Developing Automated Methods to Estimate Spectrally Resolved Direct Normal Irradiance for Solar Energy Applications

Kelvin, Tsz Hei Choi[1, 2], Helen Brindley[1, 2], N. Ekins-Daukes[3], Rodrigo Escobar[4]

[1]Space and Atmospheric Physics Group, Department of Physics, Imperial College London, United Kingdom SW7 2AZ; [1]National Centre for Earth Observation, Imperial College London; [3]School of Photovoltaic & Renewable Energy Engineering, UNSW, Sydney, NSW 2052, Australia; [4]Centro de Energa, Centro del Desierto de Atacama, Escuela de Ingeniera, Pontificia Universidad Catlica de Chile Vicua Mackenna 4860, Macul, Santiago, Chile.

We describe four schemes designed to estimate spectrally resolved direct normal irradiance (DNI) for multi-junction concentrator photovoltaic systems applications. The schemes have increasing levels of complexity in terms of aerosol and circumsolar irradiance (CSI) treatment, ranging from a climatological aerosol classification with no account of CSI, to an approach which includes explicit aerosol typing and type dependent CSI contribution. When tested against ground-based broadband and spectral measurements at five sites spanning a range of aerosol conditions, the most sophisticated scheme yields an average bias of +0.068%, well within photometer calibration uncertainties. The average spread of error is 2.5%. These statistics are markedly better than the climatological approach, which carries an average bias of −1.76% and a spread of 4%. They also improve on an intermediate approach which uses Angström exponents to estimate the spectral variation in aerosol optical depth across the solar energy relevant wavelength domain. This approach results in systematic under and over-estimations of DNI at short and long wavelengths respectively. Incorporating spectral CSI particularly benefits sites which experience a significant amount of coarse aerosol. All approaches we describe use freely available reanalyses and software tools, and can be easily applied to alternative aerosol measurements, including those from satellite.

Keywords: Multi-Jucnction Solar Cell; Concentrator Photovoltaics; Aerosols; Circumsolar Irradiance; Spectral Direct Normal Irradiance; Solar Energy Resources

1 Motivation

Virtually all photovoltaic solar energy technology installed to date uses single-junction solar cells whose energy output closely follows the solar irradiance.[1] The output of highly efficient multi-junction (MJ) solar cells, however, depends on both the solar irradiance and the spectral distribution. Developed for use on space-craft, these MJ cells are also used in solar concentrator systems where the spectral variation of the incident irradiance affects the power output. In the longer term, it is likely that MJ cells, sometimes referred to as tandem cells, will also be used in unconcentrated flat-plate solar panels, with the International Roadmap for PV suggesting a 5% market share by 2030.[2] It is therefore important to establish reliable methods for estimating spectral irradiance to ensure accurate power prediction of future PV technologies.

When MJ cells are used within a concentrating photovoltaic system (CPV), as is currently the case in commercial solar farms, one of the critical quantities needed to enable power prediction is DNI_λ, the spectrally resolved solar radiance at ground, integrated over a small angle centred at the solar disc, and projected to the direct normal.[3] We emphasise the 'spectrally resolved' and 'integration over small angle' aspects of the definition, both of which come with their own challenges.

First, MJ cells are tested and their power rated under a reference DNI$_\lambda$ spectrum based on a standard atmosphere,[4] such as the AM1.5d ASTM G-173-03 direct spectrum[5] (figure 1). However, actual operating conditions can be quite different, and assuming a fixed spectrum and operating condition may substantially impact MJ CPV power production estimations, leading to unrepresentative evaluations of energy yields. This is because MJ CPV is not only sensitive to the broadband DNI, but also its spectral variation due to the need for current matching across the MJ stack[6] (figure 1). Once geometric parameters such as the airmass factor have been specified, and with proper cloud screening in place, the most critical factors to consider include ozone and water vapour absorption, and aerosol extinction,[7] all of which can substantially modify the spectral shape of DNI$_\lambda$. Previous studies have established that using spectra without considering these atmospheric parameters and their variability can lead to inaccurate energy yield estimates for MJ CPV.[8–16]

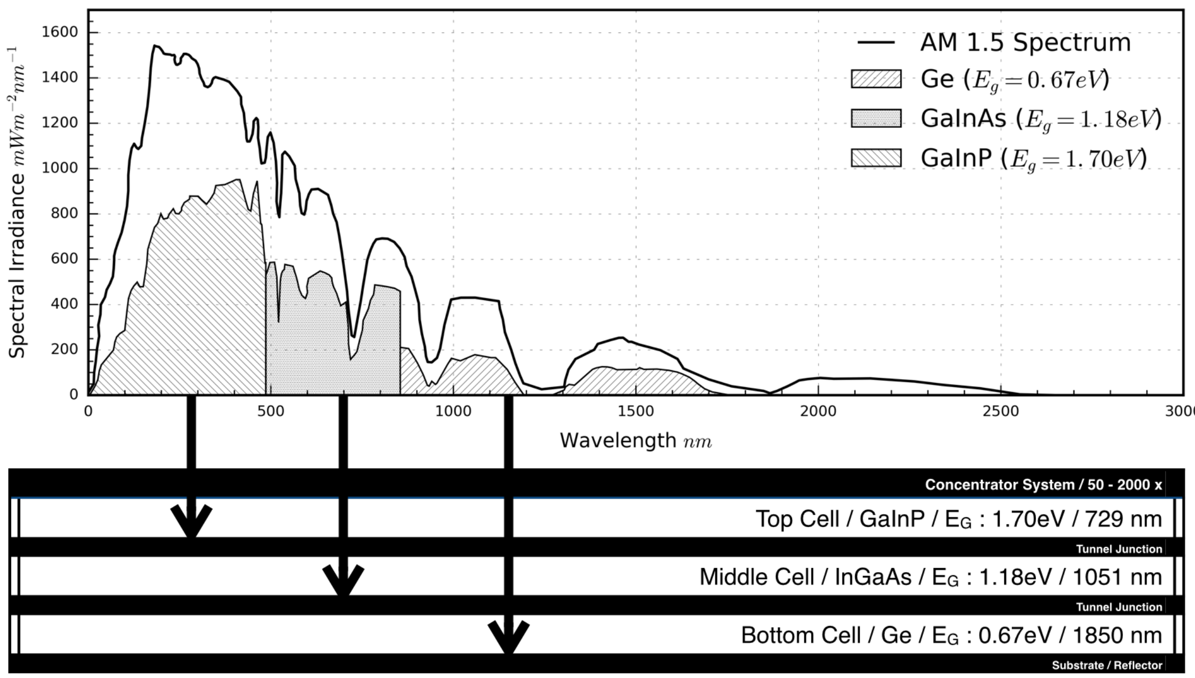

Figure 1: (Top) The AM 1.5 solar spectrum is shown as a solid line. The shaded regions are the parts of the spectrum that can be captured theoretically by a GaInP/InGaAs/Ge 3J CPV cell. (Bottom) A diagram of a typical, commercially available triple-junction GaInP/InGaAs/Ge 3J CPV cell, showing how each stacked layer is designed to absorb a specific slice of the solar spectrum.

Second, it is important to consider the optical tolerance of the concentrating system in CPV. Typical optical acceptance half-angles range from 0.7° to 3.5°.[17,18] Depending on system design, the optical tolerance may hence be inconsistent with the definition set by the World Meteorological Organisation, which defines it as ±2.5°.[3] Consideration of the acceptance angle will have implications for how much spectral circumsolar normal irradiance (CSNI$_\lambda$) one must include for realistic estimations of the exact amount of irradiance CPV can capture in the field. Since CSNI$_\lambda$ includes strong contributions from the forward scattering of aerosols, it is expected to have a strong spectral signature too. If the circumsolar contribution, and its spectral variation, is not accounted for properly, DNI$_\lambda$ could be systematically biased.[19]

In the light of these considerations, there is a need to routinely generate CSNI$_\lambda$ based on long-term observations so that these systems can be designed, optimised and deployed based on real atmospheric behaviour. We must also do so in a way that is consistent with the optical tolerances of concentrator systems. Ideally any such CSNI$_\lambda$ generation scheme should be relatively fast, flexible and use freely available data and software packages.

We introduce and evaluate four automated methods to estimate DNI_λ, each with a different way of handling the spectral signature of aerosol extinction and the spectral circumsolar irradiance (or lack thereof). In common with several existing physics-based DNI estimation techniques,[7,20,21] our methods are based on the use of radiative transfer modelling, in this case in conjunction with European Centre for Medium-Range Weather Forecasts (ECMWF) reanalyses[22,23] and AErosol RObotic NETwork (AERONET) ground-based observations.[24]

The organisation of this paper is as follows. Sections 2.1 and 2.2 introduce the tools used to develop the schemes and the validation sites, respectively. Section 3 describes the four automated schemes to estimate DNI_λ while section 4 presents the validation. Section 5 discusses the merits and limitations of each scheme before we draw conclusions in section 6.

2 Tools, Measurements & Validation Sites

2.1 Tools & Data Sources

2.1.1 Radiative Transfer Calculations

Irradiance calculations are performed using libRadtran version 2.0.2.[25] libRadtran has been validated against other radiative transfer models[26,27] and against measurements.[28] It has also been used for solar energy applications.[29,30] A two-stream radiative transfer equation solver is used to minimise computational time, with the DNI_λ obtained using

$$\mathrm{DNI}_\lambda = F_\lambda^{\downarrow} cos(\theta_z), \tag{1}$$

where $F_\lambda^{\downarrow}$ is the surface downwelling flux. The output spectral resolution of $F_\lambda^{\downarrow}$ is determined by the input extra-terrestrial solar spectrum. We use the Kurucz $0.1nm$ spectrum[31] scaled to the daily top of atmosphere total solar irradiance measurements from SORCE.[32] We account for varying site elevation using the GMTED2010 elevation dataset.[33]

2.1.2 Aerosol Optical Depth

We use aerosol optical depth (AOD, τ_λ) from the AErosol RObotic NETwork (AERONET).[24] The precision of τ_λ is stated as ± 0.01 for $\lambda > 440nm$.[24] The wavelengths at which τ_λ is measured vary from site to site depending on the exact sun-photometer model. Where needed, τ_λ is interpolated assuming a constant Angström exponent α between the nearest available τ_λ's. α is defined as:

$$\alpha_{\lambda_{1,2}} = -\frac{\log \frac{\tau_{\lambda_1}}{\tau_{\lambda_2}}}{\log \frac{\lambda_1}{\lambda_2}} \tag{2}$$

where τ_{λ_1} and τ_{λ_2} are τ_λ measured at wavelengths λ_1 and λ_2. α also captures the aerosol size distribution. Generally, a higher α implies a more pronounced spectral variation and is indicative of finer aerosols such as urban anthropogenic aerosols. Meanwhile, a lower α implies coarser aerosols, typically characteristic of aerosols of natural origin, such as desert dust or marine salts. Since α is a measure of the spectral variation of aerosol extinction, many studies have shown that it has important effects on MJ CPV performance.[34] More generally, the combination of τ and α broadly classify aerosols into different types[35,36](figure 2).

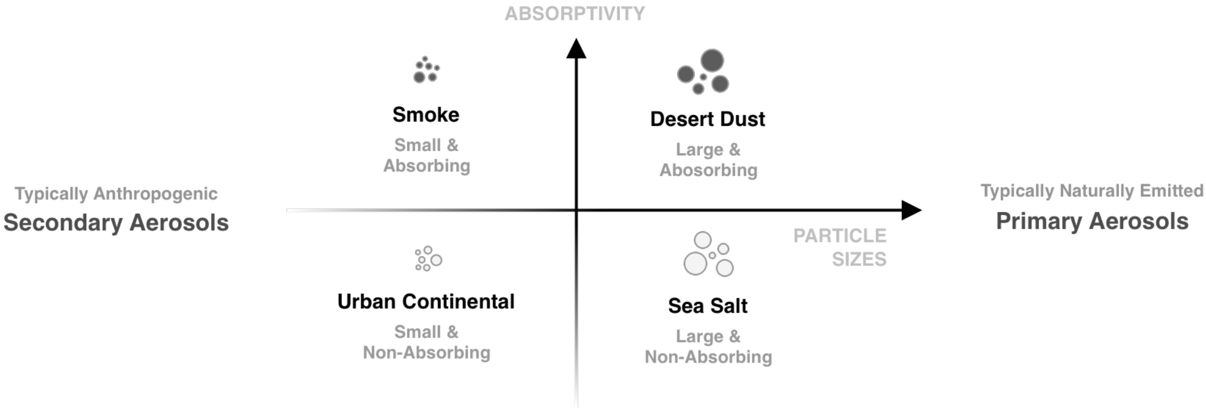

Figure 2: Aerosol classification based on discrimination in a bivariate parameter space spanned by an aerosol loading proxy and an aerosol size distribution proxy. This illustrative diagram is not drawn to any scale, and we do not imply, for example, smoke and desert dust as having similar absorptivity or smoke and urban aerosols as having similar size distributions.

2.1.3 Aerosol Optical Properties

We use aerosol properties from the Optical Properties of Aerosols and Clouds[37] (OPAC) database, conveniently available within libRadtran. OPAC aerosol types are sets of pre-defined aerosol species with pre-computed optical properties based on assumed micro-physics. The ten OPAC aerosol types are continental (clean, average and polluted), urban, maritime (clean, polluted and tropical), desert (standard and spheroids) and arctic.

The optical properties for each OPAC type are given by a weighted sum of components, such as soot, minerals and sulphate droplets. Each component is described by a size distribution and its spectral refractive index. The assumed external mixing of components means that there are no physical or chemical interactions amongst components. The optical properties, such as extinction coefficient and single scattering albedo are provided at 61 wavelengths between 0.25 and 40 μm, encompassing the solar energy relevant range. Finally, the AOD for each aerosol type is obtained by integrating the vertical distribution of the mass extinction coefficient, which follows the number density height profile from OPAC.

Figures 3(a) and 3(b) demonstrate the importance of aerosol typing for solar irradiance. We show the difference in simulated spectral and broadband DNI with varying OPAC aerosol type under an otherwise identical atmosphere and with the same AOD at $500nm$. The relative variation amongst aerosol types is substantial in the broadband, with differences of up to $100Wm^{-2}$, but is even more striking spectrally. The largest deviations occur within the top InGaP and middle InGaAs spectral range of a typical InGaP/InGaAs/Ge cell. For instance, there can be up to a $50mWm^{-2}nm^{-1}$ difference at the bandgap of the InGaP layer, and a $150mWm^{-2}nm^{-1}$ difference within the InGaAs band. Depending on the real world conditions, one of these two bands is typically the current limiting layer in a triple-junction cell, and hence caps the overall power output.

Figure 3(c) translates the observed differences in 3(a) and 3(b) to their impact on the performance of a InGaP/InGaAs/Ge MJ cell using a solar cell simulator model, SolCore [38] (see also section 5.3). There are many ways to quantify this performance impact, one of which is the spectral factor (SF), defined for each of the subcells i in an MJ cell. [12,14,34]

$$\mathrm{SF}_i = \underbrace{\frac{\int DNI_\lambda d\lambda}{\int DNI_\lambda^{\mathrm{ref}} d\lambda}}_{\mathbf{A}} \underbrace{\frac{\int R_\lambda^i DNI_\lambda^{\mathrm{ref}} d\lambda}{\int R_\lambda^i DNI_\lambda d\lambda}}_{\mathbf{B}} \tag{3}$$

where the 'ref' superscript denotes the reference spectrum (see figure 1). Since the first term **A** is simply the ratio of the broadband DNI's, it is a normalisation factor. The second term, **B**, is the ratio of the narrowband DNI with response function R_λ^i for each subcell i (shaded in grey in figure 3(a) (d) and (g)). A positive SF_i implies a spectral gain relative to the reference spectrum, whereas a negative SF_i implies a spectral loss along the spectral region covered by subcell i. As the total output current of the whole cell stack is restricted to the minimum current of the subcells due to in-series connection, SF_i is indicative of which of the subcells are current limiting. Figure 3(c) shows that an inappropriately selected aerosol type may substantially impact the estimated SF_i by up to $\pm 11\%$. For example, selecting a maritime clean aerosol gives a SF_i for the InGaP layer that is 8.6% higher and a SF_i for the Ge layer that is -8.1% lower than the baseline continental clean type. This is consistent with the fact that the spectral DNI difference between maritime clean type and the baseline continental clean type is positive at shorter wavelengths ($\lessapprox 500nm$) and negative at longer wavelengths (figure 3(a)).

2.1.4 Atmospheric Gas Profiles

Atmospheric profiles for all other species are taken from either the Copernicus Atmospheric Monitoring Service (CAMS) or Monitoring Atmospheric Composition and Climate (MACC) (prior to 2012) reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF).[22,23] Six hourly O_3, NO_x, humidity and temperature profiles at $0.125° \times 0.125°$ resolution are obtained at the closest grid point to the site.

Aside from aerosols, total column ozone (TCO) and precipitable water vapour (PWV) are also highly solar energy relevant (section 1). Rossana et al.[39] validated the TCO from the ECMWF Integrated Forecasting System (IFS) against ground-based Dobson measurements, and demonstrated a typical uncertainty of $\pm 5\%$ at low altitudes. Nock and Nuret[40] assessed PWV from the IFS against ground-based GPS receivers over Africa, and obtained a mean bias of $0.0 \pm 8\%$. We translate these columnar uncertainties into their corresponding impact on spectral DNI, broadband DNI and SF_i (figures 3(d) to (i)). In the broadband the impact of the uncertainty is $\pm 0.64 Wm^{-2}$ and $\pm 4.3 Wm^{-2}$ for TCO and PWV, respectively. The spectral impact is confined to the major water and ozone absorption bands. The impact on SF_i is limited to $\pm 0.7\%$ for PWV, and is negligible ($< 0.007\%$) for TCO. These results indicate that, compared to the impact of incorrect aerosol typing, the uncertainties induced by TCO and PWV profiles are secondary, as long as they are taken from a reliable source such as the ECMWF IFS.

2.2 Validation Sites

Five validation sites are selected for this study. They are Santiago, Chile; Niamey, Niger; Ganges Valley, India; Manacapuru, Brazil and Cape Cod, USA (figure 4). These locations are chosen because they span a range of aerosol climatologies (table 1) and have approximately co-located ground-based aerosol and broadband DNI information. Four of the five sites also have indicative measurements of the spectral variation of DNI_λ.

2.2.1 Santiago Chile

Kinne et al.[41] classified the Santiago AERONET site as having 'excellent' quality, with the ability to capture 'climatological' conditions over a range of approximately $100km$. They further indicated that

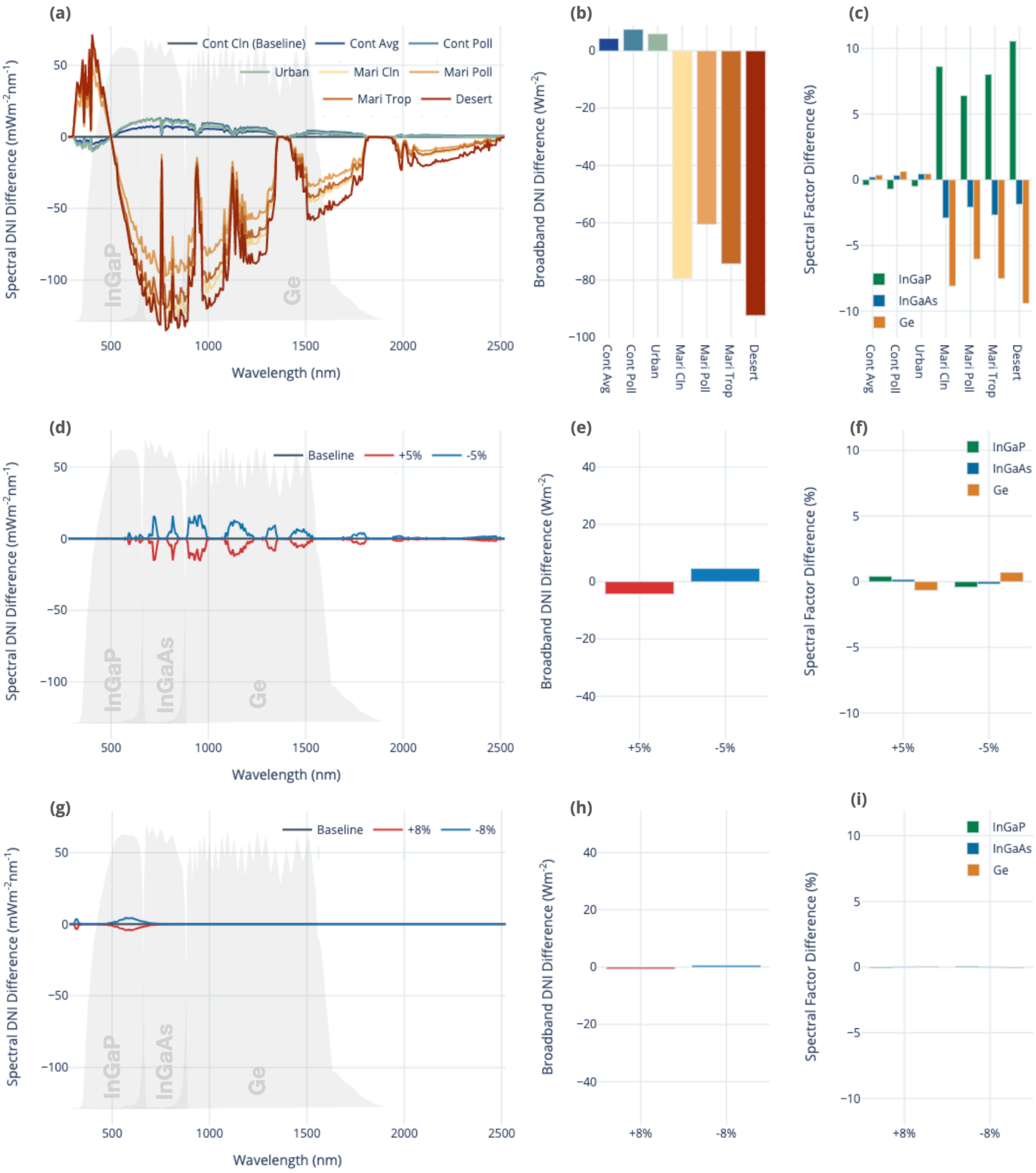

Figure 3: Simulated differences in (a) spectral, (b) integrated broadband DNI (c) spectral factor for each of the three InGaP, InGaAs and Ge subcells, simulated for different OPAC aerosol types using continental clean as the baseline. All other atmospheric parameters are held fixed for a typical day at Santiago, Chile with moderate aerosol loading ($\tau_{500} = 0.3$). Atmospheric profiles are taken from ECMWF CAMS. (d), (e) and (f) show the spectral DNI, broadband DNI and spectral factor differences when total columnar ozone is scaled by ±5%. (g) (h) and (i), as (d), (e) and (f), but with precipitable water vapour scaled by ±8%. The quantum efficiency of the absorption bands of a GaInP/InGaAs/Ge MJ cell is overlaid to indicate the spectral range of interest in (a), (d) and (g). These were used to compute the spectral factors shown in (c), (f) and (i)

Figure 4: Locations of the five validation sites for this study at (a) Santiago Chile; (b) Niamey Niger; (c) Ganges Valley India; (d) Manacapuru Brazil; (e) Cape Cod USA.

Table 1: Summary of Validation Sites and Available Instruments with Period of Operations for Validation

Site	**Climatology**	**Elevation**	**Period**	**Broadband**	**Spectral**
Santiago Chile	Maritime	579 m	Feb 14 - Jan 15	LI200/CHP	-
Niamey Niger	Desert	205 m	Aug 06 - Dec 06	SKYRAD	MFRSR
Ganges Valley India	Continental	1939 m	Sep 11 - Mar 12	SKYRAD	MFRSR
Manacapuru Brazil	Urban	50 m	Dec 13 - Dec 15	SKYRAD	MFRSR
Cape Cod USA	Maritime	47.9 m	Jul 12 - Jul 13	SKYRAD	MFRSR

the dominant aerosol type is pollution, to be expected given that the site is in an urban industrial area. However, an analysis of τ_{500} and $\alpha_{440-870}$ measurements from AERONET indicates that maritime aerosols from the Pacific Ocean, episodic dust outbreaks from the nearby Atacama Desert, and biomass burning events are also sampled (figure 5(a)).[42]

We use in-situ broadband DNI measurements, available at one-minute resolution from March 2014 to December 2014 for validation. Broadband measurements at the Pontificia Universidad Católica laboratory were conducted with pyrheliometer devices of the CHP1 type, and also with Rotating ShadowBand Irradiometers that compute DNI from global horizontal and diffuse horizontal irradiance measured with LI200 photodiode pyranometers.[43] Calibration of all sensors is achieved by direct traceability to the World Radiometric Reference. The calibration uncertainty of the rotating shadow band measurements is estimated to be typically ±3% with a maximum uncertainty of ±5%.

2.2.2 Niamey Niger

We exploit the deployment of the Atmospheric Radiation Measurement (ARM) Program's Mobile Facility (AMF)[44] near Niamey airport (13.48°N, 2.17°E) in 2006. Trajectory analysis indicates that aerosols over Niamey during this period are typically of desert type, originating from the northeast, but some trajectories originating from the Atlantic were also found.[45] Dependent on the time of year, the site also experiences biomass burning episodes.[46] Moreover, given the location of the AMF deployment, periodic increases in pollutants such as NO_x and ozone associated with local air traffic and nearby urban areas

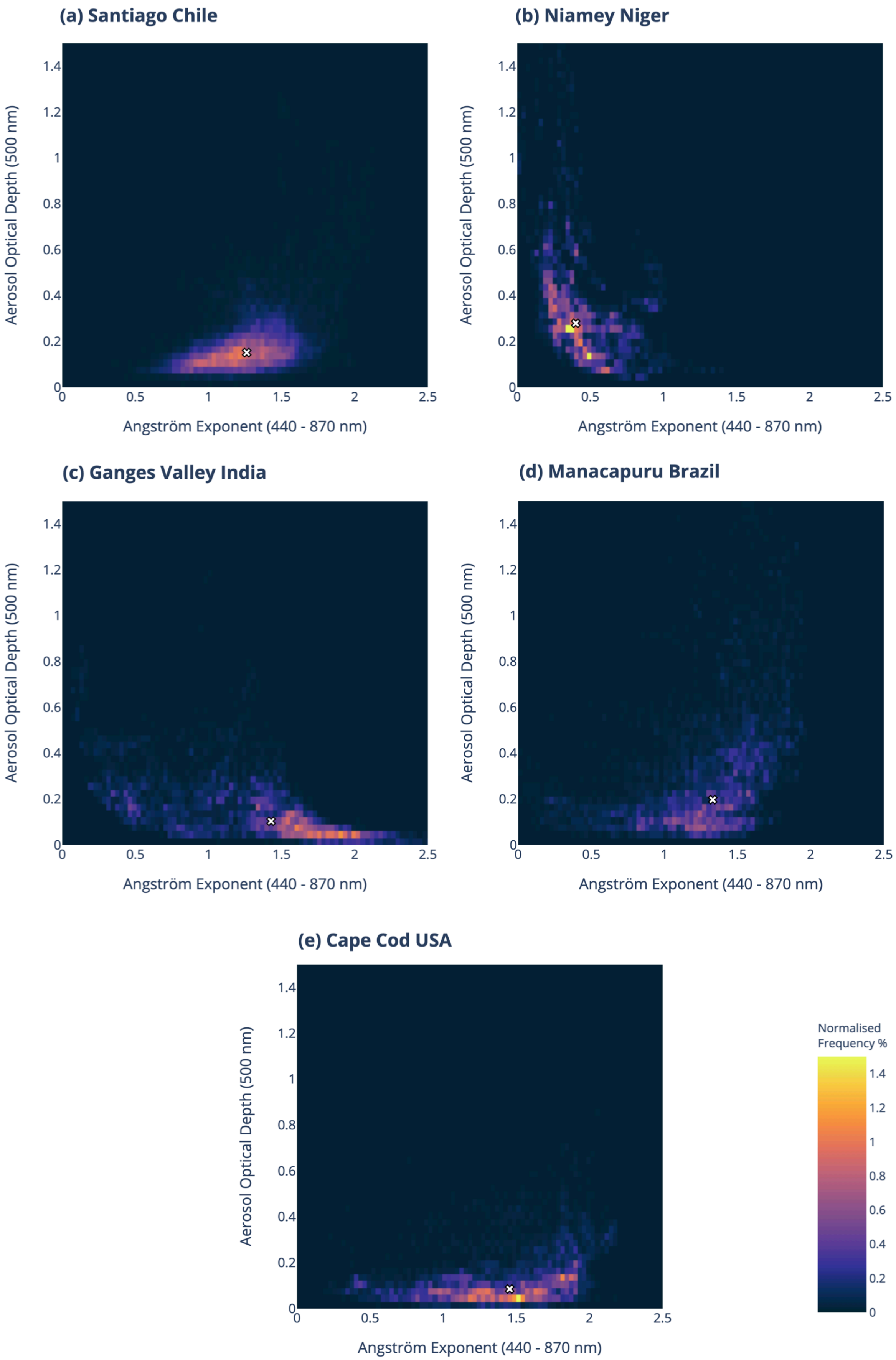

Figure 5: Aerosol climatology for each of the validation sites using τ_{500} (aerosol loading proxy) and $\alpha_{440-870}$ (aerosol size distribution proxy) from AERONET. The colour scale shows the normalised count frequency. The white crosses mark the median of the distributions.

is expected.

Broadband DNI measurements are from the Sky Radiometers on Stand for Downwelling Radiation (SKYRAD) instrument,[47] a pyrheliometer with a nearly flat spectral response from $300nm$ to $3000nm$. The aperture half angle of the pyrheliometer is 2.85°. The reported uncertainty in the broadband DNI is ±3%. Co-located with SKYRAD was a visible multi-filter rotating shadowband radiometer (MFRSR). MFRSR is a passive instrument that measures the global and diffuse components of solar irradiance in six narrowband channels at 415, 500, 615, 673, 870, and 940 nm. The blocking angle of the MFRSR shadowband is 7.8°.[48] The direct component is then obtained from the difference between the diffuse and global measurements. The uncertainty of MFRSR measurements depends on the quality of the in-situ Langley calibration, which requires stable clear-sky conditions, but is estimated to be a few percent.[48]

2.2.3 Ganges Valley India

Measurements are from the Ganges Valley Aerosol Experiment (GVAX) from June 2011 to March 2012.[49] This site is situated at $1.9km$ a.s.l. in the foothills of the Himalayans. The site is bounded by high mountains to the north and east, but is otherwise open to the Indo-Gangetic Plain towards the south. This exposes the site to transported anthropogenic air pollution arising from fossil fuel combustion and biomass burning from agricultural activities,[50] as well as mineral dust from the Thar Desert.[51] The previously described SKYRAD, AERONET and MFRSR instruments were all available during GVAX.

2.2.4 Manacapuru Brazil

Measurements are from the Observations and Modeling of the Green Ocean Amazon (GoAmazon) Experiment from 2014 to 2015.[52] The site was located downwind of the city of Manaus, Brazil and experienced highly variable aerosol conditions depending on whether heavy pollution was transported from the urban area of Manaus. The city of Manaus uses high-sulphur oil as its primary source of electricity and is also home to an industrial zone of three million people. As such, urban aerosol, soot and black carbon are expected to be transported over Manacapuru.[52] Situated in the Amazon, one would also expect biomass burning and secondary biogenic volatile organic compounds.[53] The same set of instruments were available as at Niamey.

2.2.5 Cape Cod United States

Measurements are obtained from The Two-Column Aerosol Project (TCAP).[54] This measurement site sits along the coast of North America, near Cape Cod, Massachusetts. Being a coastal site, we expect a maritime-like aerosol climatology. This site is notable because it is free from significant sources of local anthropogenic emissions. Indeed, the observed median aerosol loading from AERONET over the area is the lowest amongst the five validation sites (figure 5(e)). The same suite of instruments as at Niamey were available.

2.2.6 Data Screening

Thomalla et al.[55] showed that even the slightest cloud contamination can have substantial effects on the solar irradiance, particularly the circumsolar component. As such, we limit our analysis to clear sky conditions. Cloud screening is partly achieved by only utilising the quality-assured level 2.0 AERONET product.[56] Furthermore, in line with Kaskaoutis et al.,[57] we threshold the diffuse-to-global and direct-to-diffuse irradiance ratios for the broadband measurements. We also require the aerosol and DNI observations used for validation to be made within $\pm 10min$. These criteria, summarised in table 2, represent the best efforts to ensure only clear sky conditions are sampled and that aerosols and DNI observations are comparable. Dependent on the site, between 17% to 37% of the total number of level 2 AERONET observations are removed.

Table 2: Cloud Screening Criteria for Each Measurement Instance (*if instrument present)

Quantity	**Criteria**
AERONET Measurements	Available at level 2.0
SKYRAD Measurements*	Available to within $\pm 10min$ of an AERONET measurement
MFRSR Measurement*	Available to within $\pm 10min$ of an AERONET measurement
Solar Zenith Angle	$< 75°$, avoid deviation from a cosine directional response
Direct to Diffuse Ratio	> 0.5
Diffuse to Global Ratio	< 0.55

3 Methodology

Table 3 describes the four DNI estimation schemes. The Climatological Aerosol Scheme (CAS) represents the baseline approach where the τ_λ spectral variation is represented by an *a priori* climatological aerosol type. The Angström Exponent Scheme (AEX) is a literature-based approach where the τ_λ spectral variation is approximated by a linear fit. The Spectral Aerosol Classification Scheme (SACS) proposes a new aerosol classification method and uses the classified type to account for the τ_λ spectral variation. Finally, the Spectral Aerosol Classification Scheme with Circumsolar Factor (SACS + CSF) extends this scheme to include spectral circumsolar effects.

Table 3: Automated Spectral DNI Estimation Schemes

Scheme	Acronym	τ_λ Spectral Variation Representation
(1) Climatological Aerosol Scheme	CAS	Climatological Type
(2) Angström Exponent Scheme	AEX	Linear Fit
(3) Spectral Aerosol Classification Scheme	SACS	Optimal OPAC Type Fit
(4) SACS with Circumsolar Factor	SACS + CSF	Optimal OPAC Type Fit + CSF LUT

3.1 The Climatological Aerosol Scheme (CAS)

In the Climatological Aerosol Scheme (CAS), we fix τ_{500} to the AERONET observations. Then, the spectral variation beyond $500nm$ is accounted for by prescribing a fixed climatological aerosol type *a priori* from one of the ten OPAC types. The climatological aerosol type is selected automatically by examining the combination of AERONET τ and α measurements available (see figures 2 and 5), with thresholds on τ and α adapted from literature to match the aerosol types available in OPAC.[35,36] We then use the climatological type, along with the instantaneous values of τ_{500}, to perform radiative transfer calculations.

3.2 Angström Exponent Scheme (AEX)

Since aerosols are typically highly variable in space and time, a fixed climatological aerosol type is likely undesirable. A more fundamental limitation is related to setting *a priori* thresholds. Literature has shown that established classification thresholds are not realistically transferrable amongst sites and climatologies.[58]

To address this, the Angström Exponent Scheme (AEX) handles the spectral variation of τ_λ using a more standard approach. The Angström Exponent (α), defined in equation 2, is the slope of τ_λ against λ in log-log space. We find α by performing a linear regression of $ln(\tau_\lambda)$ against $ln(\lambda)$ for each instance. The fitted α and τ_{500} are then used for the radiative transfer calculation. Other aerosol properties, such as the single scattering albedo and the phase function, are set to those those taken from the climatological OPAC aerosol type.

3.3 Spectral Aerosol Classification Scheme (SACS)

The AEX scheme assumes α is wavelength-independent. However, this is an approximation. Postulating a second-order polynomial, the spectral variation may be captured by a curvature term.[59–61] In libRadtran, although we may specify the curvature, there are drawbacks. First, whilst solving for the curvature is possible when there are numerous spectral measurements, it presents more of a challenge as the spectral coverage and number of points reduces. This is not much of an issue for AERONET but would present a challenge for an extension of the scheme to satellite observations, where the number of channels is limited.[62,63] Second, MJ solar cells cover a wide range of wavelengths, up to $\sim 1.8\mu m$. The polynomial fit may not be suitably robust against extrapolation to such distant wavelengths compared to the channels at which measurements are made. Again, this is particularly true for satellite products.

Our third approach, the Spectral Aerosol Classification Scheme (SACS), thus automatically classifies observations into the OPAC aerosol type that has the most compatible spectral behaviour with observations. This is then used to represent the spectral variation of τ_λ. We first compute the spectral aerosol optical depth for the ten OPAC aerosol types based on the relative humidity profiles from ECMWF IFS, along with the component mixing ratios, size distributions, refractive indices and scale height defined in OPAC. For each aerosol type, a regression is performed using least-square against AERONET measurements with observational uncertainties taken into account. For these curve fitting procedures, the only independent variable is the optical depth at $500nm$ (i.e. vertical offset of the τ_λ curves). The curve, and the corresponding OPAC type, that produces a fit with the smallest root mean square error (RMSE) is then ultimately selected following equation 4. If the minimum RMSE is larger than the reported uncertainty (0.01 for AERONET), then the classification is flagged as inconclusive. If there is more than one curve that produces RMSE smaller than the observational uncertainty, then the classification is flagged as ambiguous.

Mathematically, the SACS selects an OPAC aerosol type k with an aerosol optical depth spectral variation of $A_k(\lambda)$ that minimises the following equation:

$$min(\sqrt{\frac{1}{n}\sum_{\lambda}((A_k(\lambda) - A_k(500nm) + a) - \tau_\lambda)^2} \tag{4}$$

where n is the number of τ_λ observation channels and a is an independent variable for the regressions. The parameter a can be interpreted as the optimal aerosol loading at $500nm$ that is compatible with the spectral variation $A_k(\lambda)$ based on measurements across all channels. In other words, a captures information about the aerosol loading, whilst the type k captures information about its spectral variation. SACS allows us to include knowledge about measurement uncertainty as part of the regression, while the

$A_k(\lambda)$ curves guarantee, as far as possible, robust and physically valid extrapolation to all solar energy relevant wavelengths.

3.4 Spectral Aerosol Classification Scheme with CSF (SACS + CSF)

The DNI_λ estimated in section 3.3 is still not strictly compatible with real measurements due to the presence of circumsolar irradiance (CSI). CSI includes scattered photons in the vicinity of the direct solar beam. As aerosol scattering is strongly peaked in the forward direction, the CSI contribution can be substantial. If the scattering species are evenly distributed horizontally in the atmosphere, then CSI would decrease monotonically from the centre of the Sun.[64] The rate of decrease depends strongly on the aerosol type present and is also spectrally dependent, making its inclusion crucial to DNI_λ estimation.

The presence of CSI means that the definition of DNI is somewhat ambiguous[17,65] as it involves an integration over a small angle centred at the solar disc. Different instruments have varying definitions of ‘small angle’. The varying optical acceptance aperture sizes can, at least to first order, be characterised by the aperture half angle α_0 (figure 6). It is the angle over which the CPV system can efficiently direct light onto the cells. α_0 depends on the concentration factor, and is typically 1°, but can range from 0.4° to 2.0° [66]. Therefore, α_0 for MJ CPV systems are generally smaller than that of sunphotometers ($\approx 2.5°$).

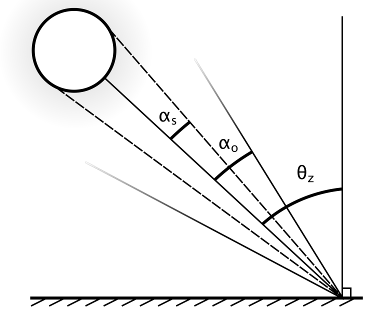

Figure 6: Angles important to the description of CSI. α_0 is the aperture half angle; $\alpha_s = 0.266°$ is the angular radius of the solar disc; θ_z is the solar zenith angle.

DNI_λ computed using equation 1 has an α_0 of 0.0°. This is further complicated by the fact that in radiative transfer calculations, scattered photons are not generally included in the direct flux output, whereas in reality, this distinction is non-existent. Photons scattered back into the direction of the direct beam are indistinguishable from the direct beam itself.

Previous authors have sought to address the circumsolar contribution by employing a correction to DNI or τ. Box et al.[67] showed that the correction is $< 1\%$ under clear conditions. Gueymard et al.[68] investigated hazy conditions and concluded there would be a few percent CSI contribution to DNI, depending on the aperture size and airmass. However, existing literature has largely focused on studying CSI due to desert dust and clouds, with little mention of other aerosol types.

Following Gueymard et al.,[68] we make a distinction between $\mathrm{DNI}_\lambda^{strict}$ and DNI_λ.

$$\mathrm{DNI}_\lambda = \mathrm{DNI}_\lambda^{strict} + \mathrm{CSNI}_\lambda(\alpha_0), \tag{5}$$

where $\mathrm{DNI}_\lambda^{strict}$ is the direct solar beam, projected to the direct normal, which has not been scattered. $\mathrm{CSNI}_\lambda(\alpha_0)$ is the circumsolar normal irradiance. This term includes all scattered photons, including those travelling along the direct beam after multiple scattering, in the vicinity of the direct beam up to $\pm\alpha_0$, projected to the direct normal.

To compute $\text{CSNI}_\lambda(\alpha_0)$, the radiance $L_\lambda(\Omega)$ needs to be integrated over a solid angle, Ω. This integration is different for different instruments. For pyrheliometers, such as the CIMEL sunphotometer from AERONET, and CPV systems, the integration is

$$\text{CSNI}_\lambda^{pyrhe}(\alpha_0) = \int_{2\alpha_0}\int_{2\alpha_0} L_\lambda(\Omega) cos(\theta - \theta_z) sin(\theta) d\theta d\phi, \tag{6}$$

where θ and ϕ are the polar and azimuthal angles respectively, and θ_z is the solar zenith angle. The integration limits are $\pm\alpha_0$ zenith and azimuthal angles of the Sun.

For measurements from pyranometers, the integration is subtly different. Unlike a pyrheliometer, a pyranometer derives the direct downward irradiance (DDI) from taking the difference between the global horizontal irradiance (GHI) measurement and the diffuse horizontal irradiance (DHI) measurement. DNI is then obtained by projecting the DDI to the direct normal. This means that the CSNI_λ must also be projected to the zenith direction first before integration, as shown in equation 8.

$$\text{CSNI}_\lambda^{pyran}(\alpha_0) = sec(\theta_z) \int_{2\alpha_0}\int_{2\alpha_0} L_\lambda(\Omega) cos(\theta) sin(\theta) d\phi d\theta \tag{7}$$

$$= \frac{sec(\theta_z)}{2} \int_{2\alpha_0}\int_{2\alpha_0} L_\lambda(\Omega) sin(2\theta) d\phi d\theta \tag{8}$$

where the integration limits are the same as in equation 6. The difference between equations 6 and 8 is only important for large α_0. As the directionally dependent radiance $L_\lambda(\Omega)$ is needed to obtain CSNI_λ, a two-stream solver is insufficient, and we use the DISORT solver instead.[69] In order for DISORT to adequately capture the strongly forward peaked $L_\lambda(\Omega)$ within the circumsolar region, 32 streams are used.[70]

Following the treatment of Reinhardt et al.,[64] we define the circumsolar ratio $\text{CSR}_\lambda(\alpha_0)$ as

$$\text{CSR}_\lambda(\alpha_0) = \frac{\text{CSNI}_\lambda(\alpha_0)}{\text{CSNI}_\lambda(\alpha_0) + \text{DNI}_\lambda^{strict}}. \tag{9}$$

Combining equations 9 and 5, we obtain

$$\text{DNI}_\lambda = \text{DNI}_\lambda^{strict} \frac{1}{1 - \text{CSR}_\lambda(\alpha_0)}. \tag{10}$$

Encompassed within $\text{DNI}_\lambda^{strict}$ is the AERONET measured τ_λ, which is also affected by CSNI_λ. Strictly speaking, when CSNI_λ is considered, the derived τ_λ will be underestimated. This is because the AERONET pyrheliometer is capturing more radiation than simply $\text{DNI}_\lambda^{strict}$, as it has an aperture half angle of 0.65°. However, in practice, 0.65° is small that this contribution is sufficiently less than the quoted uncertainty of AOD from AERONET under most plausible atmospheric conditions.[71]

Running DISORT with 32 streams is too computationally expensive for every single observation. Thus, we adopt a Look-Up Table. We precompute and parameterise CSR_λ as a function of θ_z, τ_λ and OPAC aerosol type k to generate $\text{CSR}(\theta_z, \tau_\lambda, k)$. The contribution to CSR from Rayleigh scattering has been taken into account by assuming gaseous profiles from a standard mid-latitude summer atmosphere.[72] The surface albedo is set to zero as it has been demonstrated that the effect on CSR_λ of changing surface albedo between the extreme cases of 0 and 1 is always below 0.0025.[64]

Figure 7 compares CSR_λ simulated for four OPAC aerosol types for different values of τ_λ, θ_z and two values of α_0, 2.5° and 1.0°. These two α_0 are selected to represent the typical acceptance half angles of sunphotometers and MJ CPV respectively. Under all conditions, CSR_λ increases at shorter wavelengths and can reach almost 50% in the case of $\alpha_0 = 2.5°$ and 20% in the case of $\alpha_0 = 1.0°$ at wavelengths around $400nm$ for high values of θ_z and τ_λ. The variation of CSR_λ is smooth except in bands with substantial water vapour absorption, as CSR_λ is undefined when $CSNI_\lambda(\alpha_0) + DNI_\lambda^{strict}$ approaches zero. Generally, CSR_λ is more important for coarser aerosols, such as those of maritime or desert type, compared to finer aerosols, such as those of continental or urban origin. Comparing across panels, we see that CSR_λ increases with increasing τ_{500} and increasing θ_z. We also see that as expected, CSR_λ is lower for smaller α_0, implying that for MJ CPV it only becomes important to consider these effects at relatively high aerosol loadings and solar zenith angles.

4 Results

4.1 Broadband Validation

Figure 8 shows how the broadband DNI estimated by each scheme compares with ground-based broadband DNI measurements. The distributions of the percentage error have been fitted using a gaussian, with the regressed means and standard deviations shown in table 4.

Table 4: Fitted Means and Standard Deviations of Broadband DNI Error Distribution in % For Each Validation Site and Scheme

Site	**CAS**		**AEX**		**SACS**		**SACS + CSF**	
	μ	σ	μ	σ	μ	σ	μ	σ
Santiago Chile	-3.19	3.57	0.82	2.67	-0.32	2.82	**-0.06**	2.69
Niamey Niger	-4.11	3.48	3.73	1.79	-0.7	1.94	**0.37**	1.31
Ganges Valley India	1.15	4.57	1.55	5.55	-0.58	3.52	**-0.22**	3.68
Manacapuru Brazil	-0.29	5.48	3.92	2.53	-0.88	3.38	**-0.27**	3.29
Cape Code USA	-2.35	2.91	1.78	1.41	**0.39**	1.59	0.52	1.56

For three out of the five sites, CAS shows the largest magnitude in mean bias, reaching −4.11% at Niamey, with the AEX scheme showing the largest values for the remaining 2 sites. The CAS scheme also tends to have the widest error distribution, up to 5.48% for Manacapuru. The relatively poor performance of CAS is not surprising given it only considers the climatological aerosol type. While the use of the Angström Exponent improves the error statistics for some sites, the bias is increased over Ganges Valley and Manacapuru Brazil. In addition, the AEX scheme shows a systematic overestimation across all sites, which may be attributed to the linear approximation of the spectral dependence of τ_λ.

Applying the Spectral Aerosol Classification scheme (SACS) reduces the bias for all sites to under 1%. Finally, the inclusion of the circumsolar factor in the SACS + CSF scheme yields the smallest bias for four out of the five sites. In addition, the spread of error for SACS + CSF is typically smaller compared to SACS. The magnitude of the difference between the inclusion and exclusion of CSF depends on the site. Generally, for sites with coarser and more strongly scattering aerosols, such as Niamey, Ganges Val-

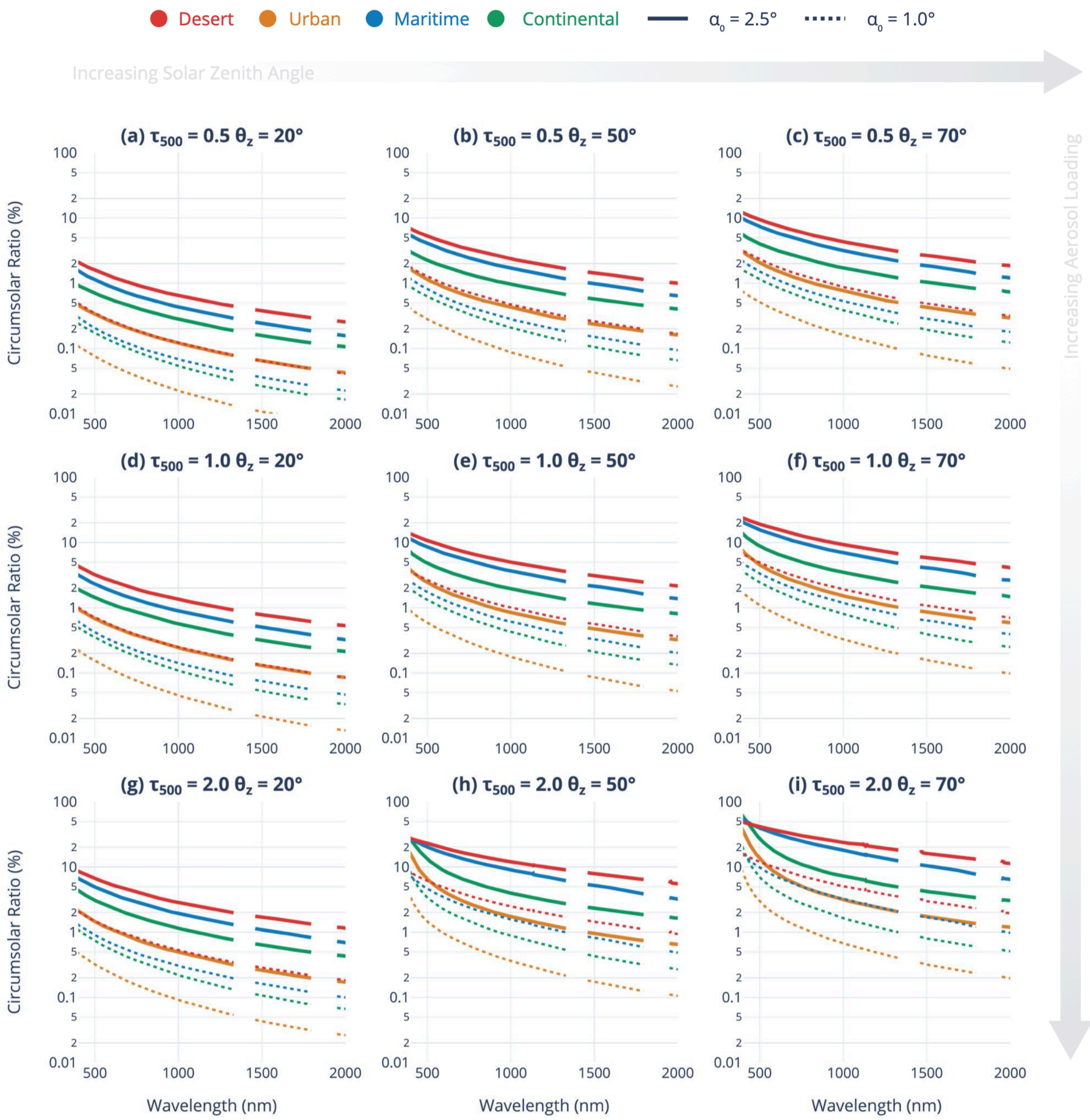

Figure 7: Comparison of CSR_λ for selected OPAC aerosol types under various combinations of θ_z and τ_λ. Regions of substantial water vapour bands are masked out as CSR_λ is undefined. Calculations assuming an aperture half angle of 2.5°, typical of pyrheliometers, are shown with solid lines. Calculations assuming an aperture half angle of 1.0°, typical of MJ CPV systems, are shown with dotted lines.

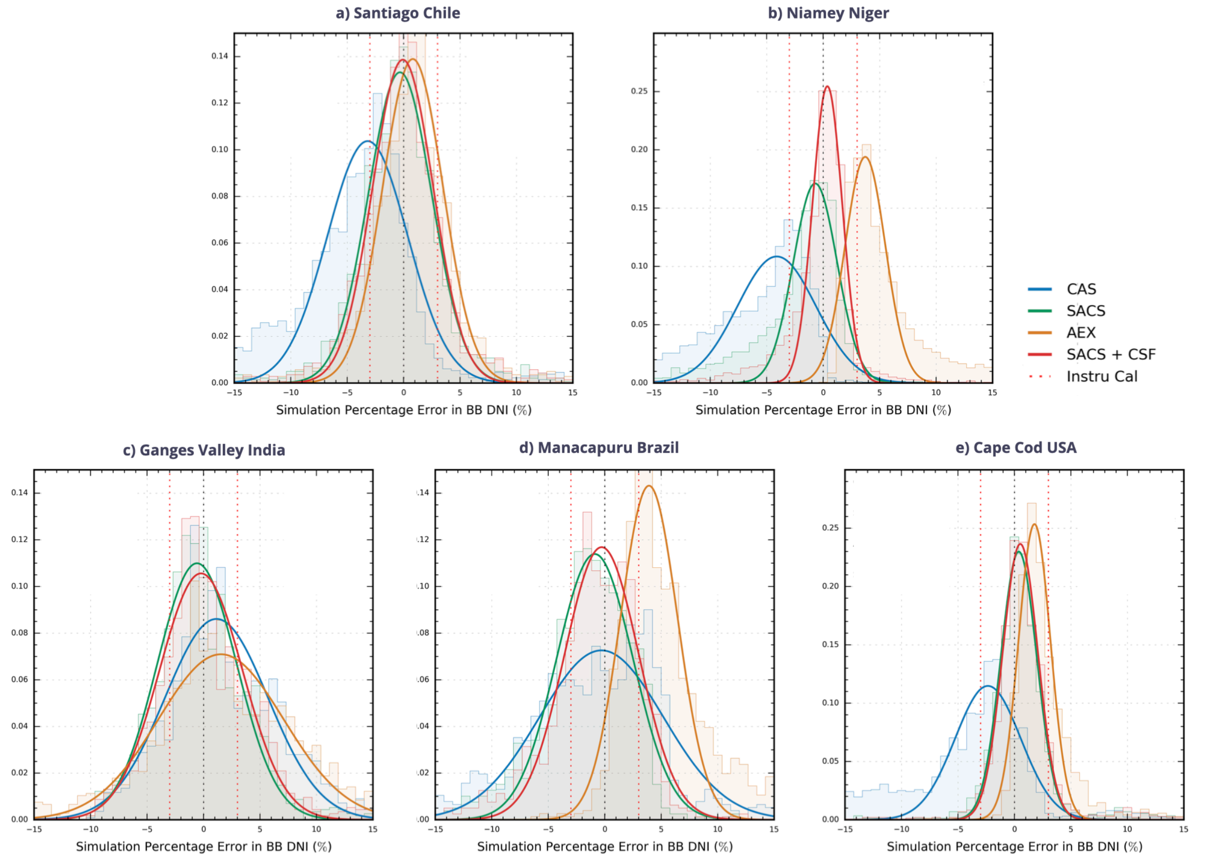

Figure 8: Histogram of percentage error against the appropriate ground-based broadband observation associated with each scheme for (a) Santiago Chile, (b) Niamey Niger, (c) Ganges Valley India, (d) Manacapuru Brazil and (e) Cape Cod USA. A gaussian regression is performed for each of the schemes, and the fitted mean and standard deviation are listed in table 4. The y-axes for all subplots are the normalised occurrence frequency.

ley and Manacapuru, the inclusion of CSF leads to a more substantial improvement. This is expected, as the CSF correction is more marked for these aerosol conditions (figure 7).

4.2 Spectral Validation

Figure 9 shows the corresponding spectral validation. Each solid line shows the median of the distribution within each MFRSR channel, whilst the error bars indicate the interquartile range. For all locations, the AEX scheme underestimates substantially in the shortwave and overestimates towards longer wavelengths due to its linear approximation. It is likely that these opposing biases are partially, but not entirely compensating when considering the broadband DNI results (table 4). Compared to AEX, the other schemes yield a flatter bias across the spectral range.

At Manacapuru (figure 9(a)), the channel dependent bias of AEX is particularly evident, with a strong underestimation at 415nm of more than 10% and an overestimation of the same magnitude at 868nm. For other schemes, the bias is less wavelength-dependent, with the SACS + CSF scheme producing the smallest bias across all channels, overestimating by approximately 1% at 500nm and underestimating by $\sim 3\%$ at 868nm. As expected from figure 7, the difference between the inclusion and exclusion of CSF is wavelength dependent, with a stronger CSF correction at shorter wavelengths. For Niamey (figure 9(b)), the SACS + CSF scheme also produces the smallest bias across all but the 868nm channel, with a bias of less than -0.2% between 500nm and 670nm. In contrast, the CAS scheme consistently underestimates for all channels whilst AEX exhibits the aforementioned wavelength dependent bias, reaching +5% at 868nm. The SACS + CSF scheme also exhibits a noticeably smaller interquartile range for all channels compared to the other schemes. At Ganges Valley (figure 9(c)), the performance of CAS, SACS and SACS + CSF is similar. All three schemes typically exhibit negative biases relative to the observations except for the 670nm channel. Again, the AEX scheme shows a positive bias at the longest wavelength, whilst all other schemes tend to underestimate the DNI. Finally, for Cape Cod (figure 9(d)), the difference in the bias distribution between SACS and SACS + CSF is negligible. This is because the circumsolar correction factor is very small due to the close to pristine skies observed through the measurement period. Both schemes produce a flatter bias curve compared to the CAS and AEX schemes.

Averaging over all sites, the average bias magnitudes for SACS+CSF show the closest match to observations across all channels. For the 415nm channel, the average bias is -0.1%. The 500nm and 614nm channels on average overestimate by approximately $+0.3\%$. The 670nm and 868nm channels exhibit the largest biases at -0.52% and -0.6% respectively, though both of these values are still within the instrument calibration uncertainties.

4.3 Comparison to Other Work

Two popular spectral DNI models, SPECTRAL2[73] and SMART2,[74] have mean bias error of about 1% for most cases as validated over Egypt[75] and around the US.[76] The SUNFLUX[77] scheme, on the other hand, yields biases of 3.03%.[75] Here, across our five validation sites sampling a wide range of aerosol conditions, the mean percentage error of our most sophisticated scheme, SACS + CSF, is +0.068%. When it comes to the spread of error, the RMSE for SPECTRAL2 and SMART2 is about 3%, while SUNFLUX carries a RMSD of about 5%.[75] The SACS + CSF produces a similar or smaller percentage spread of error with the average spread across sites approximately 2.5%. Obviously a direct comparison with previous literature is somewhat compromised by the fact that different schemes have used different inputs and their validation was performed using different sites.

Unlike fast schemes like SUNFLUX, computational speed is not our primary objective. However, as part of the development process, computational efficiency has been considered as exemplified by the use of the two-stream approximation (section 2.1.1) and of LUTs for the circumsolar factor (section 3.4). On a modern Linux server, the SACS + CSF scheme takes about 10 seconds or less to compute a DNI spectrum. This includes the conversion of raw netCDF data from ECMWF to formats and units compatible with libRadtran solvers, the SACS classification process, the radiative transfer calculation and the circumsolar ratio look-up.

5 Discussion

The spectral aerosol classification method and the circumsolar factor account for the majority of improvements over the CAS and AEX schemes. In this section, these are examined more closely.

5.1 Spectral Aerosol Classification

Using Niamey as an example, the SACS aerosol classification is examined in figure 10. Figure 10(a) shows the fitted curve for each of the OPAC types, superimposed on the AERONET observations at 21:13 UTC on 29 September 2006. On the right, we show the types as ranked by the RMSE of the fits. For this example, the aerosol type that minimises the RMSE is Maritime Polluted, but Maritime Tropical and Maritime Clean are also likely candidates as their RMSE is also below the AERONET uncertainty. The maritime typing is corroborated by figure 10(c), which shows HYSPLIT airmass trajectories terminating over Niamey that day,[78] with most originated from the Atlantic ocean. Similarly, in figure 10(b), we

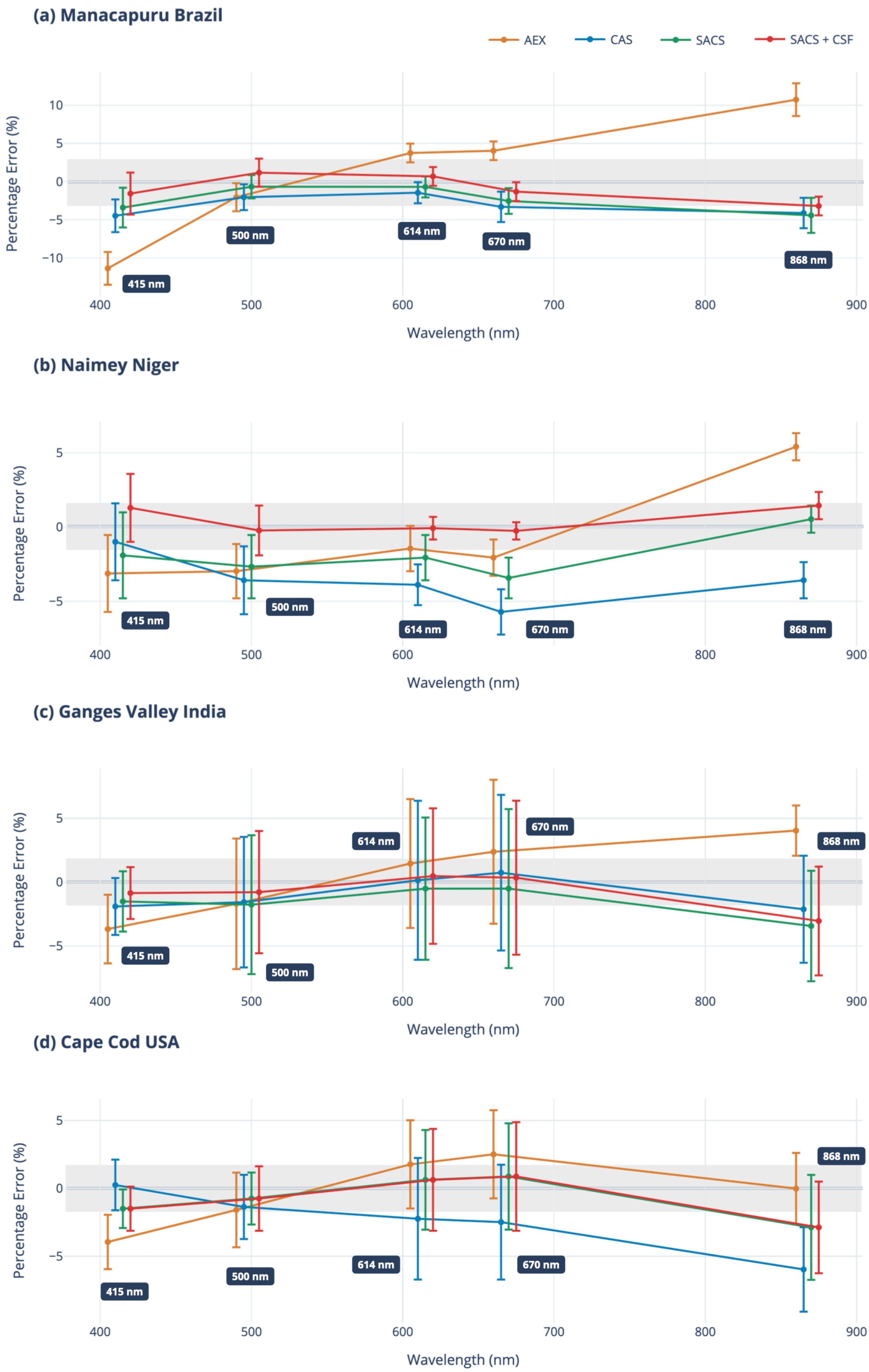

Figure 9: Spectrally resolved simulation percentage error against MFRSR measurements at (a) Manacapuru; (b) Niamey; (c) Ganges Valley and; (d) Cape Cod. Typical instrument calibration uncertainty is of a couple percent. Here, we indicate the ±3% region in grey. Note the change of scale in (a).

show the SACS typing for 13:25 UTC on 24 November 2006. On this occasion the OPAC type that minimises the fitted RMSE is Desert. This is corroborated by figure 10(d), where we observe HYSPLIT trajectories originating from the Sahara towards the Northeast.

Figure 11 compares the SACS aerosol type classification with the literature-based method of discrimination using combinations of τ_{500} and $\alpha_{440-870}$ described in section 3.1. Focusing on Santiago, we show the location of each AERONET measurement in the τ_{500}-$\alpha_{440-870}$ parameter space. Also shown are the aerosol classification thresholds from Kaskaoutis et al.[57] Of the points that fall within one of the threshold regions, the majority are classified as maritime clean. This justifies the selection of the maritime clean OPAC type for the CAS at Santiago. However, the majority of data points do not fall into any of the thresholded regions and thus are unclassified. This indicates the absence of a dominant aerosol type in the Santiago atmosphere and was one of the motivations for developing the SACS in the first place. The colours of the scattered points in figure 11 illustrate the OPAC aerosol type as determined via SACS. Not only do the classified OPAC types fall approximately in line with the threshold regions but the scheme also offers a more detailed classification.

While it seems that SACS has satisfactorily classified data points into one of the OPAC types, the issue of typecasting all data points into a finite set of ten OPAC types remain. For example, OPAC does not include biomass burning aerosols. Therefore, the merit of the classification is limited by how well OPAC captures the full range of aerosol characteristics.

5.2 Verifying the CSR LUT Parameterisation

The CSR_λ LUT is parametrised with k, θ_z and τ_λ. It is instructive to verify the parameterisation by examining how the simulation error varies as a function of these parameters. Figure 12 shows the broadband DNI bias at Santiago as a function of θ_z. Prior to the inclusion of CSF (figure 12(a)) there is a strong negative correlation between the bias and solar zenith angle with the largest biases seen at higher θ_z or equivalently, for longer optical paths. In these cases the amount of CSI will be higher due to increased scattering by both aerosols and air molecules. Figure 12(b), shows that the dependence of the bias on θ_z has been reduced with the inclusion of CSF.

Finally, figure 13(a) shows the distribution of broadband DNI bias, with aerosols of coarse and fine mode grouped separately. Prior to including CSF, the peaks of the two distributions are offset, with a stronger negative bias for the coarse mode aerosol. Figure 13(b) shows the distribution of errors after the CSI has been accounted for. Both distributions are shifted towards a less negative overall bias but the effect is larger for the coarse mode aerosol. This is expected as coarse mode aerosols scatter more strongly in the forward direction, and also do so across a broader part of the spectrum. As such, the error distributions for the two modes are now more aligned.

5.3 Impact on MJ CPV Power Estimates

While exploring in detail how the spectral DNI estimation scheme may couple with a MJ CPV system is beyond the scope of this work, we provide a first indication of the likely impact. This is achieved by feeding simulated spectra into Solcore [38], assuming a InGaP/InGaAs/Ge 3J cell (see figure 1). The system is then solved optically with a transfer matrix method, and electrically with the depletion approximation. The operating temperature of the cell is held at $300K$ and any spectral effects due to the CPV optics are ignored. The primary quantity simulated is the maximum power point (P_m) per unit cell area, which is the maximum power deliverable from a cell given a particular irradiance spectrum. Although we do

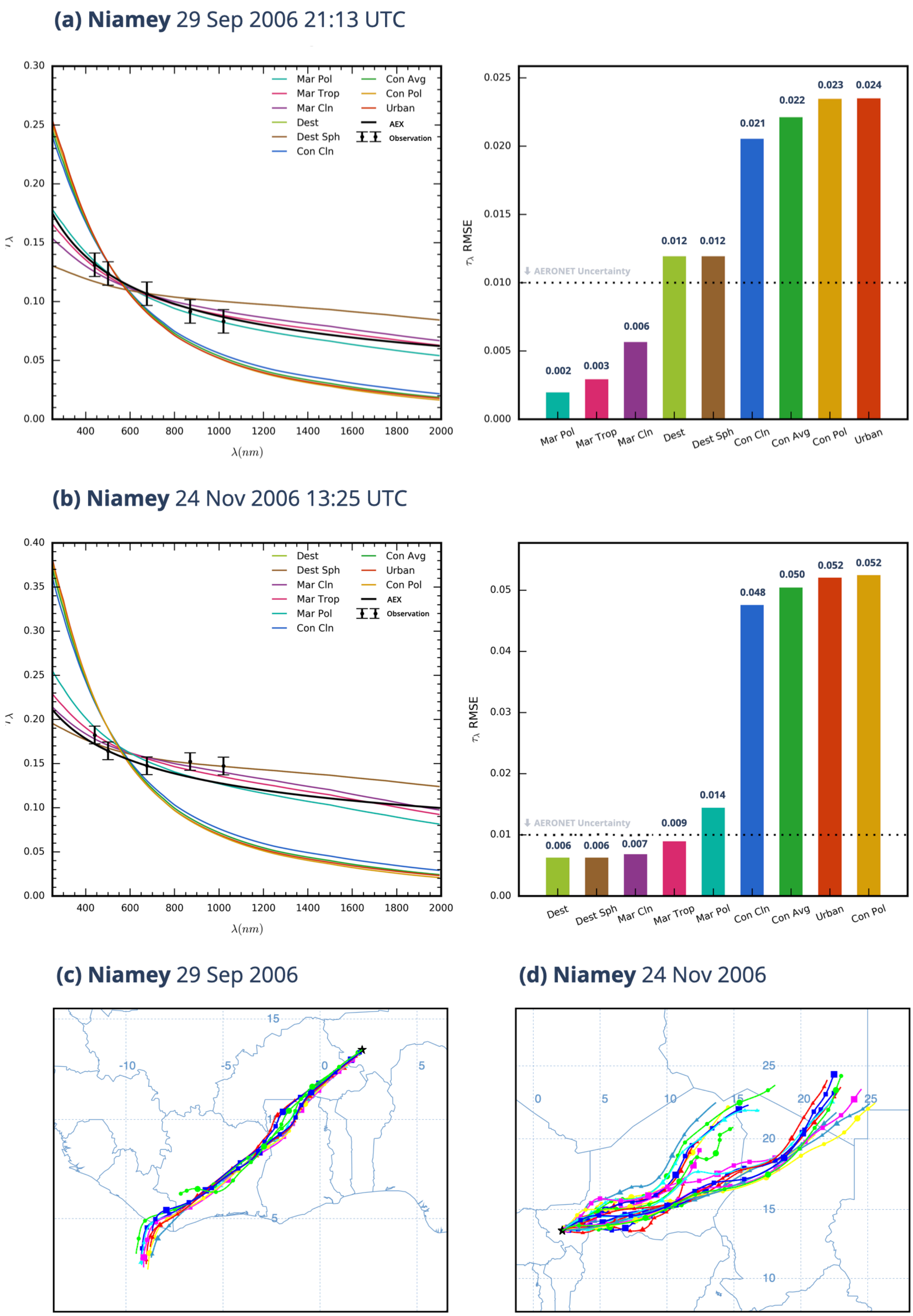

Figure 10: Examples of SACS for Niamey Niger on (a) 29 September 2006 at 21:13 UTC and (b) 24 November 2006 at 13:25 UTC. For (a) and (b), we show the curve fitting attempts for each of the OPAC aerosol types, excluding the Antarctic type, along with the AERONET observations and uncertainties on the left panels. On the right, we show the ranked fitted RMSE of the fits. (c) and (d) show airmass trajectories terminating $500m$ over the Niamey sites for the corresponding dates in (a) and (b) respectively. These 72 hours trajectories, spaced two hours apart leading up to the observation times in (a) and (b), are from the HYSPLIT model.[78]

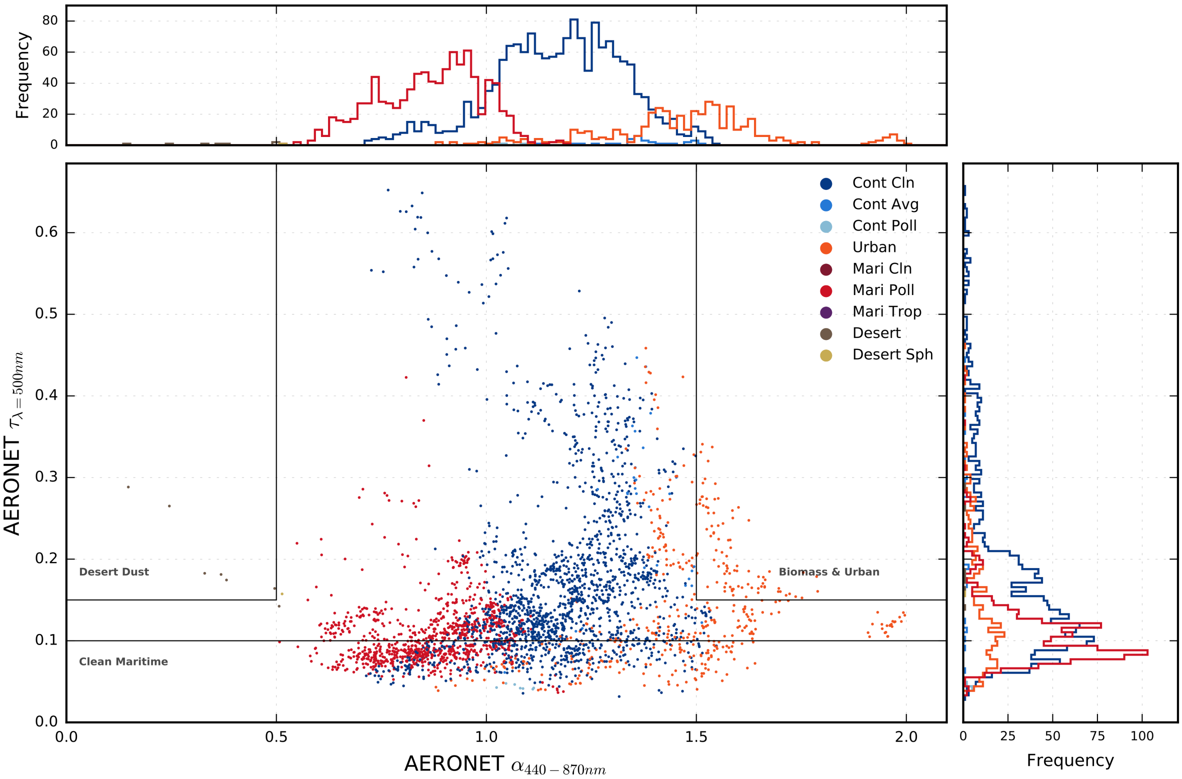

Figure 11: Comparison of the SACS types with the literature classification method based on $\tau_{\lambda=500nm}$ and $\alpha_{\lambda=440-870nm}$, with thresholds proposed by Kaskaoutis et al.[57] The colour illustrates the OPAC aerosol classification determined using SACS. Only showing results for Santiago Chile.

not have co-located MJ CPV power output measurements for any of the five sites, we may compare P_m as estimated by the SACS + CSF scheme against that as estimated by the baseline CAS scheme. The difference between the two gives a first-order indication of the impact of using a more accurate spectral DNI estimation.

Figure 14(a) shows that using the SACS + CSF scheme leads to higher estimated P_m under all circumstances over Niamey, consistent with the large underestimation in BB DNI from CAS at that location (table 4). The modal difference between the two schemes is $10Wm^{-2}$, which is approximately 6% of the average power output. Similarly, the modal difference in P_m is also $10Wm^{-2}$ at Santiago, while the differences over Cape Cod and Manacapuru are generally smaller ($\sim 4Wm^{-2}$), but more varied, reaching up to $\sim 20Wm^{-2}$ at the latter site. Of all five sites, the difference in P_m over Ganges Valley is the smallest, which is not unexpected as the relative bias between CAS and SACS + CSF is small (see figure 9).

6 Conclusions

Four automated schemes to constrain spectral DNI estimations for solar energy applications are developed. The methods handle the spectral signature of aerosols on DNI$_\lambda$ differently but are all based on the use of radiative transfer modelling in conjunction with CAMS/MACC reanalyses and AERONET observations. The schemes developed here ensure that the DNI$_\lambda$ generated is based on realistic estimations of observation geometry, atmospheric state and aerosol characteristics, and can be computed in a fashion that is consistent with the acceptance angle of real systems. The schemes have been tested against ground-based broadband and spectral DNI measurements at five sites that sample contrasting aerosol conditions.

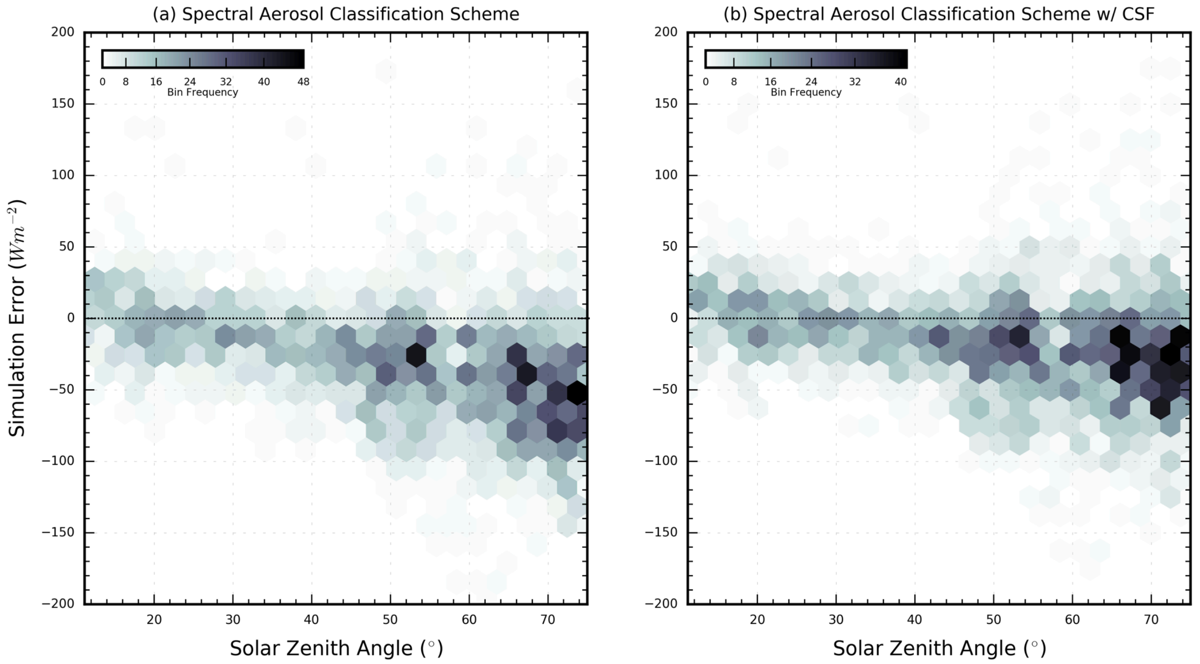

Figure 12: Simulation biases for Santiago Chile (simulated - observed) as a function of solar zenith angle (a) before and (b) after the inclusion of CSR.

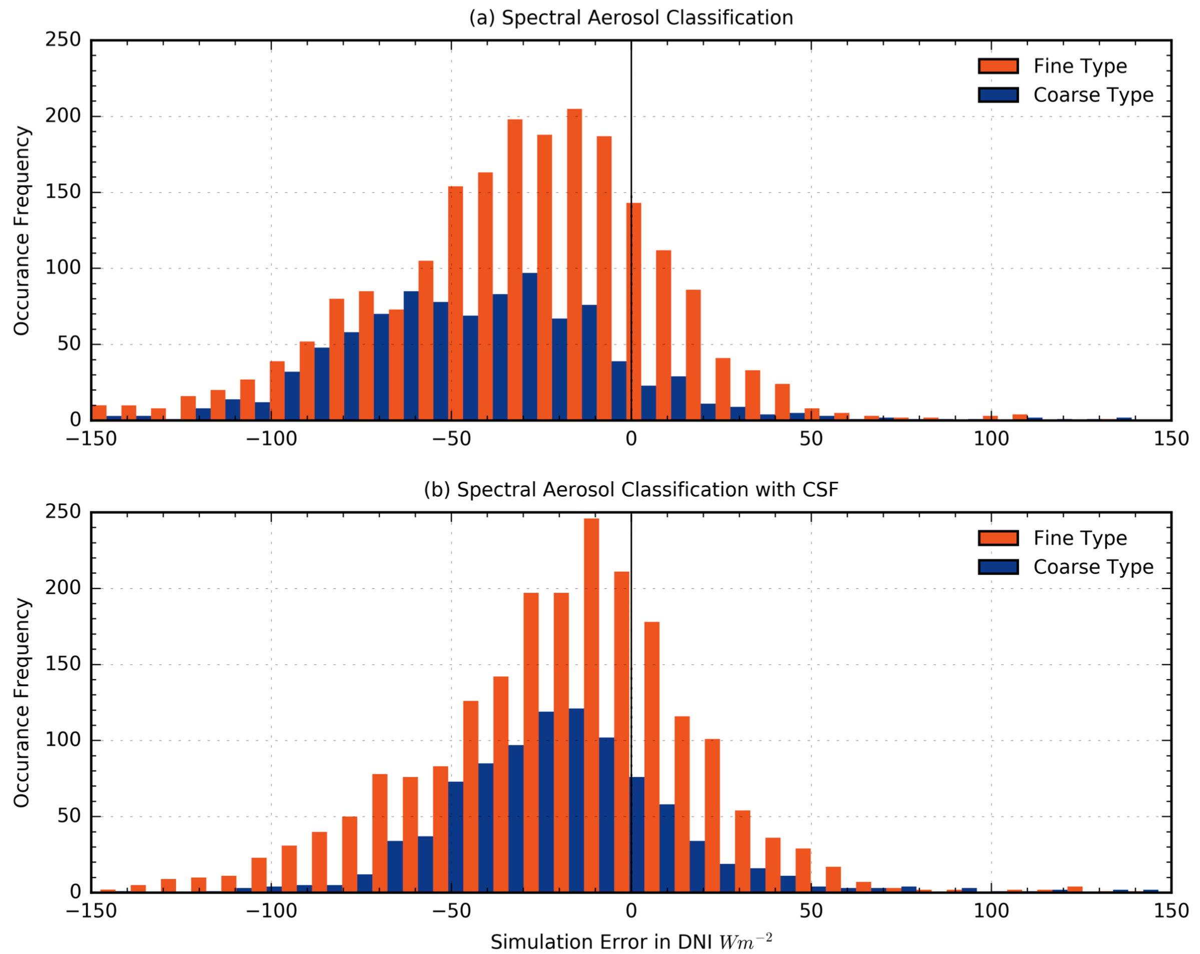

Figure 13: (a) Simulation bias (simulated - observed) distribution for Santiago Chile without CSF, binned according to the aerosol mode. Maritime-like and desert-like OPAC aerosols are classified as coarse, whereas urban and continental-like aerosols are classified as fine. (b) Same as (a), but with the inclusion of CSF.

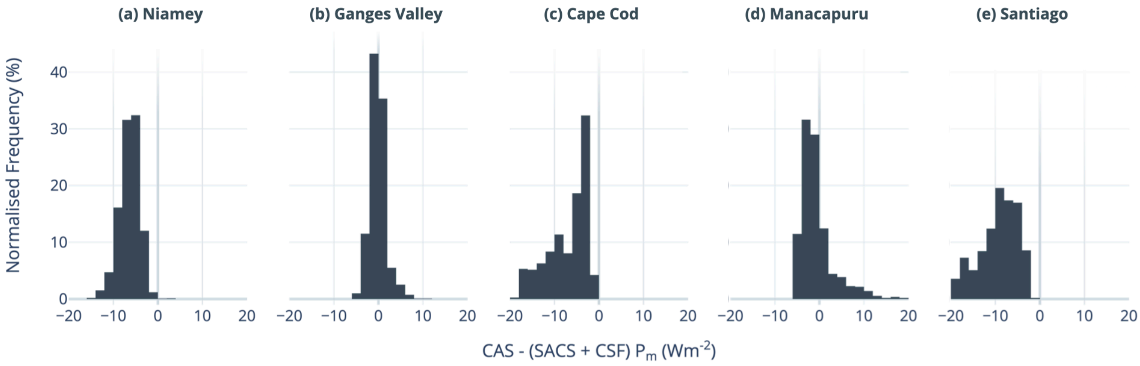

Figure 14: The normalised distribution of the difference in P_m as estimated via CAS and via SACS + CSF for (a) Niamey, Niger, (b) Ganges Valley, India, (c) Cape Cod, USA, (d) Mancapuru, Brazil, (e) Santiago, Chile. Positive differences imply P_m estimated from SACS + CSF is lower.

Amongst the schemes proposed, and for the conditions sampled, the Spectral Aerosol Classification Scheme with Circumsolar Factor (SACS + CSF) shows the closest agreement with observations. Under this scheme, the spectral variation of aerosol extinction is represented by an optimally selected OPAC aerosol type, whilst the circumsolar irradiance is accounted for using a physics-based look-up table. We do note that the merit of the scheme is limited by the representativeness of the available OPAC aerosol types. Nevertheless, using this scheme, the simulated broadband DNI carries a percentage bias of +0.068% when averaged across all five sites, well within the instrument calibration uncertainties of ±3%. The average spread of error over the five sites is about 2.5%. When validating the schemes against spectral observations, SACS + CSF also typically gives the smallest bias across all channels, with the average bias ranging from −0.6% in the $868nm$ channel to +0.3% in the $500nm$ channel. This advantage is retained even for Manacapuru, a site which might be expected to be impacted by biomass aerosol, an aerosol type which is not explicitly included in the OPAC database.

Characterising the aerosol type markedly improves the agreement with both broadband and spectral observations compared to either assuming a climatological aerosol type for a given site or linearly extending the observed Angström exponent. For these comparisons, the inclusion of the CSF is also shown to be a key factor in reducing error in the estimated DNI. It has a distinct spectral imprint, which varies according to aerosol type and appears to be more critical for coarse mode aerosol due to stronger forward scattering. Although the contribution of the CSF will be smaller for the smaller acceptance angles of CPV systems, it has a non-negligible ($> 1\%$) contribution to the DNI at shorter wavelengths, especially under high coarse aerosol loading ($\tau > 1$) and at moderate to high solar zenith angles ($\theta_z > 50$). Estimating the impact of our approach in terms of the power output of a typical 3J solar cell shows that using SACS+CSF as compared to a climatological aerosol baseline typically alters the maximum power point by the order of $10Wm^{-2}$.

There are two novel aspects to SACS + CSF. First, it enables a robust handling of the spectral variation of aerosol extinction, including extrapolation of the spectral variation of aerosol extinction across solar energy-relevant wavelengths that accounts for observational uncertainty. Second, it utilises a simple method to account for the spectral circumsolar contribution to the DNI by employing a look-up table (LUT) computed using a physics-based approach. We make explicit the spectrally resolved treatment of the circumsolar irradiance in the LUT, whereas existing work typically focuses on and validates in the broadband.[79,80] Further, whilst previous work typically only considers desert conditions, we have prepared our LUTs for a wide range of aerosol types, aerosol loading, solar zenith angle and CPV optical tolerances, meaning they can be used globally. Although tested on AERONET observations here, the scheme is designed such that it can easily accommodate satellite retrievals of aerosol optical depth: these are typically only available in a few narrow-band channels and carry relatively high associated uncertainties. The ability of the scheme to accommodate such retrievals is important in the context of generating a more complete global estimate of solar energy resources from real observations.

7 Acknowledgements

This study was part funded through NERCs support of the National Centre for Earth Observation, grant reference: NE/R016518/1

References

[1] Dirnberger, D., Blackburn, G., Muller, B. & Reise, C. On the impact of solar spectral irradiance on the yield of different pv technologies. *Solar Energy Materials and Solar Cells* (2014). URL `https://www.sciencedirect.com/science/article/pii/S0927024814005169`.

[2] Photovoltaic roadmap (itrpv): Eleventh edition online (2020). URL `https://itrpv.vdma.org/viewer/-/v2article/render/48393879`.

[3] WMO. *Guide to meteorological instruments and methods of observation* (World Meteorological Organization, 2008).

[4] Emery, K., Myers, D. & Kurtz, S. What is the appropriate reference spectrum for characterizing concentrator cells? *Conference Record of the Twenty-Ninth IEEE Photovoltaic Specialists Conference, 2002.* (2003).

[5] ASTM. Tables for reference solar spectral irradiances: Direct normal and hemispherical on 37 tilted surface. *ASTM 173-03 (ed. 1)* (2012).

[6] Dominguez, C., Anton, I., Sala, G. & Askins, S. Current-matching estimation for multijunction cells within a cpv module by means of component cells. *Progress in Photovoltaics: Research and Applications* **21**, 14781488 (2012).

[7] Ineichen, P. Validation of models that estimate the clear sky global and beam solar irradiance. *Solar Energy* **132**, 332344 (2016).

[8] Kinsey, G. S. & Edmondson, K. M. Spectral response and energy output of concentrator multijunction solar cells. *Progress in Photovoltaics: Research and Applications* **17**, 279288 (2009).

[9] Philipps, S. *et al.* Energy harvesting efficiency of iiiv triple-junction concentrator solar cells under realistic spectral conditions. *Solar Energy Materials and Solar Cells* **94**, 869877 (2010).

[10] Chan, N. L. A. *et al.* Validation of energy prediction method for a concentrator photovoltaic module in toyohashi japan. *Progress in Photovoltaics: Research and Applications* **21**, 15981610 (2012).

[11] Ghoneim, A. A., Kandil, K. M., Alzanki, T. H. & Alenezi, M. R. Performance analysis of high-concentrated multi-junction solar cells in hot climate. *International Journal of Sustainable Energy* **37**, 294310 (2017).

[12] Rodrigo, P. M., Fernndez, E. F., Almonacid, F. M. & Prez-Higueras, P. J. Quantification of the spectral coupling of atmosphere and photovoltaic system performance: Indexes, methods and impact on energy harvesting. *Solar Energy Materials and Solar Cells* **163**, 7390 (2017).

[13] Marzo, A. *et al.* Standard or local solar spectrum? implications for solar technologies studies in the atacama desert. *Renewable Energy* **127**, 871882 (2018).

[14] Caballero, J. A., Fernandez, E. F., Theristis, M., Almonacid, F. & Nofuentes, G. Spectral corrections based on air mass, aerosol optical depth, and precipitable water for pv performance modeling. *IEEE Journal of Photovoltaics* **8**, 552558 (2018).

[15] Chan, N. L. A., Brindley, H. E. & Ekins-Daukes, N. J. Impact of individual atmospheric parameters on cpv system power, energy yield and cost of energy. *Progress in Photovoltaics: Research and Applications* **22**, 10801095 (2013).

[16] Marzo, A. *et al.* Standard or local solar spectrum? implications for solar technologies studies in the atacama desert. *Renewable Energy* **127**, 871882 (2018).

[17] Blanc, P. *et al.* Direct normal irradiance related definitions and applications: The circumsolar issue. *Solar Energy* **110**, 561577 (2014).

[18] Segal-Rosenheimer, M. *et al.* Retrieval of cirrus properties by sun photometry: A new perspective on an old issue. *Journal of Geophysical Research: Atmospheres* **118**, 45034520 (2013).

[19] Xie, Y. *et al.* A physics-based dni model assessing all-sky circumsolar radiation. *iScience* **23**, 100893 (2020).

[20] SchroedterHomscheidt, M., Oumbe, A., Benedetti, A. & Morcrette, J.-J. Aerosols for concentrating solar electricity production forecasts: requirement quantification and ecmwf/macc aerosol forecast assessment. *Bulletin of the American Meteorological Society* 130109100058001 (2012).

[21] Lefvre, M. *et al.* Mcclear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmospheric Measurement Techniques* **6**, 24032418 (2013).

[22] Morcrette, J.-J. *et al.* Aerosol analysis and forecast in the european centre for medium-range weather forecasts integrated forecast system: Forward modeling. *Journal of Geophysical Research* **114** (2009).

[23] Benedetti, A. *et al.* Aerosol analysis and forecast in the european centre for medium-range weather forecasts integrated forecast system: 2. data assimilation. *Journal of Geophysical Research* **114** (2009).

[24] Holben, B. *et al.* Aeroneta federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment* **66**, 116 (1998).

[25] Mayer, B. & Kylling, A. Technical note: The libradtran software package for radiative transfer calculations description and examples of use. *Atmospheric Chemistry and Physics Discussions* **5**, 13191381 (2005).

[26] Koepke, P. *et al.* Comparison of models used for uv index calculations. *Photochemistry and Photobiology* **67**, 657 (1998).

[27] Weele, M. V. *et al.* From model intercomparison toward benchmark uv spectra for six real atmospheric cases. *Journal of Geophysical Research: Atmospheres* **105**, 49154925 (2000).

[28] Mayer, B., Seckmeyer, G. & Kylling, A. Systematic long-term comparison of spectral uv measurements and uvspec modeling results. *Journal of Geophysical Research: Atmospheres* **102**, 87558767 (1997).

[29] Breitkreuz, H., Schroedter-Homscheidt, M., Holzer-Popp, T. & Dech, S. Short-range direct and diffuse irradiance forecasts for solar energy applications based on aerosol chemical transport and numerical weather modeling. *Journal of Applied Meteorology and Climatology* **48**, 17661779 (2009).

[30] Mueller, R. Rethinking satellite-based solar irradiance modellingthe solis clear-sky module. *Remote Sensing of Environment* **91**, 160174 (2004).

[31] Chance, K. & Kurucz, R. An improved high-resolution solar reference spectrum for earths atmosphere measurements in the ultraviolet, visible, and near infrared. *Journal of Quantitative Spectroscopy and Radiative Transfer* **111**, 12891295 (2010).

[32] Rottman, G. The sorce mission. *The Solar Radiation and Climate Experiment (SORCE)* 725 (2005).

[33] Danielson, J. J. & Gesch, D. B. Global multi-resolution terrain elevation data 2010 (gmted2010). *Open-File Report* (2011).

[34] Fernndez, E. F., Almonacid, F., Ruiz-Arias, J. & Soria-Moya, A. Analysis of the spectral variations on the performance of high concentrator photovoltaic modules operating under different real climate conditions. *Solar Energy Materials and Solar Cells* **127**, 179187 (2014).

[35] Kaskaoutis, D. *et al.* Aerosol properties and radiative forcing over kanpur during severe aerosol loading conditions. *Atmospheric Environment* **79**, 719 (2013).

[36] Kalapureddy, M. C. R. *et al.* Identification of aerosol type over the arabian sea in the premonsoon season during the integrated campaign for aerosols, gases and radiation budget (icarb). *Journal of Geophysical Research* **114** (2009).

[37] Hess, M., Koepke, P. & Schult, I. Optical properties of aerosols and clouds: The software package opac. *Bulletin of the American Meteorological Society* **79**, 831844 (1998).

[38] Alonso-Álvarez, D. *et al.* Solcore: a multi-scale, python-based library for modelling solar cells and semiconductor materials. *Journal of Computational Electronics* **17**, 10991123 (2018).

[39] Dragani, R. On the quality of the era-interim ozone reanalyses. part i comparisons with in situ measurements 19 (2010).

[40] Bock, O. & Nuret, M. Verification of nwp model analyses and radiosonde humidity data with gps precipitable water vapor estimates during amma. *Weather and Forecasting* **24**, 10851101 (2009).

[41] Kinne, S. *et al.* Mac-v1: A new global aerosol climatology for climate studies. *Journal of Advances in Modeling Earth Systems* **5**, 704740 (2013).

[42] Hoyo, M. D., Rondanelli, R. & Escobar, R. Significant decrease of photovoltaic power production by aerosols. the case of santiago de chile. *Renewable Energy* **148**, 11371149 (2020).

[43] Escobar, R. A. *et al.* Solar energy resource assessment in chile: Satellite estimation and ground station measurements. *Renewable Energy* **71**, 324332 (2014).

[44] Mcfarlane, S. A., Kassianov, E. I., Barnard, J., Flynn, C. & Ackerman, T. P. Surface shortwave aerosol radiative forcing during the atmospheric radiation measurement mobile facility deployment in niamey, niger. *Journal of Geophysical Research* **114** (2009).

[45] Slingo, A. *et al.* Overview of observations from the radagast experiment in niamey, niger: Meteorology and thermodynamic variables. *Journal of Geophysical Research* **113** (2008).

[46] Mackie, A., Palmer, P. I. & Brindley, H. Characterizing energy budget variability at a sahelian site: a test of nwp model behaviour. *Atmospheric Chemistry and Physics* **17**, 1509515119 (2017).

[47] Andreas, A. *et al.* Solar infrared radiation station (sirs), sky radiation (skyrad), ground radiation (gndrad), and broadband radiometer station (brs) instrument handbook (2018).

[48] Yin, B., Min, Q. & Joseph, E. Retrievals and uncertainty analysis of aerosol single scattering albedo from mfrsr measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer* **150**, 95106 (2015).

[49] Kotamarthi, V. Ganges valley aerosol experiment: Science and operations plan (2010).

[50] Dumka, U. C., Kaskaoutis, D. G., Srivastava, M. K. & Devara, P. C. S. Scattering and absorption properties of near-surface aerosol over gangetichimalayan region: the role of boundary layer dynamics and long-range transport. *Atmospheric Chemistry and Physics Discussions* **14**, 2110121148 (2014).

[51] Chinnam, N., Dey, S., Tripathi, S. N. & Sharma, M. Dust events in kanpur, northern india: Chemical evidence for source and implications to radiative forcing. *Geophysical Research Letters* **33** (2006).

[52] Machado, L. A. T. Observations and modeling of the green ocean amazon 2014/15. chuva field campaign report (2016).

[53] Goldstein, A. H., Yee, L. D., Issacman-Vanwertz, G. & Wernis, R. A. Green ocean amazon (goamazon) 2014/15. semi-volatile thermal desorption aerosol gas chromatograph (svtag) field campaign report (2016).

[54] Dubey, M., Aiken, A., Berg, L., Freedman, A. & Gorkowski, K. Two-column aerosol project: Aerosol light extinction measurements field campaign report (2016).

[55] Thomalla, E., Kpke, P., Mller, H. & Quenzel, H. Circumsolar radiation calculated for various atmospheric conditions. *Solar Energy* **30**, 575587 (1983).

[56] Smirnov, A., Holben, B., Eck, T., Dubovik, O. & Slutsker, I. Cloud-screening and quality control algorithms for the aeronet database. *Remote Sensing of Environment* **73**, 337349 (2000).

[57] Kaskaoutis, D. G., Kambezidis, H. D., Kharol, S. K. & Badarinath, K. The diffuse-to-global spectral irradiance ratio as a cloud-screening technique for radiometric data. *Journal of Atmospheric and Solar-Terrestrial Physics* **70**, 15971606 (2008).

[58] Kalashnikova, O. V. Ability of multiangle remote sensing observations to identify and distinguish mineral dust types: Optical models and retrievals of optically thick plumes. *Journal of Geophysical Research* **110** (2005).

[59] King, M. D. & Byrne, D. M. A method for inferring total ozone content from the spectral variation of total optical depth obtained with a solar radiometer. *Journal of the Atmospheric Sciences* **33**, 22422251 (1976).

[60] Eck, T. F. *et al.* Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *Journal of Geophysical Research: Atmospheres* **104**, 313Deepak3331349 (1999).

[61] Schuster, G. L., Dubovik, O. & Holben, B. N. Angstrom exponent and bimodal aerosol size distributions. *Journal of Geophysical Research* **111** (2006).

[62] Thomas, G. E. *et al.* Oxford-ral aerosol and cloud (orac): aerosol retrievals from satellite radiometers. *Satellite Aerosol Remote Sensing over Land* 193225 (2009).

[63] Levy, R. C. *et al.* The collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech.* **6**, 2989–3034 (2013).

[64] Reinhardt, B., Buras, R., Bugliaro, L., Wilbert, S. & Mayer, B. Determination of circumsolar radiation from meteosat second generation. *Atmospheric Measurement Techniques* **7**, 823838 (2014).

[65] Sun, Z. *et al.* Determination of direct normal irradiance including circumsolar radiation in climate/nwp models. *Quarterly Journal of the Royal Meteorological Society* **142**, 25912598 (2016).

[66] Prez-Higueras, P., Ferrer-Rodrguez, J. P., Almonacid, F. & Fernndez, E. F. Efficiency and acceptance angle of high concentrator photovoltaic modules: Current status and indoor measurements (2018). URL `https://www.sciencedirect.com/science/article/pii/S1364032118304453`.

[67] Box, M. A. & Deepak, A. Atmospheric scattering corrections to solar radiometry. *Applied Optics* **18**, 1941 (1979).

[68] Gueymard, C. A., Bett, A. W., Mcconnell, R. D., Sala, G. & Dimroth, F. Spectral circumsolar radiation contribution to cpv. *AIP Conference Proceedings* **1277**, 316319 (2010).

[69] Eissa, Y., Blanc, P., Wald, L. & Ghedira, H. Can aeronet data be used to accurately model the monochromatic beam and circumsolar irradiances under cloud-free conditions in desert environment? *Atmospheric Measurement Techniques* **8**, 50995112 (2015).

[70] Zhang, F., Liu, K., Yang, Q., Wu, K. & Zhao, J.-Q. Comparison of chebyshev and legendre polynomial expansion of phase function of cloud and aerosol particles. *Advances in Meteorology* **2017**, 110 (2017).

[71] Sinyuk, A. *et al.* Assessment of error in aerosol optical depth measured by aeronet due to aerosol forward scattering. *Geophysical Research Letters* **39** (2012).

[72] Anderson & P., G. Afgl atmospheric constituent profiles (0-120km) (1986). URL `http://adsabs.harvard.edu/abs/1986aacp.book.....A`.

[73] Bird, R. & Riordan, C. Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earths surface for cloudless atmospheres (1984).

[74] Askar, H. & Batty, W. Assessing the performance of the simple model of the atmospheric radiative transfer of sunshine (smarts2) in a first tier of software using empirical weather data. *Energy* **30**, 15771588 (2005).

[75] Madkour, M., El-Metwally, M. & Hamed, A. Comparative study on different models for estimation of direct normal irradiance (dni) over egypt atmosphere. *Renewable Energy* **31**, 361382 (2006).

[76] Ruiz-Arias, J. A., Dudhia, J., Santos-Alamillos, F. J. & Pozo-Vazquez, D. Surface clear-sky shortwave radiative closure intercomparisons in the weather research and forecasting model. *Journal of Geophysical Research: Atmospheres* **118**, 99019913 (2013).

[77] Shi, G., Sun, Z., Li, J. & He, Y. Fast scheme for determination of direct normal irradiance. part i: New aerosol parameterization and performance assessment. *Solar Energy* **199**, 268277 (2020).

[78] Stein, A. F. *et al.* Noaa's hysplit atmospheric transport and dispersion modeling system. *Bulletin of the American Meteorological Society* (2015). URL `https://journals.ametsoc.org/bams/article/96/12/2059/69294/NOAA-s-HYSPLIT-Atmospheric-Transport-and`.

[79] Gueymard, C. A. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. *Solar Energy* **86**, 21452169 (2012).

[80] Ruiz-Arias, J. A. & Gueymard, C. A. Worldwide inter-comparison of clear-sky solar radiation models: Consensus-based review of direct and global irradiance components simulated at the earth surface. *Solar Energy* **168**, 1029 (2018).