-06-16
2019-06-23
Date
c)
TUCSON - CLEAR SKY
AOD @ 550nm [−]
Aeronet Obs
Climo. Fx
GEOS-5 Fx
d)
RMSE [Wm⁻²]
NAM 00Z (Climo. Fx)
NAM 18Z (GEOS-5 Fx)
Date