JGR Atmospheres

RESEARCH ARTICLE
10.1029/2022JD036822

Influence of Aerosol Embedded in Shallow Cumulus Cloud Fields on the Surface Solar Irradiance

Jake J. Gristey[1,2], Graham Feingold[2], K. Sebastian Schmidt[3,4], and Hong Chen[3,4]

[1]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA, [2]NOAA Chemical Sciences Laboratory, Boulder, CO, USA, [3]Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO, USA, [4]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO, USA

Key Points:

- Aerosol embedded in shallow cumulus cloud fields significantly perturbs the distribution of surface solar irradiance
- Representative aerosol and cloud field properties accurately predict surface solar irradiance via machine learning
- Inference techniques enable process-level understanding of the aerosol influence on surface solar irradiance

Correspondence to:

J. J. Gristey,
Jake.J.Gristey@noaa.gov

Citation:

Gristey, J. J., Feingold, G., Schmidt, K. S., & Chen, H. (2022). Influence of aerosol embedded in shallow cumulus cloud fields on the surface solar irradiance. *Journal of Geophysical Research: Atmospheres*, *127*, e2022JD036822. https://doi.org/10.1029/2022JD036822

Received 23 MAR 2022
Accepted 27 MAY 2022

Abstract Ubiquitous shallow cumulus clouds are associated with complex variability in surface solar irradiance (SSI). Aerosol embedded in the cloud field typically has a much smaller overall radiative effect, but can significantly perturb the shape of the SSI probability density function (PDF). These perturbations have important implications for several applications that utilize SSI, but are poorly quantified and are the subject of this study. Multiple cases of shallow cumulus cloud fields with embedded aerosol are simulated at the Southern Great Plains Atmospheric Observatory using large eddy simulation (LES). The LES-derived cloud and aerosol fields are then ingested into Monte Carlo three-dimensional (3D) radiative transfer to simulate SSI. We find a variety of perturbations to the SSI PDF that depend on aerosol presence and optical properties. The processes leading to these perturbations include extinction of the direct beam that often increases from the clear-sky region toward cloud edge due to aerosol hygroscopic growth, and scattering of radiation by aerosol into cloud shadows. The ability to predict the SSI PDF in the presence of aerosol is assessed by adding three representative aerosol optical properties into an existing machine learning framework. We show that machine learning accurately predicts the SSI PDF across a wide range of conditions with negligible computational expense. Importance metrics reveal the relatively high influence of aerosol optical properties in making the predictions. These new findings highlight the important role that aerosol plays in SSI variability for highly 3D cloud-aerosol environments and provides a computationally efficient route forward for its simulation.

Plain Language Summary It is evident to anyone who has witnessed a cloud shadow that broken clouds cause large variations in the amount and distribution of sunlight reaching Earth's surface. However, the influence of small atmospheric particles, which always co-exist to some extent with clouds, is less apparent. This study demonstrates that the abundance and properties of these particles play a crucial role in determining how much sunlight reaches the surface in both the cloud shadows and the clear-sky between clouds. Accounting for the relevant processes is only possible with detailed calculations that track the exact path of sunlight through the Earth-atmosphere system, which is seldom done in atmospheric models routinely used throughout the world. By utilizing machine learning, we show that knowledge of just a few key properties of the atmosphere is sufficient to reproduce the results from the detailed calculations. Machine learning also provides insights into the most important drivers of sunlight reaching the surface, highlighting the importance of particles surrounding clouds, and therefore challenging traditional assumptions. The approach provides a route forward to predict sunlight reaching the surface in broken cloud conditions for solar renewable energy and many other applications.

1. Introduction

Electromagnetic radiation emitted by the sun, known as the solar irradiance, provides four orders of magnitude more energy to drive Earth-atmosphere processes than any other energy source (Kren et al., 2017). In the absence of cloud and aerosol, most of this incoming solar energy would be transmitted directly to Earth's surface. However, we live in a world where clouds occupy the sky around two-thirds of the time (King et al., 2013) and aerosol from both natural and anthropogenic sources is omnipresent. It is crucial to understand the sometimes complex influence of cloud and aerosol on the magnitude and variability of solar energy reaching the surface (the surface solar irradiance; SSI), given the numerous and profound implications for humanity. These implications include solar renewable energy production, photochemical reactions relevant to air pollution, crop yield, human health, and the surface energy budget that ultimately determines our weather and climate (Campillo et al., 2012; Ghiasvand et al., 2019; e.g., Perez et al., 2016; Peterson & Flowers, 1977; Wild, 2017).

Table 1
Surface Aerosol Extinction β_a, Single Scattering Albedo ω_a, and Asymmetry Parameter g_a, Nominally at 415 nm and 40% Relative Humidity, for the Different Simulation Setups on Each Day

	Scattering aerosol			Absorbing aerosol		
Date (YYYY-MM-DD)	β_a (km^{-1})	ω_a	g_a	β_a (km^{-1})	ω_a	g_a
2018-05-22	0.07	1.00	0.56	0.07	0.81	0.57
2018-06-06	0.10	1.00	0.58	0.10	0.80	0.59
2018-06-18	0.06	1.00	0.62	0.06	0.70	0.65
2018-06-19	0.04	1.00	0.60	0.04	0.71	0.62
2018-07-07	0.08	1.00	0.64	0.08	0.78	0.67
2018-07-09	0.14	1.00	0.61	0.14	0.81	0.62
2018-07-12	0.10	1.00	0.61	0.10	0.75	0.63
2018-07-31	0.04	1.00	0.63	0.05	0.82	0.65
2018-09-11	0.05	1.00	0.53	0.05	0.80	0.54
2018-09-14	0.07	1.00	0.59	0.07	0.79	0.61
2018-09-16	0.08	1.00	0.60	0.08	0.78	0.63
2018-09-17	0.12	1.00	0.61	0.13	0.83	0.61
2018-09-18	0.06	1.00	0.56	0.07	0.81	0.56
2018-10-02	0.07	1.00	0.56	0.07	0.80	0.57

A frequent but challenging environment for understanding the magnitude and variability of SSI is shallow cumulus cloud fields. These cloud fields exhibit detailed three-dimensional (3D) spatial structure, are characterized by small scales, and evolve quickly (Berg & Kassianov, 2008; Lamer & Kollias, 2015). Recent work has shown that the SSI variability in this environment is captured compactly by the shape of the SSI probability density function (PDF; Gristey et al., 2020b; Riihimaki et al., 2021; Schmidt et al., 2007, 2009); the PDF is typically bimodal, representing separately the cloud shadows and the [clear-sky] gaps between. The bimodal SSI PDF is routinely observed but can only be simulated by computationally expensive 3D radiative transfer, providing rare direct observational evidence for 3D radiative effects (Villefranque & Hogan, 2021). In simulations, the expense of 3D radiative transfer can be bypassed by predicting the precise bimodal PDF shape with machine learning algorithms using just a few key properties of the cloud field as predictors (Gristey et al., 2020a). These existing works mostly focused on the 3D radiative effect of shallow cumulus clouds, but there are indications from specific cases that aerosol embedded in the cloud field can significantly perturb the SSI PDF (Schmidt et al., 2009).

The purpose of the present study is twofold. First, we aim to quantify the influence of aerosol in the presence of shallow cumulus cloud fields on the SSI PDF. As shown in Section 3.1, despite having a relatively small overall radiative effect, the redistribution of solar irradiance via aerosol scattering plays an important role in determining the location, shape, and size of both modes of the SSI PDF. Second, we determine whether this aerosol influence can be captured via an existing machine learning framework (Gristey et al., 2020a) by adding a few key aerosol properties to the list of cloud field predictors. We demonstrate in Section 3.2 that representative values of aerosol optical depth, single scattering albedo, and asymmetry parameter provide sufficient predictive power to represent the perturbed SSI PDF associated with shallow cumulus clouds in the presence of aerosol.

The remainder of the manuscript is structured as follows. Section 2 describes large eddy simulation (LES) of cloud and aerosol variability, 3D radiative transfer to determine the corresponding SSI, and machine learning to reveal the relationships between them. Section 3 examines the influence of aerosol embedded in shallow cumulus clouds on the SSI PDF, and the ability to predict and understand the SSI PDF in the presence of aerosol. A summary and concluding remarks are provided in Section 4.

2. Data and Methodology

2.1. LES of LASSO 2018 Cases With Observationally Constrained Aerosol Variability

A set of 14 nonprecipitating days dispersed over the summer of 2018 that develop shallow cumulus cloud fields at the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) atmospheric observatory (36.6°N, 97.5°W) are selected for analysis (Table 1, column 1). These days are part of the LES ARM Symbiotic Simulation and Observation (LASSO) activity for continental shallow convection (Gustafson et al., 2020). For each day, LES is run using the System for Atmospheric Modeling (SAM; Khairoutdinov & Randall, 2003) with a domain size of 24 × 24 × 15 km. Horizontal grid spacing is 100 × 100 m and vertical grid spacing is 30 m below 5 km, incrementally stretched to 300 m by 10 km, continuing to 15 km. Simulations are initialized using observed atmospheric profiles from the 1200 UTC radiosonde launch at the SGP Central Facility on each day, and driven by prescribed surface fluxes and large scale forcing from the ARM constrained variational analysis (VARANAL) product representing a 300-km scale (Xie et al., 2004). Within SAM, 2-moment bulk microphysics (Morrison & Gettelman, 2008) and the Rapid Radiative Transfer Model with improved efficiency for general circulation model applications (Clough et al., 2005; RRTMG; Mlawer et al., 1997) are used. Note that while RRTMG is used online for efficiency in the LES and could influence the cloud field evolution (Jakub & Mayer, 2017; Veerman et al., 2020), we perform Monte Carlo 3D radiative transfer offline to obtain all SSI fields analyzed in this

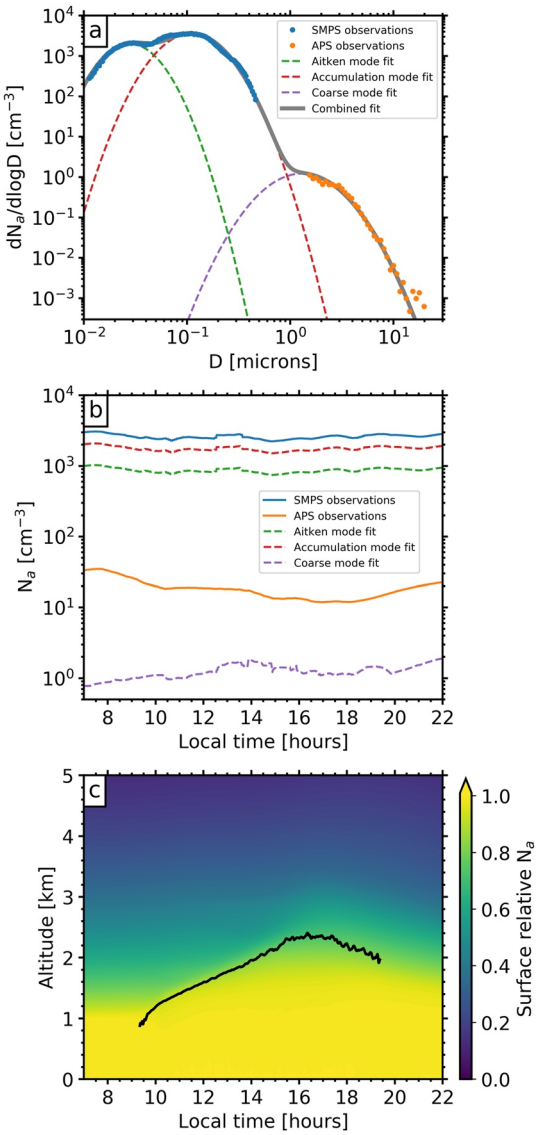

Figure 1. Observationally constrained aerosol variability implemented in large eddy simulation (LES) for an example case on 22 May 2018. (a) Aerosol size distribution from observations (dots) and fitted with a trimodal lognormal (gray line). Individual lognormal modes (corresponding to Equation 1) are shown with colored dash lines. (b) Diurnal evolution of aerosol number concentration in observations (solid lines) and for each fitted aerosol mode (dash lines). Note that the fitted modes in (a) overlap with one another and with both SMPS and APS size ranges, hence the offset between observed concentrations and each fitted mode. (c) Diurnal evolution of the vertical profile of surface-relative aerosol number concentration ($N_a(z)/N_a(z=0)$). The solid black line is the mean cloud top height for all columns with cloud optical depth greater than unity. SMPS is the Scanning Mobility Particle Sizer and APS is the Aerodynamic Particle Sizer.

study (see Section 2.2). Each simulation is run from 07:00 to 22:00 local time (Central Daylight Time or UTC − 5 hr) with a dynamical time step of half a second and a radiation time step of 1 min.

Given the focus of this study on aerosol, special attention is given to aerosol variability in our LES of LASSO 2018 cases. Following Glenn et al. (2020), aerosol size distributions are determined for each day using 07:00–08:00 local time averaged surface observations from the SGP Aerosol Observing System (AOS). Specifically, the Scanning Mobility Particle Sizer (SMPS; ARM, 2020b) and Aerodynamic Particle Sizer (APS; ARM, 2020a) provide the aerosol number size distribution in diameter ranges 0.01–0.5 and 0.5–20 μm, respectively. We fit multiple lognormal distributions to the observations, each with the functional form

$$f(D) = \frac{N_a}{\sqrt{2\pi}\ln(\sigma_g)\,D}\exp\left(-\frac{\ln^2(D/D_g)}{2\ln^2(\sigma_g)}\right) \quad (1)$$

where D is the particle diameter, N_a is the aerosol number concentration, D_g is the geometric mean diameter, and σ_g is the geometric standard deviation.

Rather than fit two lognormal modes to the observed size distribution as was done by Glenn et al. (2020), we found the observations for LASSO 2018 cases typically exhibit three distinct modes and therefore fit a trimodal lognormal to the aerosol size distribution on each day (e.g., Figure 1a). The fitted size distributions are then scaled during the course of our simulations to match the diurnal evolution of the observed total number concentration at the surface (also determined from the SMPS and APS) using a 1-hr running-mean at 1-min temporal resolution (e.g., Figure 1b). In the vertical, the aerosol concentration is assumed to be constant in an initial mixed layer and decay exponentially above with a scale height of 2 km (Turner et al., 2001; e.g., Figure 1c). The initial mixed layer depth was set to 1,000 m on each day by default, but adjusted to either 500 or 1,500 m if the observed radiosonde profile of potential temperature and specific humidity suggested otherwise. The evolution of aerosol in our LES is tracked with a numerical prognostic tracer that captures the 3D spatial distribution while being constrained by the observed diurnal evolution at the surface (further details in Glenn et al. (2020)). The parameters of each lognormal mode and their corresponding number concentration are passed to the activation scheme (Abdul-Razzak & Ghan, 2000) of the LES microphysics to represent aerosol-cloud interactions explicitly.

2.2. Monte Carlo 3D Radiative Transfer With Cloud and Aerosol

To simulate realistic SSI variability in shallow cumulus cloud fields, 3D radiative effects must be accounted for. To represent these 3D radiative effects accurately, we ingest the cloud and aerosol fields from the LES into the Education and Research 3D Radiative Transfer (EaR^{3}T) toolbox (Chen et al., 2022) that provides a Python-based interface to the Monte Carlo Atmospheric Radiative Transfer Simulator (MCARaTS; Iwabuchi, 2006). Specifically, we use 3D fields of cloud drop number concentration and effective radius, 3D fields of humidified aerosol optical properties for each size mode (details below), and a fixed surface albedo of 0.2 representative of the surface at SGP. The 3D radiative transfer simulations are performed for every tenth simulated minute (approximately the decorrelation time scale of cloud field (Kassianov et al., 2005)) that meets our "classic" shallow cumulus criteria. These criteria are identical to those defined by Gristey et al. (2020a): cloud fraction 5–40%, no significant cloud ice, >10 individual

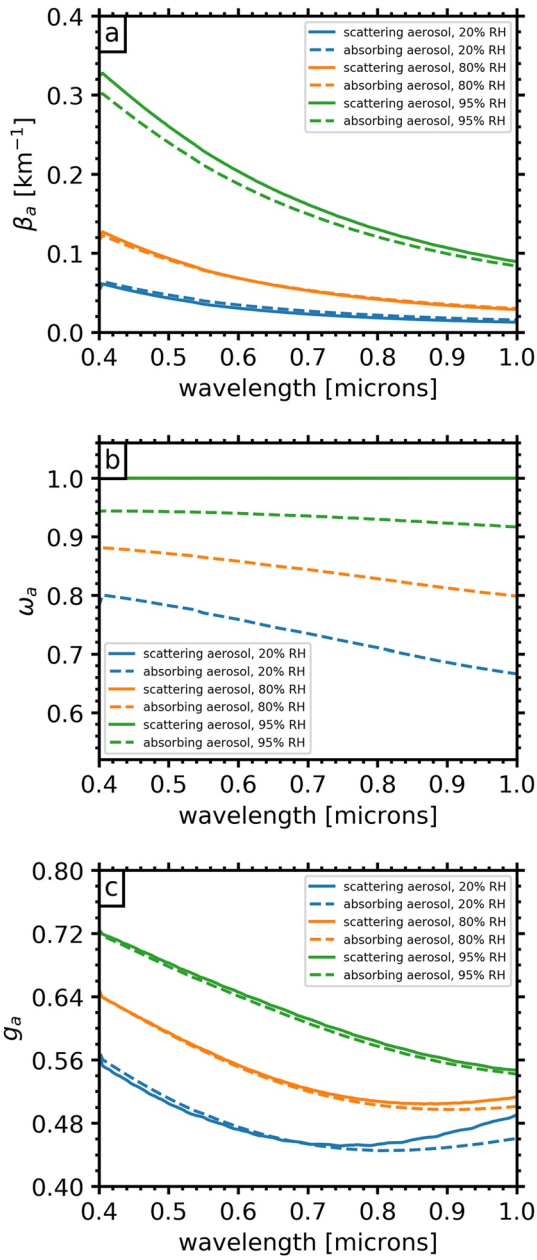

Figure 2. Surface aerosol (a) extinction β_a, (b) single scattering albedo ω_a, and (c) asymmetry parameter g_a as a function of wavelength for the simulated aerosol size distribution and number concentration averaged from 7 to 8 a.m. local time for an example case on 22 May 2018. Data are plotted at fixed relative humidity (RH) values (colors) and for scattering (solid lines) and absorbing (dash lines) aerosol. All results are based on Mie calculations.

clouds in the domain, mean cloud size <2 km, and solar zenith angle <40° to capture the afternoon evolution of shallow cumulus and the times when most energy is available at the surface. This results in 261 snapshots of cloud and aerosol fields that are used for 3D radiative transfer in this study. We also run EaR^{3}T with identical parameters as in Gristey et al. (2020a), but with the most recent version of EaR^{3}T that implements, among other relatively minor improvements, a Mie scattering phase function calculated by using Mie code from libRadtran (Emde et al., 2016; Mayer & Kylling, 2005).

Since aerosol optical and physical properties can be significantly different in the vicinity of clouds (Charlson et al., 2017; Koren et al., 2007; e.g., Marshak et al., 2021; Várnai et al., 2013), we first perform aerosol hygroscopic growth calculations for each aerosol size mode. This involves calculating the equilibrium particle sizes based on the 3D field of ambient relative humidity from the LES at each 3D output time step. We then use the optical properties of these swollen aerosol, as determined from Mie calculations, in our 3D radiative transfer. For the hygroscopic growth calculations, we assume a baseline chemical composition for all three aerosol size modes of soluble ammonium sulfate (i.e., highly hygroscopic and scattering aerosol). Given that aerosol with an absorbing component also occurs frequently at the SGP site, we repeat all calculations with a chemical composition of 10% insoluble soot by mass. This results in a single scattering albedo around 0.8 at 415 nm and 40% relative humidity (Table 1), which is at the lower end of observed values at SGP but not unrealistic (Zheng et al., 2020). We run our 3D radiative transfer separately for each snapshot with these highly scattering (henceforth "scattering aerosol") and partially absorbing (henceforth "absorbing aerosol") aerosol optical properties, therefore doubling our number of simulations to 522.

The aerosol hygroscopic growth calculations result in substantial differences in optical properties, as shown in Figure 2 for wavelengths from 0.4 to 1.0 μm where most of the energy resides. One obvious feature is the nonlinear increase in extinction with relative humidity; swollen particles are larger and therefore extinguish more radiation (Figure 2a). This humidification effect on extinction dominates the differences associated with scattering and absorbing aerosol (the minor offset between the two at high relative humidity is due to the reduced hygroscopic growth of the [soot-containing] absorbing aerosol). For the scattering aerosol, single scattering albedo is close to unity regardless of relative humidity (Figure 2b). For absorbing aerosol, however, increased particle water content at higher relative humidity significantly increases the single scattering albedo. The larger particles at increased relative humidity also typically lead to a larger asymmetry parameter at a given wavelength (Figure 2c). These differences highlight the importance of aerosol hygroscopic growth calculations for assessing aerosol radiative effects.

2.3. Machine Learning to Map Between Aerosol and Cloud Field Properties, and SSI

Machine learning offers an attractive approach to address the computational expense of 3D radiative transfer while simultaneously providing valuable physical insights via feature importance metrics. Gristey et al. (2020a) showed that two machine learning approaches, a random forest (RF) and an artificial neural network (ANN), can successfully map between a handful of shallow cumulus cloud field properties and the corresponding bimodal shape of the SSI PDF. The RF consists of a set of decision trees with each node in each decision tree making binary separations of data, whereas the ANN consists of a series of layers with nodes in each layer connected by weights. The two approaches

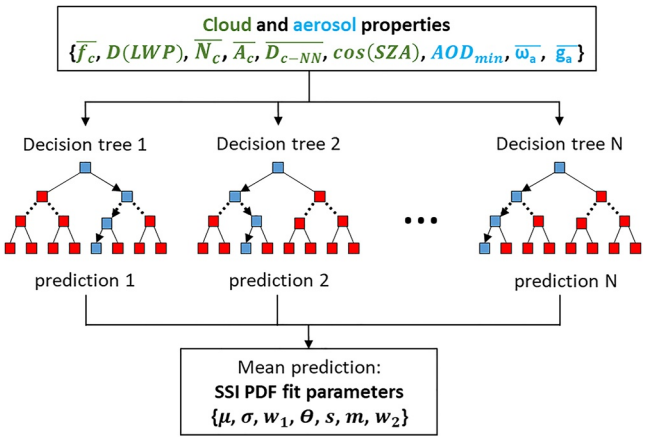

Figure 3. Schematic of random forest architecture employed in this study. Explanations of the input and output parameters are given in Table 2. Input parameters in green are cloud field properties and are identical to those used in Gristey et al. (2020a) whereas input parameters in cyan are aerosol properties added for the present study. Three dots represent extensions to the random forest (RF) architecture that are not depicted.

provided indistinguishable performance and utilized cloud field properties similarly, so we proceed only with the RF. Here, the RF is adapted to include aerosol properties in addition to cloud field properties (Figure 3).

Three representative values of aerosol properties are chosen to add to the list of features: the minimum aerosol optical depth at 500 nm for cloud-free columns $\boldsymbol{AOD_{min}}$, the mean extinction-weighted aerosol single scattering albedo at 500 nm for cloud-free grid boxes $\overline{\boldsymbol{\omega_a}}$, and the mean extinction-weighted aerosol asymmetry parameter at 500 nm for cloud-free grid boxes $\overline{\boldsymbol{g_a}}$. Unlike the cloud field properties, these aerosol properties are fundamentally radiative properties but are chosen because they are easily accessible from model output and are routinely observed. The $\boldsymbol{AOD_{min}}$ occurs in columns far from cloud and is chosen to be more consistent with observations at SGP that do not provide a retrieval in the vicinity of cloud. The full list of features, used as input for the RF, are largely uncorrelated and therefore each bring independent information to the prediction. They are given in Table 2, column 1.

The RF is used to predict the perturbed SSI PDF shape in the presence of aerosol. This PDF shape is quantified in a similar manner as described in Gristey et al. (2020a) so the reader is referred there for full details. A brief summary is provided again here for completeness. For the small irradiance mode, a two-parameter normal distribution is fitted to the data

$$f(x) = \frac{1}{\boldsymbol{\sigma}\sqrt{2\pi}} \exp\left(-\frac{(x-\boldsymbol{\mu})^2}{2\boldsymbol{\sigma}^2}\right) \tag{2}$$

where $\boldsymbol{\mu}$ is the location parameter (the mean of the distribution) and $\boldsymbol{\sigma}$ is the shape parameter (the standard deviation of the distribution). For the large SSI mode, a three-parameter lognormal distribution is fitted to the data

$$f(x) = \frac{1}{(x-\boldsymbol{\theta})\boldsymbol{s}\sqrt{2\pi}} \exp\left(-\frac{\ln^2\left((x-\boldsymbol{\theta})/\boldsymbol{m}\right)}{2\boldsymbol{s}^2}\right) \tag{3}$$

where $\boldsymbol{\theta}$ is the location parameter (a horizontal shift of the distribution), $\boldsymbol{s}$ is the shape parameter (the standard deviation of the log of the distribution), and $\boldsymbol{m}$ is the scale parameter (the median of the distribution). The only difference compared with Gristey et al. (2020a) is that here we fit the mirror image of the large SSI mode in the presence of aerosol. This is because the presence of aerosol causes the tail of the large SSI mode to extend in the opposite direction (see Section 3.1 and Figure 4b). As a result, the parameters of the large SSI mode have

Table 2
Random Forest Features (Inputs) and Labels (Outputs)

Cloud and aerosol properties (inputs)	SSI PDF fit parameters (outputs)
Mean cloud fraction: $\overline{\boldsymbol{f_C}}$ [5.2–34.0%]	Normal location parameter: $\boldsymbol{\mu}$
Dispersion in cloud liquid water path: ***D(LWP)*** [1.0–2.3]	Normal shape parameter: $\boldsymbol{\sigma}$
Mean in-cloud drop number concentration: $\overline{\boldsymbol{N_C}}$ [312–1,540 cm^{-1}]	Weight of small SSI mode: $\boldsymbol{w_1}$
Mean projected cloud area: $\overline{\boldsymbol{A_C}}$ [0.15–1.56 km^2]	Lognormal location parameter: $\boldsymbol{\theta}$
Mean distance to nearest cloud: $\overline{\boldsymbol{D_{C-NN}}}$ [0.80–1.23 km]	Lognormal shape parameter: $\boldsymbol{s}$
Cosine of solar zenith angle: ***cos(SZA)*** [0.77–0.97]	Lognormal scale parameter: $\boldsymbol{m}$
Minimum aerosol optical depth at 500 nm: $\boldsymbol{AOD_{min}}$ [0.06–0.91]	Weight of large SSI mode: $\boldsymbol{w_2}$
Mean aerosol single scattering albedo at 500 nm: $\overline{\boldsymbol{\omega_a}}$ [0.75–1.00]	
Mean aerosol asymmetry parameter at 500 nm: $\overline{\boldsymbol{g_a}}$ [0.50–0.66]	

Note. The range of values of each input is indicated in square brackets.

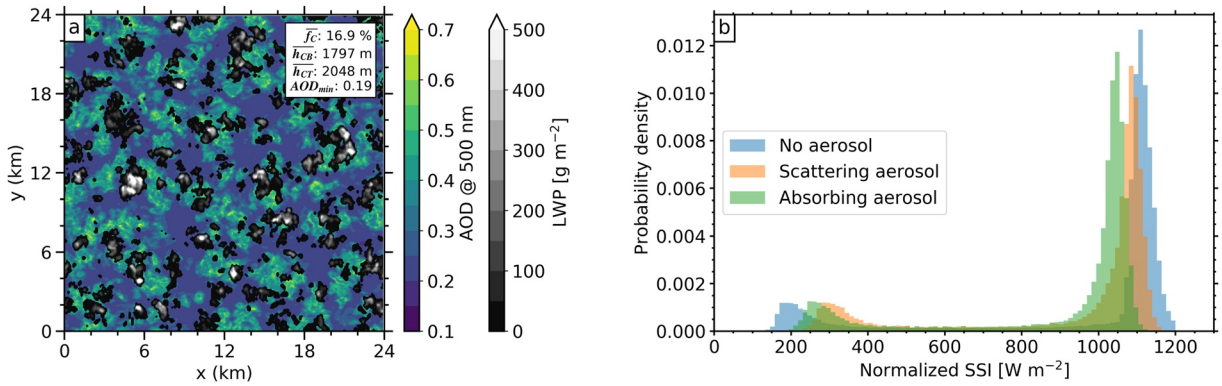

Figure 4. Simulated (a) spatial maps of cloud liquid water path (LWP; grayscale) and aerosol optical depth at 500 nm (AOD; colors), and (b) corresponding SSI PDF, for an example snapshot at 14:30 local time on 22 May 2018. The SSI PDFs are calculated for no aerosol (blue), scattering aerosol (orange) and absorbing aerosol (green). Also given in (a) are domain means of cloud fraction $\overline{f_C}$, cloud-base height $\overline{h_{CB}}$, and cloud-top height $\overline{h_{CT}}$, all defined based on a cloud optical depth of 1, and the domain minimum aerosol optical depth at 500 nm AOD_{min}.

a different physical meaning: θ is now the maximum SSI rather than the start of the large SSI mode, and s and m dictate the extent that the data extends to smaller SSI rather than larger SSI. The parameters of these fitted distributions from Equations 2 and 3, along with the weights w_1 and w_2 associated with each mode (given by the integral of the PDF under each mode), are referred to collectively as the SSI PDF fit parameters and are used as the RF labels. The full list of labels, used as output for the RF, are given in Table 2, column 2.

The data set is randomly split into 70% training data, 15% validation data, and 15% test data. The validation data are used to tune RF hyperparameters including the number of decision trees (780), the maximum depth of decision trees (50), the minimum number of samples required to split a node (2), and the minimum number of samples in a leaf node (1). Tuning is performed by, first, selecting 100 random combinations of hyperparameters from a wide range and assessing them using k-fold cross-validation with five folds. We then narrow the hyperparameter search around the most promising random combination, and repeat the k-fold cross-validation with every combination of hyperparameters in this narrower search space, and select the best combination. The trained RF associated with the best hyperparameters is, finally, used to make independent predictions on test data.

To infer the relative value of each input aerosol or cloud field property for the RF, we consider two importance metrics: impurity-importance and permutation-importance (Breiman, 2001). The impurity-importance quantifies how important a given feature is by calculating the mean reduction in mean-square-error (MSE) in the training data for questions asked about that feature. For example, all nodes of the random forest that are split using cloud fraction will determine the impurity-importance of cloud fraction based on how effectively those questions split the data. The permutation-importance works by randomly perturbing the features one-by-one, and assessing the degradation in the MSE of predictions. Only training data are used in the permutation-importance to ensure consistency with the impurity-importance, and degradations across all features are scaled to a total of 100%. We also use the impurity-importance to assess the relative importance of each feature separately for each label.

3. Results

3.1. Perturbations to the Shallow Cumulus SSI PDF in the Presence of Aerosol

3.1.1. Case Study: 14:30 Local Time on 22 May 2018

To demonstrate the various influences of aerosol embedded in shallow cumulus cloud fields on SSI, and to highlight the relevant processes at play, we first analyze a snapshot of a single case: 14:30 local time on 22 May 2018. This snapshot is associated with typical aerosol and cloud conditions among the wider data set: cloud fraction of 17% with numerous shallow clouds that each rarely exceed 2-km in horizontal extent, and a minimum aerosol optical depth of 0.19 (Figure 4a). The spatial variability in aerosol optical depth is mostly due to aerosol swelling in higher relative humidity regions, rather than increased number concentrations (not shown). The corresponding SSI PDF is bimodal (Figure 4b), which is due to the presence of the shallow cumulus clouds as revealed in previous studies. However, we see here that the precise location, shape, and size of each mode is perturbed by the

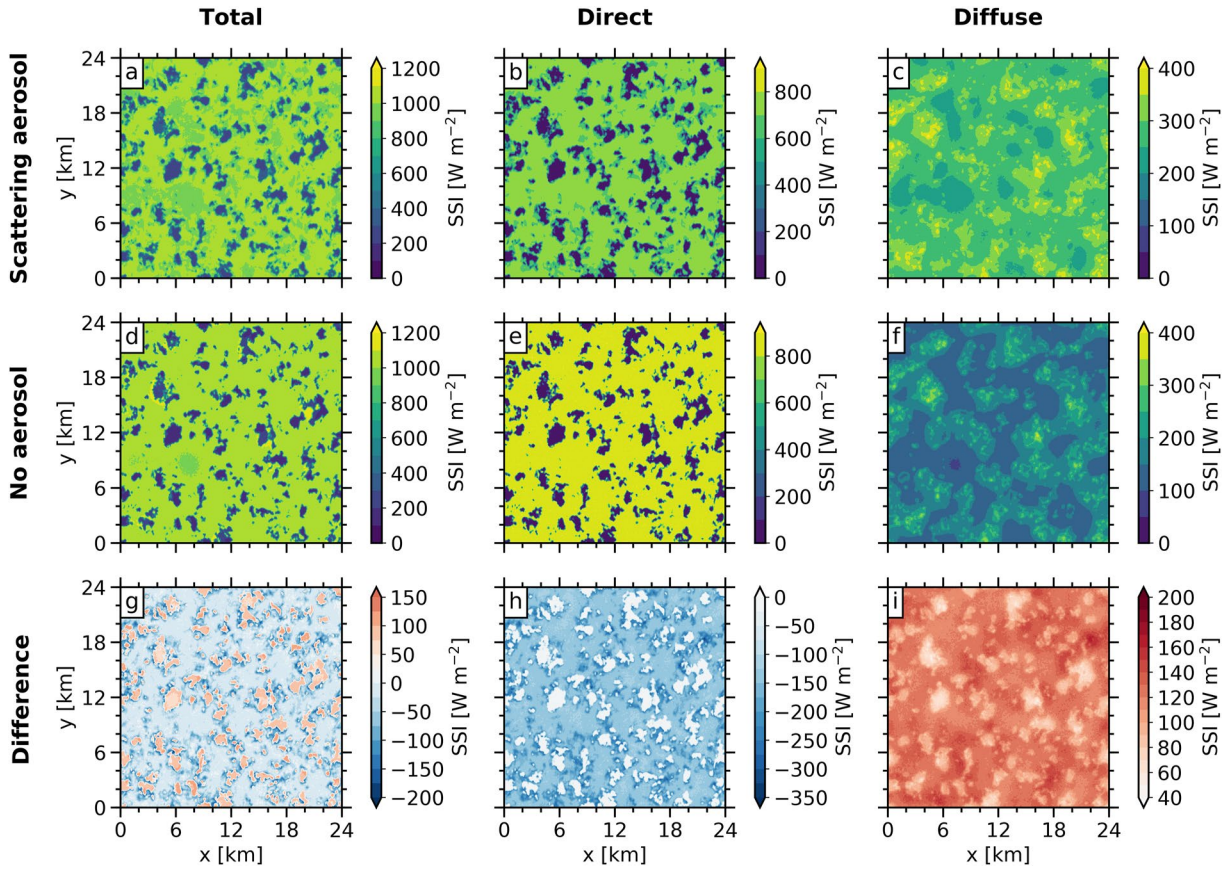

Figure 5. Simulated surface solar irradiance (SSI) maps for an example snapshot at 14:30 local time on 22 May 2018 when (a–c) including scattering aerosol, (d–f) neglecting aerosol, and (g–i) their difference. Data are shown for (a, d, g) total SSI, and then broken down separately into (d, e, h) direct SSI and (c, f, i) diffuse SSI components.

presence of aerosol, and the extent of this perturbation depends on the aerosol optical properties. There is a clear contraction of both SSI modes for our baseline 3D radiative transfer that ingests scattering aerosol. A similar shift occurs for the SSI PDF modes from simulations that instead use absorbing aerosol in the 3D radiative transfer. However, for absorbing aerosol, the shift in the large SSI mode is greater and the shift in the small SSI mode is less. This result is in line with observations and simulations from the Gulf of Mexico Atmospheric Composition and Climate Study (GoMACCS) of polluted shallow cumulus (Figure 1 in Schmidt et al. (2009)).

To provide an explanation for the SSI PDF differences seen in Figure 4b, we examine spatial maps of the SSI for each simulation, and separate the SSI into its direct and diffuse components. When comparing simulated direct SSI maps with scattering aerosol (Figure 5b) and without aerosol (Figure 5e), we see large differences in the regions between cloud shadows. This is due to the presence of aerosol between the clouds that extinguishes the direct beam. Looking closely at these differences in the direct SSI (Figure 5h), we see that the differences are not uniform but often increase approaching the edges of cloud shadows. This is due to the increase in relative humidity toward cloud edge that causes the aerosol to swell and therefore increase aerosol extinction. This aerosol humidification effect on SSI occurs in the same locations where the SSI is at a maximum in the absence of aerosol (Figure 5d), which is a 3D cloud radiative effect caused by the horizontal transport of photons scattered by clouds into the surrounding clear-sky regions. The result of the combination of these two radiative processes is that the humidification of aerosol toward cloud edge partially counteracts the 3D cloud radiative effect, and the maximum in total SSI occurs further away from the cloud shadow in the presence of aerosol (Figure 5a), rather than immediately outside the cloud shadow edge. There are also regions where the difference in direct SSI is large away from cloud shadows, which could be due to regions of enhanced humidity left behind by recently dissipated clouds.

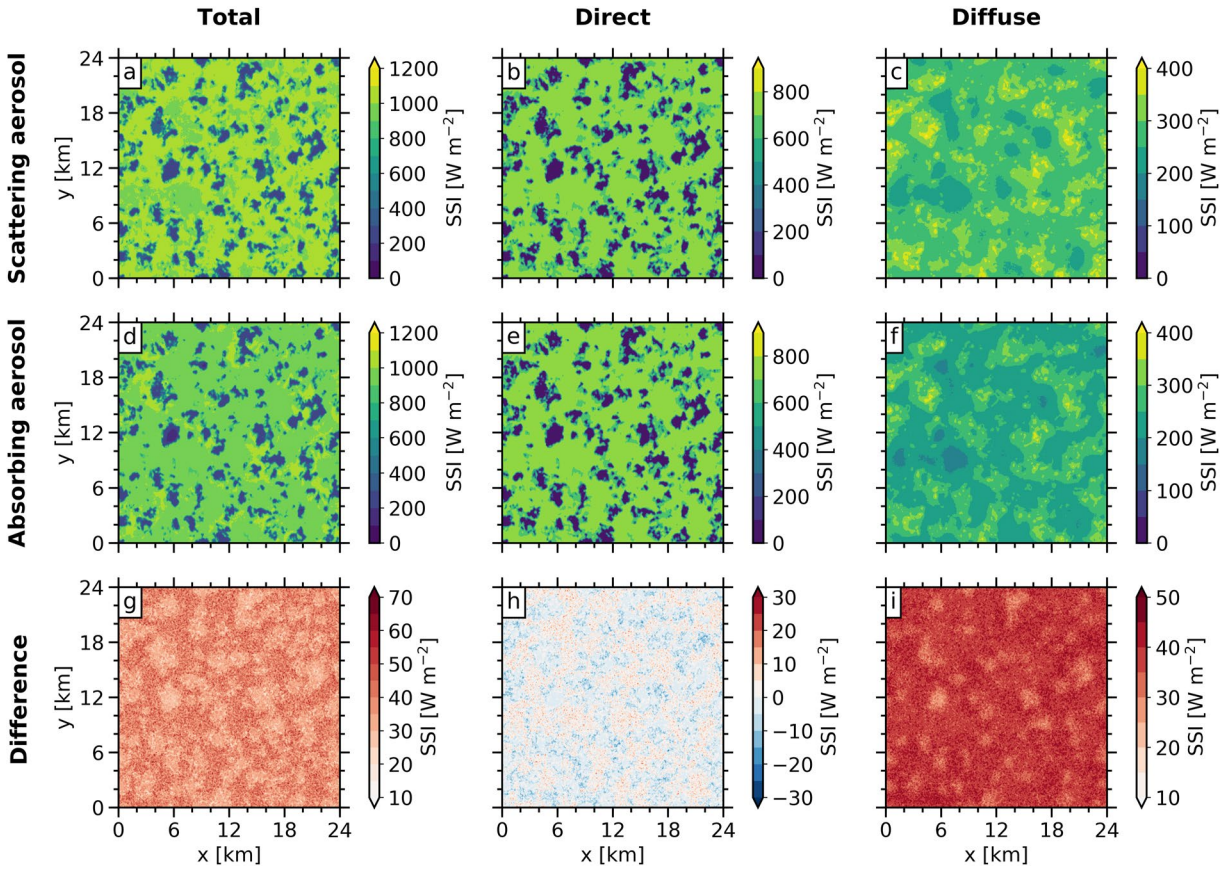

Figure 6. Simulated surface solar irradiance (SSI) maps for an example snapshot at 14:30 local time on 22 May 2018 when including (a–c) scattering aerosol, (d–f) absorbing aerosol, and (g–i) their difference. Data are shown for (a, d, g) total SSI, and then broken down separately into (d, e, h) direct SSI and (c, f, i) diffuse SSI components.

Most of the radiation extinguished by the scattering aerosol is scattered forwards and therefore toward the surface leading to an increase in the diffuse SSI (Figure 5c) compared to the simulation without aerosol (Figure 5f). The difference map (Figure 5i) shows that, although most of this diffuse SSI enhancement occurs between the cloud shadows, a substantial amount of energy is also scattered into the cloud shadows by the aerosol. This is because the aerosol phase function dictates that radiation is not always scattered *exactly* forwards, but often into other directions within the forward hemisphere. The diffuse SSI enhancement and direct SSI diminution, when combined (Figure 5g), results in cloud shadows becoming brighter (hence the shift in the small SSI mode to larger SSI) and the gaps between cloud shadows becoming darker (hence the shift in the large SSI mode to smaller SSI) in the presence of aerosol.

Spatial maps of direct SSI with scattering aerosol (Figure 6b) and absorbing aerosol (Figure 6e) are very similar. The difference between these fields (Figure 6h) shows physical patterns that are associated with the fact that the absorbing aerosol used here contains 10% insoluble soot. The insoluble soot causes the absorbing aerosol to exhibit reduced hygroscopic growth approaching cloud edge and therefore less extinction at a given relative humidity. On the other hand, spatial maps of diffuse SSI with scattering aerosol (Figure 6c) and absorbing aerosol (Figure 6f) show substantial differences. This is due to the reduced single scattering albedo of the absorbing aerosol meaning that less radiation is scattered toward the surface to appear as diffuse SSI. The difference field (Figure 6i) follows a similar pattern as the diffuse SSI difference between scattering aerosol and no aerosol (Figure 5i), but the cloud shadows do not stand out as clearly. This is because the smaller size of absorbing aerosol (from reduced hygroscopic growth) results in a less intense forward scattering peak, making it more likely for photons to be scattered not exactly forward but into the cloud shadows.

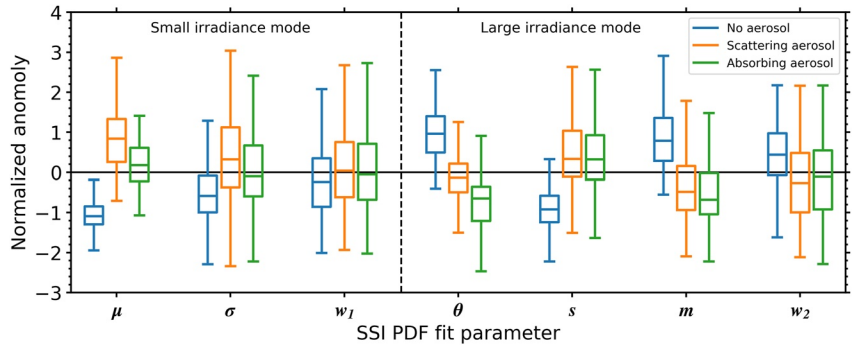

Figure 7. The distribution of surface solar irradiance (SSI) probability density function (PDF) fit parameters across all simulated cases for no aerosol (blue), scattering aerosol (orange), and absorbing aerosol (green). Normalized anomalies are calculated by subtracting the mean and dividing by the standard deviation for each parameter. Each box shows the median bounded by the interquartile range, with the whiskers extending to the last data point within an additional 1.5 × the interquartile range from the box edge.

3.1.2. Statistics Across All Simulations

To provide a broader perspective to the case study presented in the previous section, we now consider the SSI PDF in the presence of aerosol across all simulated snapshots. Figure 7 shows the differences in SSI PDF fit parameters for simulations with no aerosol, scattering aerosol, and absorbing aerosol, where the size of the boxes and whiskers represent the range of values across the simulations. The largest changes are seen in the location parameters $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$ which correspond to a translation along the horizontal axis of the SSI PDF modes. This is consistent with the case study (Figure 4b) where the small SSI mode shifts to larger SSI and the large SSI mode shifts to smaller SSI in the presence of aerosol. The shifts occur due to scattering of photons by aerosol into the cloud shadows and extinction of the direct beam between clouds, respectively. For the small SSI mode, the shift in $\boldsymbol{\mu}$ is more substantial for scattering aerosol than absorbing aerosol because the brightening of cloud shadows requires aerosol scattering. For the large SSI mode, the shift in $\boldsymbol{\theta}$ is more substantial for absorbing aerosol than scattering aerosol because less of the extinguished radiation is being scattered toward the surface. These effects are shown schematically by comparing Figures 8b and 8c.

Other SSI PDF fit parameters also show differences in Figure 7 with different aerosol configurations. For the small SSI mode, the standard deviation $\boldsymbol{\sigma}$ is increased in the presence of aerosol. This is because aerosol does not brighten the cloud shadows uniformly since, geometrically, it is more difficult to scatter radiation into the center of the cloud shadows than the edges. The increase in $\boldsymbol{\sigma}$ is larger for scattering aerosol because there is more brightening of the cloud shadow edge in general, but also because of the reduced hygroscopic growth of soot-containing aerosol resulting in a phase function where relatively more of the scattered radiation can reach the cloud shadow center. For the large SSI mode, the shape parameter $\boldsymbol{s}$ is increased and scale parameter $\boldsymbol{m}$ is decreased in the presence of aerosol. This corresponds to a sharper decline toward maximum SSI and a more pronounced tail toward smaller SSI, which is also consistent with that seen in the case study (Figure 4b). The sharper decline toward maximum SSI occurs due to hygroscopic growth of aerosol toward cloud edge, causing a reduction in SSI immediately outside the cloud shadow. This is also where the maximum SSI occurs without aerosol, and therefore a compensation occurs that erodes the tail toward maximum SSI (shown schematically by comparing Figures 8a and 8d). The pronounced tail toward smaller SSI arises from the inhomogeneity in AOD (shown schematically in Figure 8d), such that a consistent value of SSI far from clouds is not reached as readily. The weight parameters $\boldsymbol{w_1}$ and $\boldsymbol{w_2}$ are relatively insensitive to aerosol. This is because, although the aerosol can change the magnitude and variability of SSI within and between cloud shadows, the presence of aerosol does not change the fraction of the domain covered by cloud shadows.

Note that the SSI PDF fit parameters for simulations with no aerosol in Figure 7 are determined in an identical manner to those with aerosol. This provides a consistent definition of the parameters enabling a meaningful comparison. However, as noted in Section 2.3, the tail of the large SSI mode typically extends in the opposite direction with and without aerosol, meaning that the updated fit to the data used here does not work as well with no aerosol. If the goal were to provide the most accurate prediction in all cases, this could be addressed in future

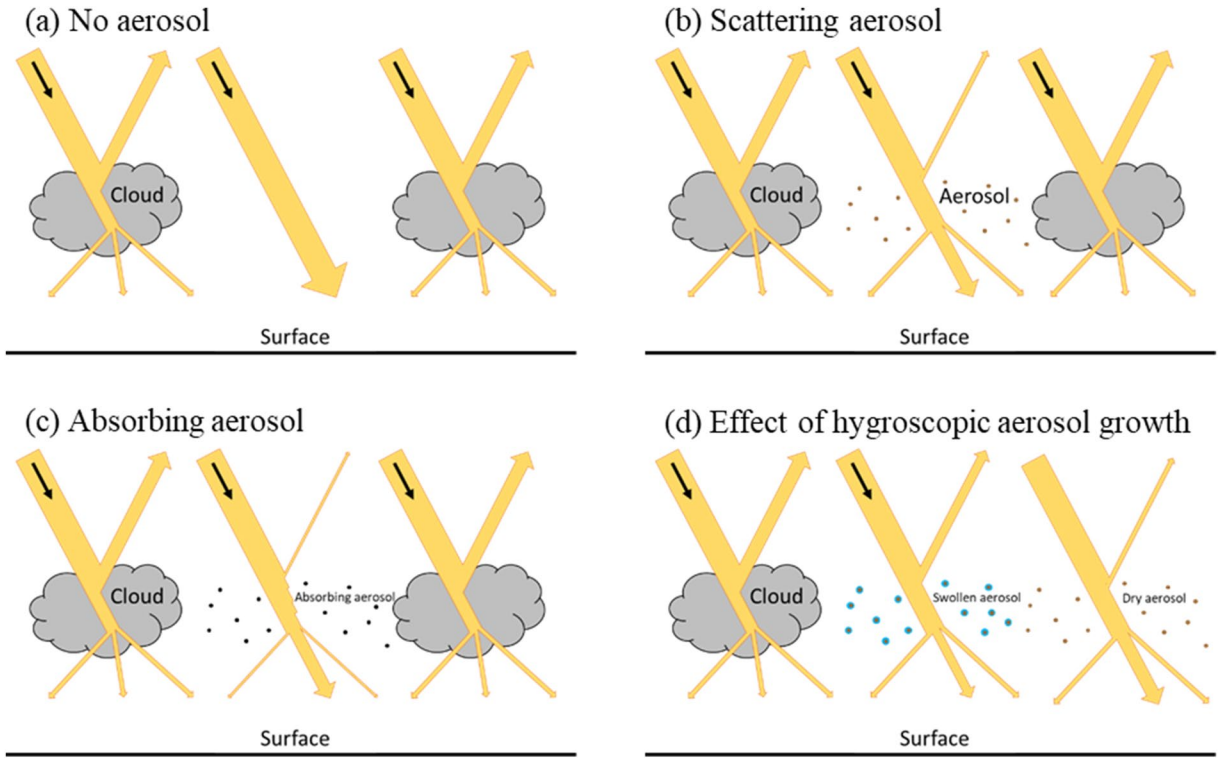

Figure 8. Schematic representation of 3D shortwave radiative processes of aerosol embedded in shallow cumulus cloud fields. The width of the yellow arrows indicates the relative magnitude of solar irradiance. Each panel shows (a) direct transmission of solar irradiance to the surface between clouds in the absence of aerosol, (b) scattering and redirection of solar irradiance for conservative scattering aerosol, (c) absorption in the aerosol layer (represented by the jagged edges), and (d) influence of hygroscopic aerosol growth in the vicinity of clouds. Gaseous scattering and absorption and surface reflection are ignored for the purpose of illustration.

work by exploring different fitting approaches based on a threshold for very clean atmospheres (i.e., introducing an initial decision of which trained RF to use). Given this caveat, and the purpose of this study, the simulations with no aerosol are not considered in the next section.

3.2. Machine Learning of the Relationship Between Shallow Cumulus Clouds and Aerosol, and the SSI PDF

3.2.1. Random Forest Prediction

In this section, the trained RF algorithm, described in Section 2.3 and depicted schematically in Figure 3, is evaluated on independent test data. The predictions for all SSI PDF fit parameters against their true values generally follow the 1:1 line indicating predictive skill (Figure 9). Upon closer inspection, some parameters are better predicted than others. The accuracy is highest for the location parameters (Figures 9a and 9d) and the weight of the large SSI mode (Figure 9g). This can be interpreted physically as a relatively good prediction of the mean SSI in cloud shadows, the maximum SSI in shadow-free regions, and the total amount of data in shadow-free regions, respectively. Accuracies are worse for the scale and shape parameters (Figures 9b, 9e, and 9f), and the weight of the small SSI mode (Figure 9c). This suggests that it is relatively difficult to predict the spread and total amount of data in the cloud shadows, and the distribution of SSI in cloud-free regions, respectively. We found no systematic association between the quality of individual predictions and any particular aerosol or cloud field property (not shown), but predictions could potentially be improved by further expanding the training data set.

Although examining the predictions for each parameter separately is useful for quantifying the performance of the predictions, the capability of the RF for predicting the SSI PDF is perhaps demonstrated best by using the predicted parameters to reconstruct the SSI PDF. Figure 10 shows six random examples of these reconstructed PDFs for test data that represent a range of aerosol and cloud conditions. Across these examples, the RF is able to capture variations in the location, shape, and size of the two SSI PDF modes in the presence of aerosol. While

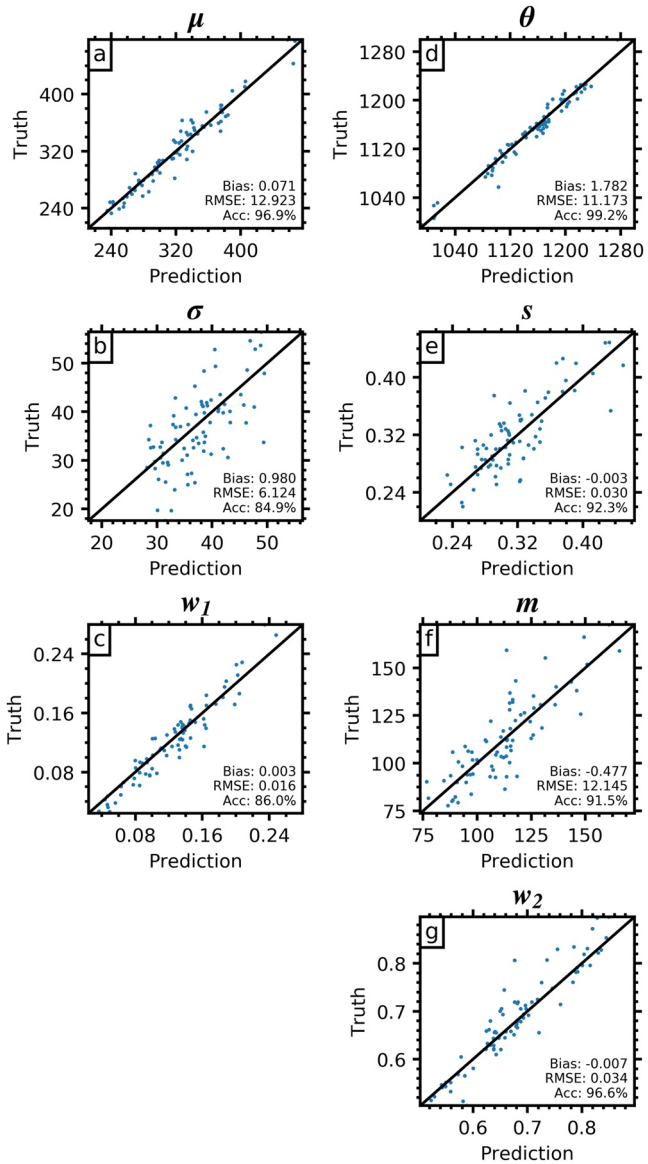

Figure 9. Predictions of surface solar irradiance (SSI) probability density function (PDF) fit parameters (defined in Table 2) for the (a–c) small SSI mode and (d–g) large SSI mode by the random forest on independent test data. RMSE is the root-mean-square-error and accuracy (Acc) is 100 minus the mean-absolute-percentage-error.

imperfections exist (e.g., the small SSI mode in Figure 10f), this is a drastic improvement compared with commonly used 1D radiation schemes that do not capture the bimodal SSI PDF shape at all (see Gristey et al. (2020b) Figure 4 or Schmidt et al. (2007) Figure 10b). In other words, the RF accurately predicts SSI in complex cloud-aerosol environments as determined from Monte Carlo 3D radiative transfer but with negligible computational expense.

3.2.2. Feature Importance

A valuable attribute of the RF approach is that it does not just make these predictions blindly, but it is also possible to extract the relative value of each input aerosol and cloud field property for making the predictions. Averaged over all of the output SSI PDF fit parameters, the most valuable predictor is the aerosol optical depth $\boldsymbol{AOD_{min}}$ (Table 3). Both the impurity and permutation importance metrics indicate that $\boldsymbol{AOD_{min}}$ provides just less than half of the overall predictive importance. This is somewhat surprising given that the bimodal shape of the SSI PDF is established by the shallow cumulus cloud field and does not require the presence of aerosol. The physical reason for this aerosol dominance is that, within the bounds of shallow cumulus where clouds are already optically thick, there is relatively little change in the transmitted radiation field due to small changes in the cloud properties. In contrast, small changes in the properties of optically thin aerosol have a relatively large impact on the transmitted radiation field via the processes shown schematically in Figure 8. Aerosols also exert their influence across a more extensive area (cloud fraction does not exceed 40% in the cases considered), can take on different composition, and undergo hygroscopic growth, all adding to the leverage of aerosol on SSI. The result suggests that, once the cloud field is established, the presence of aerosol plays a crucial role in determining how SSI is spatially distributed, which dominates over variations in the cloud properties. It follows that the aerosol optical properties $\overline{\omega_a}$ and $\overline{g_a}$ also show high importance given that they determine how much radiation is scattered toward the surface by aerosol and where it is scattered, respectively.

Among the cloud field properties, the cloud fraction $\overline{f_C}$ is the dominant controlling factor, consistent with findings in the absence of aerosol (Gristey et al., 2020a). In contrast, the dispersion in cloud liquid water path $\boldsymbol{D(LWP)}$ shows a much reduced importance in the presence of aerosol. Other cloud field properties also provide much less importance than $\overline{f_C}$, but note the standard deviation of the importance values from the permutation importance (Table 3, last column), which suggest that all inputs provide predictive value significantly greater than zero. When comparing the impurity and permutation importance metrics, the permutation importance tends to give more value to the less important inputs, but otherwise they are similar in terms of the relative order of input importance.

Further process information can be gleaned by assessing the value of each input aerosol and cloud field property for each output SSI PDF fit parameter separately (Table 4). The cloud fraction $\overline{f_C}$ shows highest importance for the weight parameters $\boldsymbol{w_1}$ and $\boldsymbol{w_2}$, which is expected given that the weight parameters control how much of the data is partitioned into cloud shadows and cloud-free regions. The dispersion in cloud liquid water path $\boldsymbol{D(LWP)}$ shows highest importance for the standard deviation of the small SSI mode $\boldsymbol{\sigma}$ due to the cloud liquid water path being an important control on how much radiation is ultimately scattered through the cloud and downward toward the surface to brighten the cloud shadow. Now we see that $\boldsymbol{D(LWP)}$ has an overall reduced importance in the presence of aerosol because aerosol plays a key role in brightening the cloud shadows (see Section 3.1). This can also be seen by examining the relationship between the mean of the small SSI mode $\boldsymbol{\mu}$ and it's most important input properties (Figure 11). The cloud fraction is the only cloud field property that shows

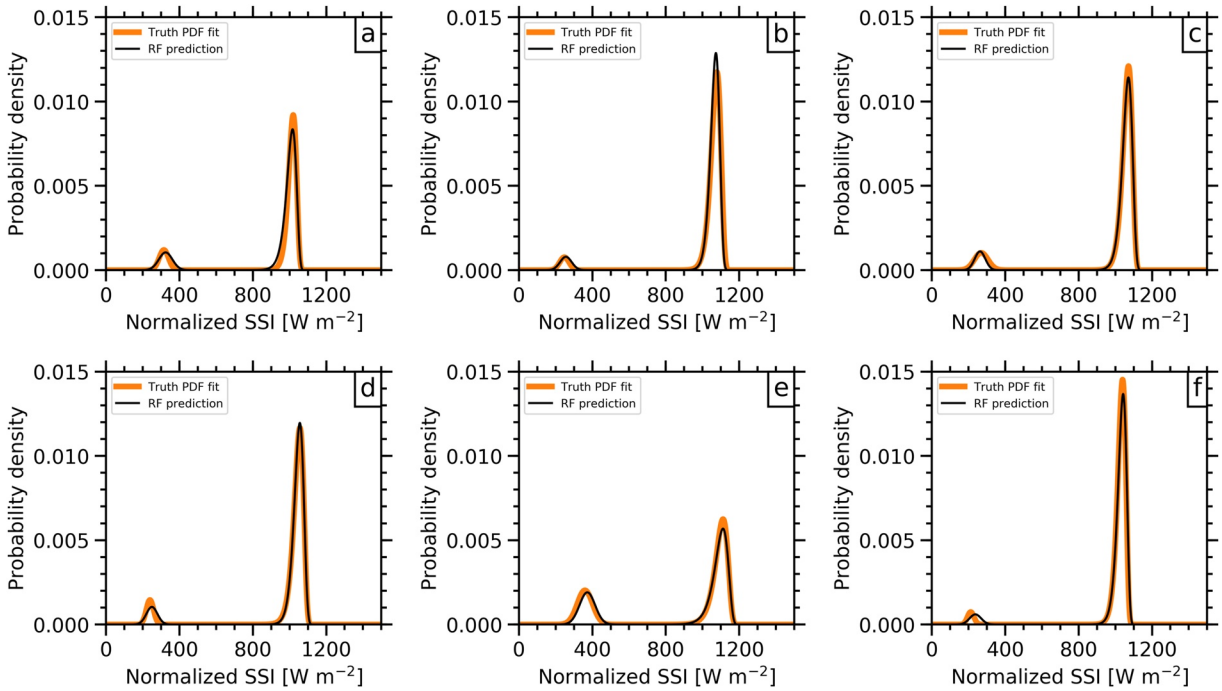

Figure 10. Reconstructed surface solar irradiance (SSI) probability density functions (PDFs) using predictions by the random forest (RF; black line) on independent test data. (a)–(f) show six random examples with the truth PDF fit (thick orange line) plotted for reference.

elevated importance for $\boldsymbol{\mu}$ but still shows a lot of scatter when plotted against $\boldsymbol{\mu}$ due to confounding variables (Figure 11a). Only when plotting against the most important variable, the domain minimum aerosol optical depth at 500 nm $\boldsymbol{AOD_{min}}$, does a functional form start to emerge (Figure 11b). The separate branches for scattering and absorbing aerosol demonstrate the only other input that shows large importance: the single scattering albedo $\overline{\boldsymbol{\omega_a}}$. This suggests, for the shallow cumulus cloud fields considered, that radiation scattered around individual clouds and back into their shadows via aerosol is a more important source of variability for the magnitude of SSI within cloud shadows than changes in the radiation transmitted through the cloud itself; a process that is entirely neglected by traditional 1D radiative transfer. Other cloud field properties and their relationship with individual PDF fit parameters also provide physically meaningful insights, including the distance between clouds $\overline{\boldsymbol{D_{C-NN}}}$ with $\boldsymbol{\sigma}$ (clouds close together can brighten the edges of each other's shadows via 3D radiative effects) and $\overline{\boldsymbol{A_C}}$ with $\boldsymbol{w_1}$ (wider clouds are also taller, meaning larger cloud shadows for the same cloud fraction when the sun is not directly overhead).

Table 3
The Relative Value of Each Input Aerosol and Cloud Field Property for the Trained RF

Input	Impurity-importance (%)	Permutation-importance (%)
$\overline{f_C}$	11.8	18.4 ± 1.2
$\boldsymbol{D(LWP)}$	1.2	1.4 ± 0.1
$\overline{N_C}$	1.8	2.8 ± 0.2
$\overline{A_C}$	2.8	3.8 ± 0.2
$\overline{D_{C-NN}}$	0.9	1.3 ± 0.1
$\boldsymbol{cos(SZA)}$	2.6	4.2 ± 0.3
$\boldsymbol{AOD_{min}}$	49.4	47.7 ± 2.0
$\overline{\omega_a}$	21.7	14.5 ± 1.0
$\overline{g_a}$	7.8	5.9 ± 0.3

Note. The permutation-based importance metric includes ±1 standard deviation from 100 random permutations.

The aerosol optical depth $\boldsymbol{AOD_{min}}$ shows highest importance for the location parameters $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$ and the shape parameter $\boldsymbol{s}$ of the large SSI mode. This supports the arguments made in Section 3.1 that the presence of aerosol primarily causes a shift in both SSI PDF modes and causes the tail of the large SSI mode to extend toward smaller irradiance. The mean aerosol single scattering albedo $\overline{\omega_a}$ also shows elevated importance for the location parameters, which occurs because the mean SSI in cloud shadows and the maximum SSI between cloud shadows are both dependent on the scattering of radiation to the surface by aerosol. The mean aerosol asymmetry parameter $\overline{g_a}$ shows highest importance for the scale parameter of the large SSI mode $\boldsymbol{m}$, which is a more nuanced effect relating to the smoothing of the SSI field near clouds. In this case, an increase in $\overline{g_a}$ causes an increase in the fraction of radiation scattered just outside the cloud shadows, changing the median of the large SSI mode. Overall, we find that assessing the relative importance of each input for each output separately provides a wealth of process-level

Table 4
Permutation Importance of Each Input for Each Output Separately (i.e., a Decomposition of the Final Column in Table 3)

	μ	σ	w_1	θ	s	m	w_2
$\overline{f_C}$	9.5 ± 0.9	20.4 ± 1.6	41.1 ± 3.1	10.5 ± 0.8	10.8 ± 0.9	17.7 ± 1.5	28.6 ± 2.1
$D(LWP)$	0.3 ± 0.0	3.1 ± 0.2	1.3 ± 0.1	0.2 ± 0.0	2.0 ± 0.1	1.3 ± 0.1	1.4 ± 0.1
$\overline{N_C}$	0.6 ± 0.0	3.7 ± 0.2	2.3 ± 0.2	0.4 ± 0.0	3.1 ± 0.2	4.2 ± 0.3	6.0 ± 0.6
$\overline{A_C}$	0.7 ± 0.1	4.9 ± 0.3	8.4 ± 0.7	1.3 ± 0.1	3.4 ± 0.2	4.6 ± 0.4	5.6 ± 0.4
$\overline{D_{C-NN}}$	0.2 ± 0.0	2.6 ± 0.1	1.0 ± 0.1	0.2 ± 0.0	1.9 ± 0.1	1.7 ± 0.2	1.4 ± 0.2
$cos(SZA)$	1.0 ± 0.1	6.0 ± 0.5	2.9 ± 0.3	0.5 ± 0.0	4.2 ± 0.3	10.0 ± 1.0	4.7 ± 0.5
AOD_{min}	62.7 ± 4.4	37.9 ± 2.8	28.3 ± 2.2	58.8 ± 4.3	60.9 ± 3.5	37.7 ± 2.7	39.1 ± 2.6
$\overline{\omega_a}$	22.8 ± 2.1	14.8 ± 1.4	8.6 ± 1.2	23.5 ± 2.0	7.6 ± 0.7	14.1 ± 1.8	5.5 ± 0.4
$\overline{g_a}$	2.3 ± 0.1	6.6 ± 0.3	6.1 ± 0.5	4.6 ± 0.3	6.2 ± 0.4	8.7 ± 0.8	7.6 ± 0.5

Note. For each input, darker shading indicates relatively higher importance.

information, and we advocate for this type of analysis in future studies that seek to utilize machine learning to understand physical processes.

4. Summary and Conclusions

Aerosol particles are frequently embedded in widespread shallow cumulus cloud fields, creating a complex 3D cloud-aerosol-radiation environment for understanding variability in SSI. This complexity is captured concisely by the shape of the SSI PDF, which is typically bimodal representing separately the cloud shadows and the cloud-free regions between. In this study, LES is used to produce observationally constrained 3D cloud and aerosol fields at the SGP Atmospheric Observatory that are ingested into 3D radiative transfer to investigate the role of aerosol in this problem. Although the aerosol has much lower optical depth than the adjacent clouds, we demonstrate the inordinately large leverage of the aerosol on the shape of the SSI PDF.

Focusing on a typical snapshot as a case study, 14:30 local time on 22 May 2018, we find that the SSI PDF is substantially perturbed in the presence of aerosol. The small SSI mode shifts to larger SSI as aerosol scatters radiation into cloud shadows, and the large SSI mode shifts to smaller SSI as aerosol attenuates the direct beam between clouds. More subtle changes to the SSI PDF shape are also identified such as a sharper edge at maximum SSI due to aerosol hygroscopic growth in the vicinity of clouds, and a longer tail in the large SSI mode toward smaller SSI due to the additional inhomogeneity introduced by aerosol in cloud-free regions. By fitting distributions to the SSI PDFs and examining changes to the parameters of those distributions, we find that the SSI PDF perturbations are consistent across hundreds of simulated cloud and aerosol fields, and the extent of the perturbations is strongly dependent on the aerosol optical properties.

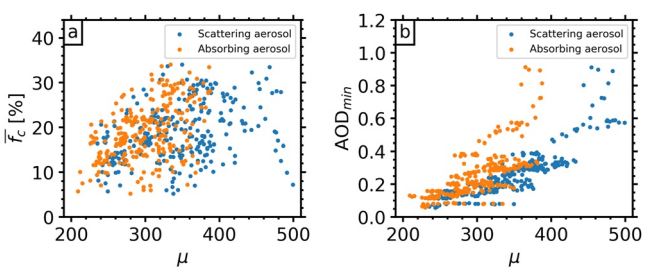

Figure 11. The dependence of the mean of the small SSI mode (μ) on (a) the domain-mean cloud fraction ($\overline{f_C}$), and (b) the domain minimum aerosol optical depth at 500 nm (AOD_{min}). Scattering aerosol (blue) and absorbing aerosol (orange) are plotted separately.

To provide an efficient calculation of SSI in this environment, and to facilitate further insights into the radiative processes involved, we develop a machine learning framework based on an existing RF algorithm that predicts the SSI PDF using a handful of key aerosol and cloud field properties. Representative values of aerosol optical depth, single scattering albedo, and asymmetry parameter are found to be sufficient to capture the perturbations to the SSI PDF in the presence of aerosol under a wide range of shallow cumulus conditions at the SGP site. The ML approach offers a drastic improvement in realism compared with commonly used 1D radiative transfer, while bypassing the computational expense of 3D radiative transfer. A major attribute

of the approach is that it enables the relative importance of aerosol and cloud field properties to be determined via feature importance metrics. These metrics reveal the crucial role that aerosol plays in determining the SSI variability, which extends beyond what might be expected from the relatively modest overall shortwave radiative effect of aerosol.

Our results highlight the consequence of neglecting aerosol in 3D shortwave radiative transfer involving shallow cumulus cloud fields and provide a route forward for efficient simulation. The RF approach developed has several potential applications such as the assessment of solar renewable energy potential by combining SSI PDFs with photovoltaic cell characteristics, and the parameterization of shortwave 3D radiative effects in LES. Any LES parameterization development will need to consider how to redistribute the SSI PDF over the domain, which is a key outstanding issue and could be important for feedbacks on the cloud field (Veerman et al., 2020). Given the demonstrated capabilities of the RF for aerosol embedded in shallow cumulus scenes at SGP, we recommend that future efforts focus on extending the range of aerosol and cloud conditions in the training data set to make the approach more widely applicable (we do not advise applying the model outside of the range of conditions on which it was trained). This could include other cloud regimes over different surfaces, with careful consideration of the properties needed to represent these different scenes.

Data Availability Statement

LASSO data (LASSO, 2022), the SAM LES model (SAM, 2022), the EaR^{3}T 3D radiative transfer package (EaR^{3}T, 2022), and the scikit-learn tool for the implementation of the random forest in Python (Scikit-learn, 2022) are all openly available.

Acknowledgments
This work was supported by the National Oceanic and Atmospheric Administration (NOAA) Atmospheric Science for Renewable Energy (ASRE) program, and by the Office of Biological and Environmental Research of the U.S. Department of Energy (DOE) Atmospheric System Research (ASR) Program under Interagency Award 89243020SSC000055. J.J.G. acknowledges support from the NOAA Cooperative Agreement with the Cooperative Institute for Research in Environmental Sciences (CIRES), NA17OAR4320101. The authors thank Ian Glenn, whose initial implementation of observationally constrained aerosol variability in SAM was a great help to the updated approach taken in this study, and three anonymous reviewers for their valuable suggestions that helped to improve this manuscript.

References

Abdul-Razzak, H., & Ghan, S. J. (2000). A parameterization of aerosol activation: 2. Multiple aerosol types. *Journal of Geophysical Research*, *105*(D5), 6837–6844. https://doi.org/10.1029/1999JD901161

ARM. (2020a). *Aerosol Observing System (AOS): Aerodynamic Particle Sizer (APS), Southern Great Plains (SGP)*. https://doi.org/10.5439/1407135

ARM. (2020b). *Aerosol Observing System (AOS): Scanning-Mobility Particle Sizer (SMPS), Southern Great Plains (SGP)*. https://doi.org/10.5439/1476898

Berg, L. K., & Kassianov, E. I. (2008). Temporal variability of fair-weather cumulus statistics at the ACRF SGP site. *Journal of Climate*, *21*(13), 3344–3358. https://doi.org/10.1175/2007JCLI2266.1

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Campillo, C., Fortes, R., & del Henar Prieto, M. (2012). Solar radiation effect on crop production. *Solar Radiation*, (p. 494). https://doi.org/10.5772/34796

Charlson, R. J., Ackerman, A. S., Bender, F., Anderson, T. L., & Liu, Z. (2017). On the climate forcing consequences of the albedo continuum between cloudy and clear air. *Tellus B: Chemical and Physical Meteorology*, *59*(4), 715–727. https://doi.org/10.1111/J.1600-0889.2007.00297.X

Chen, H., Schmidt, S., Massie, S. T., Nataraja, V. H., Norgren, M., Holz, R., et al. (2022). *The Education and Research 3D Radiative Transfer Toolbox (EaR^{3}T)—Towards the Mitigation of 3D Bias in Airborne and Spaceborne passive imagery cloud retrievals*. Under Review.

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., et al. (2005). Atmospheric radiative transfer modeling: A summary of the AER codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *91*(2), 233–244. https://doi.org/10.1016/J.JQSRT.2004.05.058

EaR3T. (2022). Education and Research 3D radiative transfer toolbox. [Software]. EaR3T. Retrieved from https://github.com/hong-chen/er3t

Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., et al. (2016). The libRadtran software package for radiative transfer calculations (version 2.0.1). *Geoscientific Model Development*, *9*, 1647–1672. https://doi.org/10.5194/gmd-9-1647-2016

Ghiasvand, R., Robsahm, T. E., Green, A. C., Rueegg, C. S., Weiderpass, E., Lund, E., & Veierød, M. B. (2019). Association of phenotypic characteristics and UV radiation exposure with risk of melanoma on different body sites. *JAMA Dermatology*, *155*(1), 39. https://doi.org/10.1001/jamadermatol.2018.3964

Glenn, I. B., Feingold, G., Gristey, J. J., & Yamaguchi, T. (2020). Quantification of the radiative effect of aerosol-cloud interactions in shallow continental cumulus clouds. *Journal of the Atmospheric Sciences*, *77*(8), 2905–2920. https://doi.org/10.1175/JAS-D-19-0269.1

Gristey, J. J., Feingold, G., Glenn, I. B., Schmidt, K. S., & Chen, H. (2020a). On the relationship between shallow cumulus cloud field properties and surface solar irradiance. *Geophysical Research Letters*, *47*, e2020GL090152. https://doi.org/10.1029/2020GL090152

Gristey, J. J., Feingold, G., Glenn, I. B., Schmidt, K. S., & Chen, H. (2020b). Surface solar irradiance in continental shallow cumulus fields: Observations and large-eddy simulation. *Journal of the Atmospheric Sciences*, *77*(3), 1065–1088. https://doi.org/10.1175/JAS-D-19-0261.1

Gustafson, W. I., Vogelmann, A. M., Li, Z., Cheng, X., Dumas, K. K., Endo, S., et al. (2020). The large-eddy simulation (LES) Atmospheric Radiation Measurement (ARM) Symbiotic simulation and observation (LASSO) activity for continental shallow convection. *Bulletin of the American Meteorological Society*, *101*(4), E462–E479. https://doi.org/10.1175/bams-d-19-0065.1

Iwabuchi, H. (2006). Efficient Monte Carlo methods for radiative transfer modeling. *Journal of the Atmospheric Sciences*, *63*(9), 2324–2339. https://doi.org/10.1175/JAS3755.1

Jakub, F., & Mayer, B. (2017). The role of 1-D and 3-D radiative heating in the organization of shallow cumulus convection and the formation of cloud streets. *Atmospheric Chemistry and Physics*, *17*, 13317–13327. https://doi.org/10.5194/acp-17-13317-2017

Kassianov, E., Long, C. N., & Ovtchinnikov, M. (2005). Cloud sky cover versus cloud fraction: Whole-sky simulations and observations. *Journal of Applied Meteorology*, *44*(1), 86–98. https://doi.org/10.1175/JAM-2184.1

Khairoutdinov, M. F., & Randall, D. A. (2003). Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities. *Journal of the Atmospheric Sciences*, *60*(4), 607–625. https://doi.org/10.1175/1520-0469(2003)060<0607:crmota>2.0.co;2

King, M. D., Platnick, S., Menzel, W. P., Ackerman, S. A., & Hubanks, P. A. (2013). Spatial and temporal distribution of clouds observed by MODIS onboard the terra and aqua satellites. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(7), 3826–3852. https://doi.org/10.1109/TGRS.2012.2227333

Koren, I., Remer, L. A., Kaufman, Y. J., Rudich, Y., & Martins, J. V. (2007). On the twilight zone between clouds and aerosols. *Geophysical Research Letters*, *34*, 8805. https://doi.org/10.1029/2007GL029253

Kren, A. C., Pilewskie, P., & Coddington, O. (2017). Where does Earth's atmosphere get its energy? *Journal of Space Weather and Space Climate*, *7*, A10. https://doi.org/10.1051/swsc/2017007

Lamer, K., & Kollias, P. (2015). Observations of fair-weather cumuli over land: Dynamical factors controlling cloud size and cover. *Geophysical Research Letters*, *42*, 8693–8701. https://doi.org/10.1002/2015GL064534

LASSO. (2022). LES ARM Symbiotic Simulation and Observation Bundle Browser. [Data Set]. LASSO. Retrieved from https://archive.arm.gov/lassobrowser

Marshak, A., Ackerman, A., Silva, A. D., Eck, T., Holben, B., Kahn, R., et al. (2021). Aerosol properties in cloudy environments from remote sensing observations: A review of the current state of knowledge. *Bulletin of the American Meteorological Society*, *102*(11), E2177–E2197. https://doi.org/10.1175/BAMS-D-20-0225.1

Mayer, B., & Kylling, A. (2005). Technical note: The libRadtran software package for radiative transfer calculations—Description and examples of use. *Atmospheric Chemistry and Physics*, *5*(7), 1855–1877. https://doi.org/10.5194/ACP-5-1855-2005

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research*, *102*(D14), 16663–16682. https://doi.org/10.1029/97JD00237

Morrison, H., & Gettelman, A. (2008). A new two-moment bulk stratiform cloud microphysics scheme in the community atmosphere model, version 3 (CAM3). Part I: Description and numerical tests. *Journal of Climate*, *21*(15), 3642–3659. https://doi.org/10.1175/2008JCLI2105.1

Perez, R., David, M., Hoff, T. E., Jamaly, M., Kivalov, S., Kleissl, J., et al. (2016). Spatial and temporal variability of solar energy. *Foundations and Trends in Renewable Energy*, *1*(1), 1–44. https://doi.org/10.1561/2700000006

Peterson, J. T., & Flowers, E. C. (1977). Interactions between air pollution and solar radiation. *Solar Energy*, *19*(1), 23–32. https://doi.org/10.1016/0038-092X(77)90085-8

Riihimaki, L. D., Flynn, C., McComiskey, A., Lubin, D., Blanchard, Y., Chiu, J. C., et al. (2021). The shortwave spectral radiometer for atmospheric science: Capabilities and applications from the ARM user facility. *Bulletin of the American Meteorological Society*, *102*(3), E539–E554. https://doi.org/10.1175/BAMS-D-19-0227.1

SAM. (2022). System for Atmospheric Modeling [software]. Retrieved from http://rossby.msrc.sunysb.edu/~marat/SAM.html

Schmidt, K. S., Feingold, G., Pilewskie, P., Jiang, H., Coddington, O., & Wendisch, M. (2009). Irradiance in polluted cumulus fields: Measured and modeled cloud-aerosol effects. *Geophysical Research Letters*, *36*, L07804. https://doi.org/10.1029/2008GL036848

Schmidt, K. S., Venema, V., Giuseppe, F., Di, Scheirer, R., Wendisch, M., & Pilewskie, P. (2007). Reproducing cloud microphysical and irradiance measurements using three 3D cloud generators. *Quarterly Journal of the Royal Meteorological Society*, *133*(624), 765–780. https://doi.org/10.1002/qj.53

Scikit-learn. (2022). Random Forest Regressor in Scikit-learn tool in Python. [Software]. Retrieved from https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html

Turner, D. D., Ferrare, R. A., & Brasseur, L. A. (2001). Average aerosol extinction and water vapor profiles over the Southern Great Plains. *Geophysical Research Letters*, *28*(23), 4441–4444. https://doi.org/10.1029/2001GL013691

Várnai, T., Marshak, A., & Yang, W. (2013). Multi-satellite aerosol observations in the vicinity of clouds. *Atmospheric Chemistry and Physics*, *13*(8), 3899–39808. https://doi.org/10.5194/ACP-13-3899-2013

Veerman, M. A., Pedruzo-Bagazgoitia, X., Jakub, F., Vilà-Guerau de Arellano, J., & Heerwaarden, C. C. (2020). Three-dimensional radiative effects by shallow cumulus clouds on dynamic heterogeneities over a vegetated surface. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001990. https://doi.org/10.1029/2019MS001990

Villefranque, N., & Hogan, R. J. (2021). Evidence for the 3D radiative effects of boundary-layer clouds from observations of direct and diffuse surface solar fluxes. *Geophysical Research Letters*, *48*, e2021GL093369. https://doi.org/10.1029/2021GL093369

Wild, M. (2017). Towards global estimates of the surface energy budget. *Current Climate Change Reports*, *3*(1), 87–97. https://doi.org/10.1007/s40641-017-0058-x

Xie, S., Cederwall, R. T., & Zhang, M. (2004). Developing long-term single-column model/cloud system-resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations. *Journal of Geophysical Research*, *109*, D01104. https://doi.org/10.1029/2003JD004045

Zheng, X., Xi, B., Dong, X., Logan, T., Wang, Y., & Wu, P. (2020). Investigation of aerosol-cloud interactions under different absorptive aerosol regimes using Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) ground-based measurements. *Atmospheric Chemistry and Physics*, *20*(6), 3483–3501. https://doi.org/10.5194/ACP-20-3483-2020