Solar Energy 224 (2021) 1063–1078

Contents lists available at ScienceDirect

Solar Energy

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

Worldwide performance evaluation of ground surface reflectance models

Arttu Tuomiranta [a,b,c], Pierre-Jean Alet [b], Christophe Ballif [b,c,*], Hosni Ghedira [a,d]

[a] Research Center for Renewable Energy Mapping and Assessment, Khalifa University, Masdar City, PO Box 127788, Abu Dhabi, United Arab Emirates
[b] CSEM PV-center, Rue Jaquet-Droz 1, 2002 Neuchâtel, Switzerland
[c] École polytechnique fédérale de Lausanne (EPFL), Photovoltaics and Thin-Film Electronics Laboratory, Rue de la Maladière 71b, 2002 Neuchâtel, Switzerland
[d] Mohamed bin Zayed University of Artificial Intelligence, Masdar City, Abu Dhabi, United Arab Emirates

ARTICLE INFO

Keywords:
Model validation
Albedo
Bi-hemispherical reflectance
Time-variant models
Solar elevation
Bifacial photovoltaics

ABSTRACT

To improve the accuracy of bifacial gain estimation, recent radiative models of solar energy systems have abandoned the traditional assumption of isotropic ground-reflected radiance. However, surface reflectance itself is still commonly considered as a constant — partly because of the recommendations of previous evaluations of reflectance models. This paper presents the findings of a new model evaluation based on a large database of measurements from 26 sites, which are representative of major land covers and climates. Both novel and previously reviewed formulations are validated with the data. On a global average, data-based estimation reduces mean absolute error by 22%, 29%, and 39% with constant, univariate, and bivariate models, respectively, compared to literature-based estimates. Only at the urban and snow-and-ice sites does time-variant estimation not notably improve accuracy. Arid sites tend to favour univariate models based on solar elevation, and diffuse fraction adds little value as the second predictor. By contrast, bivariate estimation clearly improves accuracy at vegetated and water sites. When considering the best-performing model for each site, the global average mean absolute error is 11%. Two novel formulations, univariate and bivariate, provide superior performance at many sites. The proposed 3-parameter bivariate model is one of the top performers at 19 out of the 26 considered sites.

1. Introduction

The emergence of bifacial photovoltaics has motivated a variety of recent studies on ground-reflected irradiance estimation. The traditional modelling approach based on the assumption of spatially uniform ground radiosity (Moon and Spencer, 1942; Liu and Jordan, 1960) has been replaced by models that capture the spatial distribution of irradiance incident on the ground surface surrounding the system (e.g., Yusufoglu et al., 2015; Mermoud and Wittmer, 2016; Hansen et al., 2017; Ziar et al., 2019). This new generation of models grid the surface into unshaded and shaded, internally homogeneous patches with varying levels of incident irradiance. Thereby, a more realistic, anisotropic construction of ground-reflected radiance distribution at the plane of the array becomes possible. However, it is still common when modelling photovoltaic (PV) system performance to assume that for any point on the ground the surface reflectance (ρ) i.e., the ratio of the reflected irradiance to the total incident irradiance is a constant over time. In other words, the anisotropy of ground-reflected radiance is assumed to be only caused by the spatial variability of ground-incident irradiance, and not by the optical properties of the surface itself (with exceptions to this including, e.g., Chiodetti et al. (2016)).

Despite the abundance of publicly available in-situ measurements, only two previous studies have evaluated formulations for ρ estimation (Ineichen et al., 1990; Psiloglou and Kambezidis, 2009). Their

Abbreviations: BSRN, Baseline Surface Radiation Network; cRMSE, centred RMSE; D, Kolmogorov–Smirnov statistic; d, diffuse irradiance fraction; G, downwelling i.e., global horizontal irradiance; $G_{b,n}$, beam normal irradiance; G_d, diffuse horizontal sky irradiance; G_r, upwelling i.e., ground-reflected horizontal irradiance; IGBP, International Geosphere-Biosphere Programme; MAE, mean absolute error; MBE, mean bias error; PV, photovoltaic; r, Pearson correlation coefficient; rMAE, relative MAE; RMSE, root-mean-square error; SURFRAD, Surface Radiation Budget Network; θ_z, solar zenith angle; ρ, global bi-hemispherical reflectance aka blue-sky albedo; ρ_{60°, ρ at $\theta_z = \pi/3$; ρ_b, directional-hemispherical reflectance aka black-sky albedo; $\rho_{b,60^\circ}$, ρ_b at $\theta_z = \pi/3$; $\rho_{b,n}$, ρ_b at $\theta_z = 0$; ρ_d, diffuse bi-hemispherical reflectance aka white-sky albedo; ρ_n, ρ at $\theta_z = 0$; σ, standard deviation.

* Corresponding author at: École polytechnique fédérale de Lausanne (EPFL), Photovoltaics and Thin-Film Electronics Laboratory, Rue de la Maladière 71b, 2002 Neuchâtel, Switzerland.

E-mail addresses: arttu.tuomiranta@gmail.com (A. Tuomiranta), pierre-jean.alet@csem.ch (P.-J. Alet), christophe.ballif@epfl.ch (C. Ballif), hosni.ghedira@mbzuai.ac.ae (H. Ghedira).

https://doi.org/10.1016/j.solener.2021.06.023
Received 27 April 2021; Received in revised form 6 June 2021; Accepted 8 June 2021
Available online 2 July 2021

recommendation for the use of constant ρ is somewhat contradictory with the findings of the numerous published validations of satellite-data-based ρ models (reviewed by, e.g., Yang et al., 2008; Gueymard et al., 2019). Indeed, no natural surface is a perfect diffuse reflector i.e., perfectly Lambertian: the direction of the incident ray of light influences the angular distribution of reflected radiance. This effect results in temporally variable ρ under the field conditions of solar energy systems.

Fig. 1 illustrates the impact of an assumed constant ρ on the accuracy of bifacial PV power estimation at two example sites: a low-latitude desert site and a temperate vegetated site with yearly mean ρ of 0.33 and 0.23, respectively. Fig. 1a shows that a deviation of 0.01 in the assumed ρ from the yearly mean results in an error of 0.3 % in the energy yield. For instance, a plausible error of 0.05 in ρ can, therefore, lead to a deviation of 1.5%-1.6 % in the levelised cost of electricity estimated for a bifacial PV system at the example sites. In the currently common project scale of hundreds of megawatts, this deviation represents millions of dollars worth of revenue over the project lifetime. The error can also trigger the decision to switch between bifacial and monofacial modules.

Fig. 1b shows the effect of the assumed constant ρ on the mean-normalised root-mean-square error (RMSE) of PV power estimates. At their best, the constant estimators can reach an RMSE of 1%-3%. The remaining error is due to the temporal variability of ρ. Although the integrated yearly contribution of the temporal variability appears comparatively low, its consideration can notably improve the instantaneous accuracy of system performance simulation as shown in Fig. 2. The figure shows time series of standard system performance ratio (i.e., capacity factor normalised by front in-plane irradiance) for the system at the desert site in a summer week in 2013. Performance ratio is commonly used for detecting faults and other system anomalies. The instantaneous error in performance ratio caused by considering the yearly mean ρ instead of the real, temporally variable ρ can reach 20%.

The twofold objective of the study reported in this paper is, therefore, 1) to propose novel formulations for modelling ρ and 2) to evaluate both the novel and previously published models on the basis of measurements from a large variety of sites which represent different surface and climate types. Ultimately, the study aims to improve the understanding of the suitability of the various ρ models for each type of site.

Section 2 introduces all the models considered in the evaluation. The data and methods are described in Section 3 and Section 4 presents the results.

2. Models

This section describes all the models evaluated in the present study. Subsection 2.1 briefly reviews the formulations available in the preceding literature, and subsection 2.2 proposes novel adaptations that aim to address gaps and shortcomings identified in the state of the art.

2.1. Prior art

Table 1 lists the references, abbreviations, and mathematical forms of published formulations for ρ. The formulations are sorted in the order of increasing complexity. All the 20 reviewed models are considered in the evaluation reported in this paper. The models are divided into five groups: literature-based constants, data-based constants, univariate functions of solar zenith angle (θ_z), univariate functions of diffuse irradiance fraction (d), and bivariate functions of both θ_z and d. The various symbols used in the formulations of Table 1 are described in Table 2. The parameters of the models can be classified into base ρ instances, which can be interpreted as colour parameters of the models, and into fitting coefficients, mostly denoted as b, which are supposed to reflect Lambertianity and spectral effects in most of the models.

The first three models, M0_1, M0_2, and M0_3, are particularly simple to use as they do not require any measurements. M0_1 refers to the constant ρ of 0.2 and M0_2 to a value derived from the albedo library of Gueymard (2005) based on the type of site. This library actually specifies black-sky albedos $\rho_b(\theta_z = 20°)$ for snow and $\rho_b(\theta_z = 53°)$ for other surface types. M0_3 involves a look-up table of month-specific coefficients of a third-degree polynomial fit of ρ to site latitude. The following two models represent the constants derived out of the measured data. M0_4 stands for the arithmetic mean and M0_5 for the geometric mean.

The univariate models (#6–10 in Table 1) can be divided into two categories: functions of θ_z and those of d. θ_z and d are known to be correlated (e.g., Liu and Jordan, 1960) and, therefore, the empirical models involving either one of the two capture part of both effects. However, since d and θ_z are not proportional to each other, they are not interchangeable as the predictors of ρ. The models which are functions of θ_z offer the simplest means of estimating continuously variable ρ as the only measured parameters required for their calibration are global downwelling horizontal irradiance (G) and upwelling i.e., ground-reflected horizontal irradiance (G_r). Apart from Mz_8, they involve ρ_n, whose direct measurement is only possible at sites located between the tropics. Hence, ρ_n is for most sites a fitting coefficient rather than a physical constant. $\rho_{60°}$ can be directly measured at many more sites than ρ_n. At those sites, Mz_8, which is based on the measured mean of $\rho_{60°}$, can be considered to have only one fitting coefficient. Mz_9 refers to the purely empirical extension of Mz_7 with an additional constant term. As

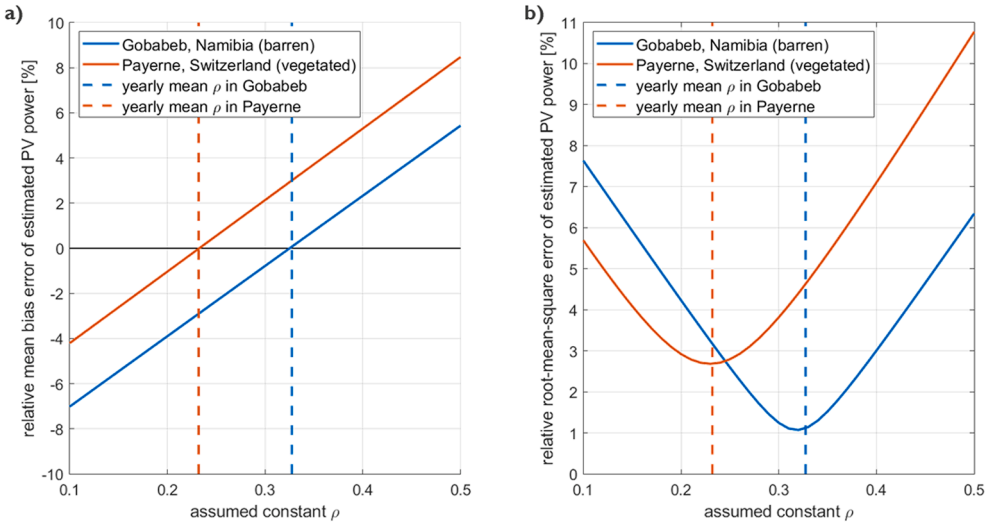

Fig. 1. Dependence of the accuracy of bifacial photovoltaic (PV) power estimation on the assumed constant ground reflectance (ρ) in terms of a) relative mean bias error (i.e., error of energy yield estimation) and b) mean-normalised root-mean-square error at PV systems located at a desert site in Gobabeb, Namibia and at a vegetated site in Payerne, Switzerland. (The PV systems were assumed to be equipped with polar-aligned modules with 80-% bifaciality coefficients at a ground clearance height of 2 metres and a ground coverage ratio of 0.3. Bifacial PV power was simulated by System Advisor Model (Gilman et al., 2018) based on in-situ measurements from Gobabeb (Vogt, 2019) and Payerne (Vuilleumier and Heimo, 2019)).

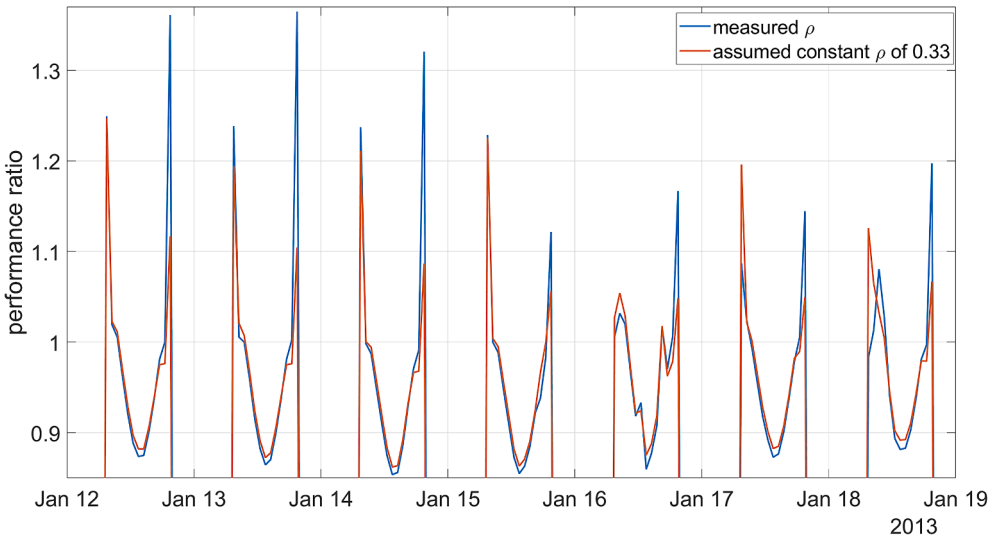

Fig. 2. Performance ratio estimated based on front in-plane irradiance for a bifacial photovoltaic system in Gobabeb, Namibia in a summer week in 2013 considering measured ground reflectance (ρ) data (blue) and the constant yearly mean ρ of 0.33 (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Overview of the models considered for evaluation.

#	Reference	Abbr.	Equation
1	Liu and Jordan (1963)	M0_1	0.2
2	Gueymard (2005)	M0_2	library constant
3	Gueymard (1993)	M0_3	$\sum_{j=0}^{3} b_j(\text{month})\phi^j$
4	Psiloglou and Kambezidis (2009)	M0_4	$\sum_{i=1}^{N}\rho_i/N$
5	Psiloglou and Kambezidis (2009)	M0_5	$\sqrt[N]{\prod_{i=1}^{N}\rho_i}$
6	Temps and Coulson (1977)	Mz_6	$\rho_n(1+\frac{1-\cos\theta_z}{2})$
7	Nkemdirim (1972)	Mz_7	$\rho_n e^{b\theta_z}$
8	Dickinson (1983)	Mz_8	$\rho_{60^\circ}\frac{1+b}{1+2b\cos\theta_z}$
9	Psiloglou and Kambezidis (2009)	Mz_9	$b_0+b_1 e^{b_2\theta_z}$
10	Ineichen et al. (1990)	Md_10	$(1-d)\rho_b+d\rho_d$
11	Gardner and Nadeau (1988)	M2_11	$(1-d)\left(\frac{1-\cos\theta_z}{2}\cos\theta_z+(1-\frac{1-\cos\theta_z}{2})\rho_n\right)+d\rho_n$
12	Ross (1981)	M2_12	$(1-b)\left(\frac{1-d}{1+2b\cos\theta_z}+\frac{d}{b^2}(b-\frac{\ln(1+2b)}{2})\right)$
13	Schlick (1994) & Gueymard (1987)	M2_13	$(1-d)(\rho_{b,n}+(1-\rho_{b,n})(1-\cos\theta_z)^5)+d\rho_d$
14	Larsen and Barkstrom (1977) & Gueymard (1995)	M2_14	$\rho_d(b(1-d)(1-\cos\theta_z\ln(1+1/\cos\theta_z))+d)$
15	Hou et al. (2002)	M2_15	$\rho_d\left(\left(1-d\right)\frac{1+b}{1+2b\cos\theta_z}+d\right)$
16	Schlick (1994) & Gardner and Nadeau (1988)	M2_16	$(1-d)\left(\rho_n+\frac{1-\cos\theta_z}{2}\left((\rho_{b,n}+(1-\rho_{b,n})(1-\cos\theta_z)^5)\cos\theta_z-\rho_n\right)\right)+d\rho_n$
17	Gueymard (1987)	M2_17	$(1-d)(\rho_n+(1-\rho_n)e^{b(\pi/2-\theta_z)})+d\rho_d$
18	Larsen and Barkstrom (1977) & Gueymard (1995)	M2_18	$\rho_d(b_1(1-d)(1+b_2\cos\theta_z)+d)$
19	Wang et al. (2007) & Chiodetti et al. (2016)	M2_19	$(1-d)\rho_{b,60^\circ}\frac{1+b}{1+2b\cos\theta_z}+d\rho_d$
20	Gueymard (1987)	M2_20	$(1-d)(\rho_n+f_{fs}e^{b_0+b_1\theta_z+b_2\theta_z^2})+d(\rho_n+0.023(f_{fs}+f_{bs}))$

opposed to the other considered univariate models, the calibration of d-based models requires the measurements of either diffuse horizontal sky irradiance (G_d) or beam normal irradiance ($G_{b,n}$) in addition to G and G_r. This category is only represented by Md_10, which involves two constants: the physically sound ρ_d and the rather statistical ρ_b.

The bivariate models (#11–20 in Table 1) involve both θ_z and d as independent variables. Models M2_11-15 involve at most one and models M2_16-19 two fitting coefficients. The most complex model is M2_20 with six coefficients at maximum.

2.2. Novel Formulations

The present evaluation covers six models that are novel to ρ estimation. Two of them are simple constants. First, while arithmetic mean (M0_4) is the minimum mean square error estimator, the median minimises mean absolute error (MAE) for any dataset. That is why in this study, median is also considered as an option for data-based ρ estimation (model M0_21). Additionally, a variant of M0_3 corrected by the arithmetic mean is evaluated (model M0_22).

The new model Mz_23 aims at improving Mz_8. The most commonly cited versions of Mz_8 define it as a function of ρ_{60°. The choice originates from the white-sky albedo approximation, $\rho_d \approx \rho_{b,60^\circ}$, proposed by

Table 2
Nomenclature of Table 1.

Symbol	Description
b	fitting coefficient
d	diffuse irradiance fraction
f_{bs}	M2_20 backscatter coefficient
f_{fs}	M2_20 forward scatter coefficient
i	time step index
j	general index
N	number of data records
θ_z	solar zenith angle
ρ	global bi-hemispherical reflectance (blue-sky albedo)
ρ_{60°	ρ at $\theta_z = \pi/3$
ρ_b	directional-hemispherical reflectance (black-sky albedo)
$\rho_{b,60^\circ}$	ρ_b at $\theta_z = \pi/3$
$\rho_{b,n}$	ρ_b at $\theta_z = 0$
ρ_d	diffuse bi-hemispherical reflectance (white-sky albedo)
ρ_i	ρ at time step i
ρ_n	ρ at $\theta_z = 0$
ϕ	latitude

Dickinson (1983). However, ρ_d can be derived as roughly equal to ρ_{50° when applying Mz_8 to isotropic sky radiance. This isotropy was assumed when deriving Mz_8 itself. Hence, the approximation results in an internal contradiction in the model, which makes ρ_{60° a problematic choice as the fixed base ρ instance. Model (Mz_23) uses ρ_n instead of ρ_{60° effectively setting a minimum level for ρ. Although ρ_n does not solve the problem of potential d sensitivity, its maximum value is physically bound by ρ_d, which may increase the stability of the model under real conditions.

$$\rho = \rho_n \frac{1+b}{1+b\cos\theta_z}, b \in [0,2], \quad \text{(Mz23)}$$

where b is a coefficient increasing with the non-Lambertian behaviour of the surface.

The new models M2_24 and M2_25 start from the M2_11 formulation. M2_11 decomposes beam reflectance into purely specular i.e., bidirectional and diffuse i.e., directional-hemispherical components by utilising the anisotropy factor proposed by Temps and Coulson (1977) as a weighting function. The specular component is calculated by multiplying the specularity weight function ($(1-\cos\theta_z)/2$) by the beam horizontal irradiance fraction ($1-d$) and the transposition factor of the receiving surface ($\cos\theta_z$). However, when striking the ground, a specularly reflected ray is attenuated only through absorption, which is considered by means of ρ. The cosine effect does not attenuate the ray as it is reflected directionally. Therefore, the use of the multiplier $(1-d)\cos\theta_z$ is questionable as the intensity of an incident ray is $G_{b,n} = G_b/\cos\theta_z = (1-d)G/\cos\theta_z$ rather than its horizontally transposed counterpart, G_b. In other words, model M2_11 appears to curtail its specular term with a redundant ground transposition factor $\cos\theta_z$. That is why we introduced model (M2_24) where this term is divided by $\cos\theta_z$.

$$\rho = (1-d)\left(\frac{1-\cos\theta_z}{2} + (1-\frac{1-\cos\theta_z}{2})\rho_n\right) + d\rho_n \quad \text{(M224)}$$

By substituting the specular term of M2_11 with Schlick's approximation (Schlick, 1994), one obtains a second novel formulation, M2_25, specified in (3).

$$\rho = (1-d)\left(\rho_n + \frac{1-\cos\theta_z}{2}\left((1-b\rho_n)(1-\cos\theta_z)^5 - \rho_n(1-b)\right)\right) + d\rho_n, b \in [0,1] \quad \text{(M225)}$$

In the adaptation, $\rho_{b,n}$, which is immeasurable with conventional instruments for ρ monitoring, is additionally replaced by $b\rho_n, b \in [0,1]$ on the basis that $\rho_{b,n} \leqslant \rho_n \leqslant \rho_d$. In the present study, for the sake of consistency, the substitutions $\rho_{b,n} = b\rho_d, b \in [0,1]$ and $\rho_{b,n} = b\rho_n, b \in [0,1]$ are also made in M2_13 and M2_16, respectively. In this paper, b in the subscript always refers to *beam* whereas b is used as a general symbol for coefficients.

Finally, the new model M2_26 starts from the M2_19 formulation. M2_19 uses $\rho_{b,60^\circ}$ as the base ρ instance of its beam term. The choice can be traced back to Mz_8 and the $\rho_d \approx \rho_{b,60^\circ}$ approximation of Dickinson (1983). In the case of a univariate model, it is well-justified to use a base ρ value whose covariance with d is as low as possible since in that way, the physically less meaningful, empirically derived coefficients integrate the maximum share of d-induced variability. Bivariate models, however, separate the beam and diffuse components and weight the corresponding terms with $(1-d)$ and d, respectively. Hence, the use of $\rho_{b,60^\circ}$ or ρ_d as the fixed base of the beam term cannot be similarly justified. As the (realistic) minimum ρ, ρ_n may offer a better base for the beam component. As discussed above, it is capped by ρ_d to prevent the fitting procedure from setting the value unrealistically high. The use of bivariate models requires d measurements which enable the evaluation of ρ_d. The novel bivariate model (M2_26) is accordingly based on the ρ_n-based formulation of M2_26.

$$\rho = (1-d)\rho_n \frac{1+b}{1+b\cos\theta_z} + d\rho_d, b \in [0,2] \quad \text{(M226)}$$

3. Methodology

This section describes the utilised data and the methods deployed to process and analyse them. The sources and scope of the data are described in subsection 3.1. Subsection 3.2 explains the post-processing steps undertaken to merge the time series of in-situ measurements into a single database of quality-assured observations. A general overview of the database is given in subsection 3.3. The objective of this study, model evaluation, was achieved through cross-validation. The details of the adopted technique are given in subsection 3.4.

3.1. Data Sources

The analysis reported in this paper is based on ground measurements made at the 26 sites specified in Table 3. The geographical distribution of the sites is shown for North America in Fig. 3 and for the rest of the world in Fig. 4. The sites were chosen from the latest snapshot databases of the Baseline Surface Radiation Network (BSRN) of the World Radiation Monitoring Center (König-Langlo et al., 2015; Driemel et al., 2018) and the Surface Radiation Budget Network (SURFRAD) of the National Oceanic and Atmospheric Administration of the United States Department of Commerce (Augustine et al., 2000; Augustine et al., 2005). The selection was limited to the sites with decomposed irradiance (G_d and/or $G_{b,n}$) measurements. All the considered BSRN and SURFRAD data are available at a resolution of 1 to 5 min. The ground surface types were categorised as per the land cover classification system of the International Geosphere-Biosphere Programme (IGBP). The sites with the land cover type "Snow and ice" have continuously snow or ice-covered ground surfaces while the other sites may also experience temporary periods of snow coverage. The sites can also be divided into climatic groups. The climatic division specified in Table 3 is based on the Köppen climate classification. In this classification system, climate types are divided into five main groups indicated by the capital letters A-E: tropical (A), dry (B), temperate (C), continental (D), and polar (E). In the three-letter class abbreviations, the second and third letters refer to the typical levels of precipitation and temperature, respectively. As a whole, the considered sites represent 10 land cover types and 11 climates. As a result, 19 different surface-climate combinations were considered in the study. Land cover type can be expected to be the primary factor when classifying sites into different ground reflectance categories. However, because of its link to the type and seasonality of the local vegetation, climate should also be regarded as a potential factor.

Table 3
Overview of the measurement sites.

#	Site	Coordinates	Land cover type (IGBP)	Climate type (Köppen)
1	Goodwin Creek, US-MS	34.25°N,89.87°W 98 m AMSL	Woody savannas	Cfa
2	Rutland, US-VT	43.64°N,72.97°W 184 m AMSL	Savannas	Dfb
3	Boulder, US-CO	40.13°N,105.24°W 1689 m AMSL	Grasslands	BSk
4	Fort Peck, US-MT	48.31°N,105.10°W 634 m AMSL		
5	Table Mountain, US-CO	40.12°N,105.00°W 1689 m AMSL		
6	Töravere, Estonia	58.25°N,26.46°E 70 m AMSL		Dfb
7	Alamosa, US-CO	37.7°N,105.92°W 2317 m AMSL	Croplands	BSk
8	Wasco, US-OR	45.59°N,120.67°W 200 m AMSL		Csb
9	Bondville, US-IL	40.05°N,88.37°W 213 m AMSL		Cfa
10	Penn State, US-PA	40.72°N,77.93°W 376 m AMSL		Cfb
11	Sioux Falls, US-SD	43.73°N,96.62°W 473 m AMSL	Croplands	Dfa
12	Tateno, Japan	36.06°N,140.13°E 25 m AMSL	Cropland/ natural vegetation mosaics	Cfa
13	Cabauw, Netherlands	51.97°N,4.93°E 0 m AMSL		Cfb
14	Payerne, Switzerland	46.82°N,6.94°E 491 m AMSL		
16	Tiksi, Russia	71.59°N,128.92°E 48 m AMSL	Permanent wetlands	Dsd
17	Barrow, US-AK	71.32°N,156.61°W 8 m AMSL		ET
18	Ny-Ålesund, Norway	78.93°N,11.93°E 11 m AMSL		
19	Gobabeb, Namibia	23.56°S,15.04°E 407 m AMSL	Barren	BWh
20	Red Lake, US-AZ	35.66°N,114.07°W 843 m AMSL		BWh
21	Desert Rock, US-NV	36.62°N,116.02°W 1007 m AMSL		BWk
22	Alert, Canada	82.49°N,62.42°W 127 m AMSL		ET
23	Neumayer, Antarctica	70.65°S,8.25°W 42 m AMSL	Snow and ice	EF
24	Syowa, Antarctica	69.01°S,39.59°E 18 m AMSL		
25	South Pole, Antarctica	89.98°S,24.80°W 2800 m AMSL		
26	Eugene, US-OR	44.05°N,123.07°W 150 m AMSL	Urban and built-up	Csb
27	Chesapeake Light, US-VA	36.91°N,75.71°W 37 m AMSL	Water bodies	Cfa

3.2. Data Processing

Parallel time series of solar geometric parameters were first generated to complement the time series of G, G_r, G_d and/or $G_{b,n}$, and basic meteorological parameters retrieved from the BSRN and SURFRAD databases. In this paper, the term *measured* ρ is used to refer to the ratio of the measured G_r to the measured G. Similarly, *measured* d denotes either the ratio of the measured G_d to the measured G or one minus the ratio of the measured $G_{b,n}$ to the measured G multiplied by the cosine of the computed θ_z. The chosen d estimation method depends on the availability of the measurements of G_d and $G_{b,n}$ at each time step. Solar geometry was calculated by means of the SG2 algorithm of Blanc and Wald (2012). The quality of the measurements was assessed based on the criteria proposed by Espinar et al. (2011). The time steps, at which one or more of the parameters did not meet the quality control criteria for sub-hourly extrema, were omitted. The quality of G_r, which is not addressed by Espinar et al. (2011), was simply assured by excluding any remaining time steps, at which $G_r \leqslant 0$ or $G_r > G$.

Snow coverage drastically intensifies the reflectance of a natural ground surface. However, with the exception of M0_3 and M0_22, the considered models do not address the seasonal variation between snow-free and snowy conditions. Hence, model evaluation under all-ground conditions might result in poor performance indicators of a model that would perform well when the surface of a site is snow-free. That is why we chose to test the models separately for periods with and without snow coverage. To determine these periods, the presence of snow had to be determined indirectly because neither databases contain direct information about snow coverage.

Since some of the covered sites do not have continuous temperature records available, we used daily mean ρ as a proxy for snow coverage. Indeed, snow has a significantly higher ground reflectance than most other surfaces. Fig. 5 shows the daily means of ρ for two seasonally snowy sites over the entire timeframes of each site's database. The figure shows numerous days when the daily mean ρ is somewhere between the presumed ρ levels of fresh snow and the snow-free surface type.

The seasonal variability in daily mean ρ is primarily due to solar elevation and both natural and anthropogenic changes in vegetation. Fig. 5 shows that the pattern of seasonal variability varies greatly with the surface type. Hence, the maximum level for a daily mean ρ of snow-free ground surface will also depend on the surface type. While 0.2 would appear as a good choice for the ground surface of Ny-Ålesund, a permanent wetland site, the maximum level of 0.3 would be more appropriate for the surface of Bondville, a highly variable cropland site. Analysing the daily mean ρ values has shown that 19 out of the 26 considered sites experience both snowy and snow-free periods. We divided the 19 sites into one very low-ρ site (urban and built-up), five low-ρ sites, ten medium-ρ sites, and three high-ρ sites with respective snow-free daily mean ρ thresholds of 0.15, 0.2, 0.25, and 0.3. The data points with complete snow coverage were identified with a minimum daily mean ρ of 0.7. Since ρ of snow decreases when the snow is ageing or melting (e.g., Hall, 1985, p. 29), snowmelt may result in a significant reduction in daily mean ρ even if the surface is still entirely covered by snow. Hence, most of the data points with an aged or melting, yet complete snow coverage were likely to be excluded from this set. The snow-free thresholds are listed for each site in Table 4. The table also specifies the covered years and the data references divided into BSRN (various references) and SURFRAD (Augustine et al., 2000; Augustine et al., 2005) sources. The sites are shown in the same order as in Table 3.

3.3. Database Overview

Fig. 6 shows the numbers of quality-assured daytime data points for each site and for each of the three categories of snow coverage. Red Lake has the smallest database with its 9868 data points considered in the analysis. Fig. 7 illustrates the variability of the filtered ρ measurements. This variability is primarily due to diurnal and seasonal changes in solar elevation, land cover, and diffuse irradiance fraction. In the case of multi-annual time series, inter-annual variability — particularly in land cover — can have an additional impact.

Fig. 8 shows the mean-normalised standard deviation (σ) of the measured monthly mean ρ for each site. In relative terms, four very different sites, Wasco, Alert, Eugene, and Chesapeake Light, stand out with their high variability. Desert and snow-and-ice sites, by contrast, appear to have comparatively low seasonal variability levels. The average normalised σ across all 26 sites is 6.7%. Part of this variability is due to changes in θ_z and/or d — the changes whose impact the reviewed models aim to capture.

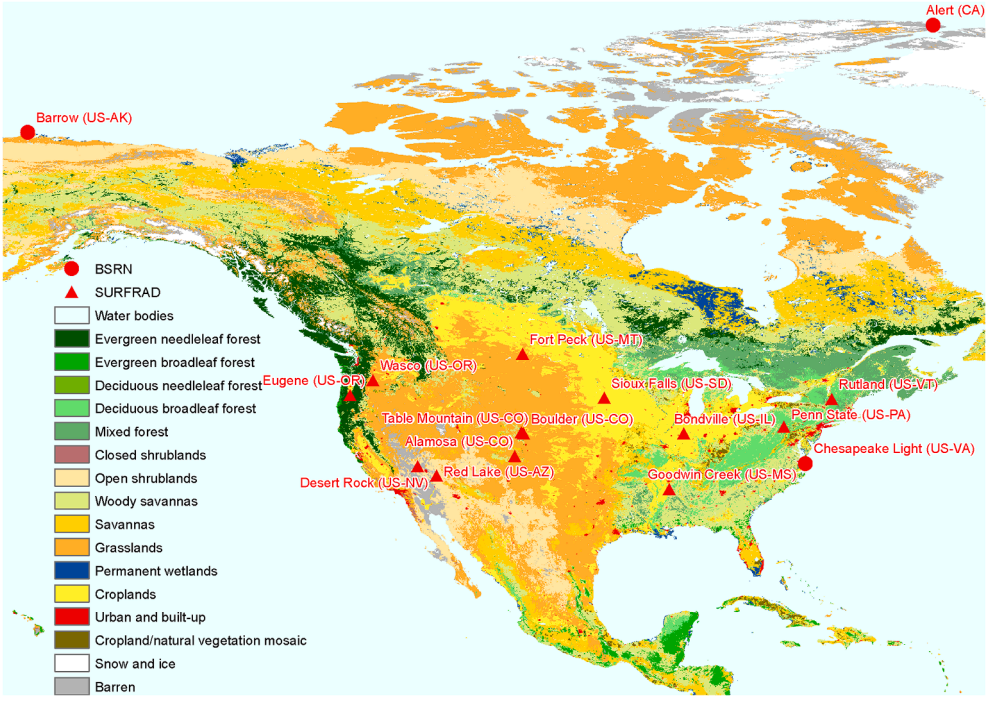

Fig. 3. North American sites covered in the analysis. (Source: IGBP base map from NASA (2012)).

3.4. Cross-Validation

The performance of the considered models was evaluated by a 10-fold cross-validation procedure performed separately for each site. The procedure involved ten runs of consecutive calibration (aka training) and validation. Here, only the snow-free data points were considered with all the datasets other than those of the land cover category "Snow and ice". In the case of the three "Snow and ice" datasets, only the records with a complete snow coverage were included. For each of the 26 sites, the resulting dataset was divided into ten subsets. Each subset was used once as the validation dataset while the remaining nine sets were allocated to calibration. In this way, the performance of the models was tested based on datasets that were independent of their calibration. This ensures that the performance indicators were not boosted by overfitting, for example. Random sampling was used for distributing the time steps into the ten datasets. Consequently, the chosen cross-validation procedure cannot be used to identify models whose parametrisations are robust against seasonal effects. The calibration step was performed through the method of least squares using MATLAB's Trust-Region-Reflective algorithm.

We limited the values of the fitting parameters to physically meaningful levels. The results of this constrained method can be expected to be more generalisable due to the maintained link to physics. The reference instances, ρ_n and ρ_{60°, were estimated by averaging measured ρ in subsets with $\theta_z \in [0^\circ, 5^\circ]$ and $\theta_z \in [55^\circ, 65^\circ]$, respectively. When the size of the subsets was insufficient, ρ_n and ρ_{60° were obtained through regression at the calibration step but keeping them subject to the following constraint: $\{\rho_n, \rho_{60^\circ}\} \in [0, 1]$. When evaluating the d-dependent models (Md_10 and M2_11–20,24–26), the following more restrictive constraint was used for $\rho_n : \rho_n \in [0, \rho_d]$. For these models, ρ_d (and ρ_b with Md_10) was estimated based on the d measurements as per the method proposed by Ineichen et al. (1990). In the rare case of an insufficient number of records available for this method, it was assumed that $\{\rho_d, \rho_b\} \in [0, 1]$. The range of $\rho_{b,60^\circ}$, in turn, was always limited to $[0, 1]$. The other parameters were constrained as specified in the original description of the models.

The performance indicators used for identifying the best-performing models for a site were the worst values out of the ten obtained for each of the following six statistics:

1. root-mean-square error (RMSE),
2. centred root-mean-square error (cRMSE),
3. mean absolute error (MAE),
4. mean bias error (MBE),
5. Pearson correlation coefficient (r), and
6. Kolmogorov–Smirnov statistic (D).

RMSE, MAE, MBE, and D are commonly used in the literature of irradiance modelling (e.g., Mieslinger et al., 2014; Cebecauer and Šúri, 2015; Gueymard and Ruiz-Arias, 2015). RMSE and MAE quantify the overall uncertainty of the estimation procedure, which is influenced by both random and systematic errors. Because of the quadratic formulation of RMSE, the weight of outliers is higher in RMSE than in MAE. Out of the constant estimators, M0_4 minimises RMSE and M0_21 MAE. RMSE and MAE are always positive unlike MBE. MBE is used to assess the model bias. r is the most commonly used correlation coefficient and can take on a value between −1 and 1. cRMSE only addresses the random error. It also reflects the fit in amplitude between the respective variability of the estimates and observations whereas r can be used as a pure measure of the similarity of the temporal patterns. Finally, D is used to quantify the similarity of the empirical distribution functions of the modelled estimates and observations. Both random and systematic errors affect D.

4. Model Performance

This section presents the results of the model evaluation. Prior to the actual evaluation, the observed correlation of ρ with θ_z and d is discussed in subsection 4.1. Subsection 4.2 delves into the comparison of aggregated errors and subsection 4.3 into the deviations of empirical distribution functions. Finally, the assessed overall performance is discussed for each site and model category in subsection 4.4.

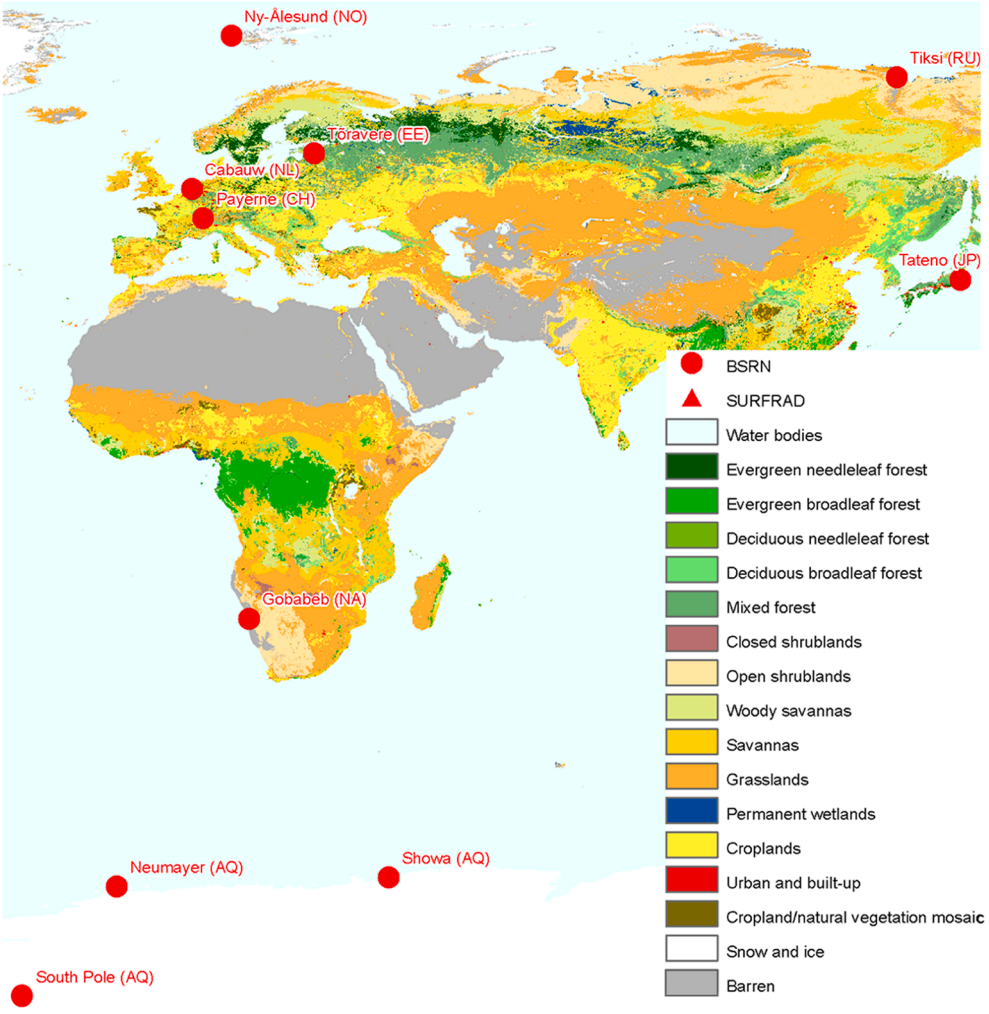

Fig. 4. Sites outside North America covered in the analysis. (Source: IGBP base map from NASA (2012)).

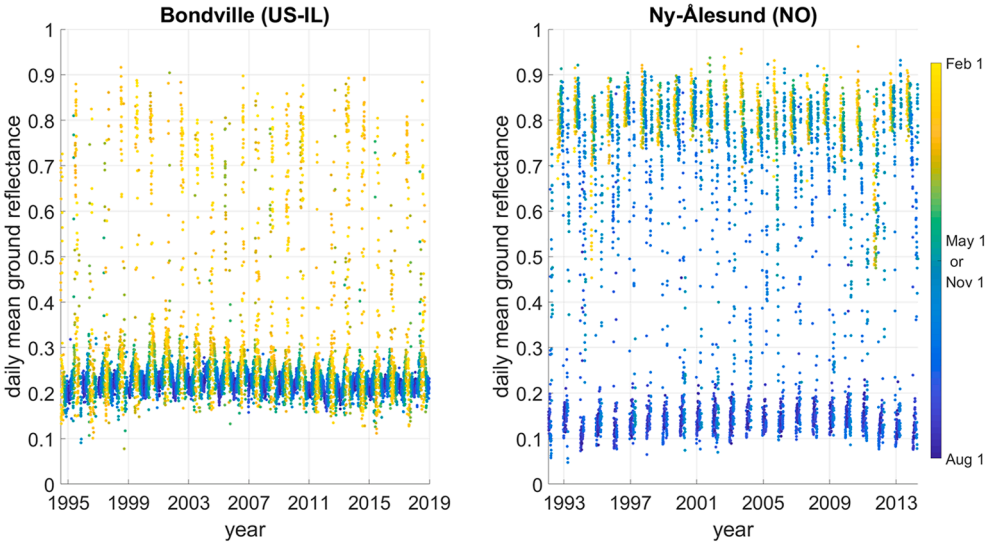

Fig. 5. Daily means of measured ground reflectance at two seasonally snowy sites. (The four-yearly ticks of the figure's time axes are placed at mid-year, the beginning of July).

Table 4
Description of the data by site (see Table 3 for more detailed site information).

Site	Reference	Considered timeframe	Snow-free daily mean ρ threshold
Goodwin Creek	Augustine et al. (2000, 2005)	1995–2019	0.25
Rutland		2014–2015	0.25
Boulder		2014	1*
Fort Peck		1995–2019	0.25
Table Mountain		1995–2019	0.20
Tõravere	Kallis (2019)	1999–2015	0.25
Alamosa	Augustine et al. (2000, 2005)	2014–2016	0.20
Wasco		2016–2017	0.25
Bondville		1995–2019	0.30
Penn State		1998–2019	0.25
Sioux Falls		2003–2019	0.25
Tateno	Ijima et al. (2019)	1996–2015	0.25
Cabauw	Knap (2019)	2005–2015	0.30
Payerne	Vuilleumier et al. (2014, 2019)	1992–2011	0.25
Tiksi	Kustov (2019)	2006–2014	0.20
Barrow	Dutton (2019)	1992–1010	0.25
Ny-Ålesund	Maturilli et al. (2015, 2019)	1992–2014	0.20
Gobabeb	Vogt (2019)	2012–2015	1*
Red Lake		2017	1*
Desert Rock	Augustine et al. (2000, 2005)	1998–2019	0.30
Alert	Cox and Halliwell (2019)	2004–2014	0.20
Neumayer	König-Langlo (2019)	1992–2015	0**
Syowa	Fukuda et al. (2019)	1994–2014	0**
South Pole	Long and Michalsky (2019)	1992–2015	0**
Eugene	Augustine et al. (2000, 2005)	2016–2017	0.15
Chesapeake Light	Denn et al. (2019)	2000–2015	1*

* all data points snow-free
** all data points with a snow coverage

4.1. Correlation Analysis

In order to pre-evaluate the level of predictability for each site, r was calculated between the measured monthly mean ρ and the monthly means of d and daily minimum θ_z. The coefficients are plotted in Fig. 9.

The preceding literature suggests that the correlation between ρ and daily minimum θ_z should be positive, increase with the latitude of the site, and decrease with an increasing Lambertianity of the surface. Therefore, a low correlation of ρ with daily minimum θ_z may indicate an approximately Lambertian surface, in which case standard deviation (σ) should be low, too. This would not be a problem for estimation accuracy. However, low or negative correlation can also suggest seasonally variable d or land cover that oppose the variability driven by solar geometry. However, apart from M0_3 and M0_22, the reviewed models do not capture seasonal changes in land cover even if these changes might dominate as the factor of ρ. Further research will address this topic in more detail.

The expected effect of d is dependent on θ_z. Indeed, at high solar elevations, an increasing d decreases the average grazing angle of ground-incident light thereby increasing ρ. At low solar elevations, the effect of d would be reverse. Since d does not vary deterministically — unlike θ_z — the annual mean d does not have an expected effect that can be generalised to all sites.

For a worst-case example, ρ estimation can be expected to be particularly difficult for a high-latitude site whose seasonal ρ variability is high and correlation between the monthly mean ρ and daily minimum θ_z is negative. As can be seen in Fig. 9, Tõravere, Wasco, Barrow, and Eugene are such sites. In Boulder and Payerne, the correlation is clearly negative but the variability is lower (Fig. 8). While also having negative correlations, Goodwin Creek, Table Mountain, and Red Lake have seasonally stable ρ and furthermore, they are located at comparatively low latitudes. Hence, the negative correlations are less alarming there.

4.2. Comparison of Aggregated Error

Fig. 10 shows the gain in performance when moving from literature-based, potentially severely biased estimates to data-based estimation (i.e., data-based constants and the estimates from univariate and bivariate models for each site).

On average, the reduction in RMSE is 16% with the data-based constants, 22% with the univariate models, and 31% with the bivariate models. In terms of MAE, the corresponding average differences are 22%, 29%, and 39%. As can be seen in the figure, the gain varies greatly between the sites. While the RMSE reduction is only 11% at maximum in Syowa, it exceeds 75% in Chesapeake Light. As opposed to Syowa, another snow-and-ice site, South Pole, also shows a reduction of almost 60%. In Syowa, the assumed library value, 0.75, does not result in underestimation as significantly as it does in South Pole and Neumayer. This can be seen more clearly in Fig. 11 that shows the MBE i.e., bias levels resulting from the usage of literature-based estimates. Much of the gain variability reported in Fig. 10 can be explained by the variation in this bias.

In addition to the MBE levels, Fig. 11 specifies both the most appropriate literature-based ρ levels and the practically unbiased, data-drawn values. The reflectance values of snow and barren ground are particularly difficult to estimate with a general constant. Among the desert sites, the data-based constants range from 0.21 to 0.34. In the case of bifacial PV systems for example, this variability would result in significant differences in bifacial gain (Fig. 1) and manifests the importance of site campaigns at desert sites in particular. Given the same ground surface and sky clarity, ρ levels tend to increase with site latitude. This is not the case with the considered desert sites (Table 3). Thus, the wide variety of the long-term average ρ is clearly due not to the solar geometry but to the differences in the soil type or the level of vegetation. The levels given in albedo libraries for wetland sites seem overestimated. Though, all the wetland sites considered in the study represent only subarctic or tundra climates. The results might be different if also lower-latitude wetland sites were included in the analysis. The large difference between the literature and measured values at Chesapeake Light may be explained by the highly non-Lambertian nature of water surfaces or by experimental considerations. The library value is valid for beam incidence of 53° (Gueymard, 2005) but at higher incidence angles, the ρ of water steeply increases. However, as pointed out by Fabbri et al. (2016), part of the field of view of the sensor in that station covers the monitoring station itself instead of the sea. The latter may be the main reason for M2_20 clearly outperforming the models specifically developed for semi-specular, water-like surfaces there.

Fig. 12 shows more explicitly the added value of time-variant modelling. It compares the RMSE reductions due to using the best-performing d-independent and d-dependent models instead of M0_4, which is the minimum-RMSE constant estimator. Most of the considered classes of land cover include sites where time-variant ρ modelling notably improves estimation accuracy. The only exceptions are the single urban site, Eugene, and snow-and-ice sites where the gain is only modest at two of the three sites. In terms of RMSE, the use of d as a predictor improves the estimation accuracy at all sites. It at least doubles the RMSE reduction at 18 out of the 26 sites. The RMSE improvement is insignificant only in Gobabeb.

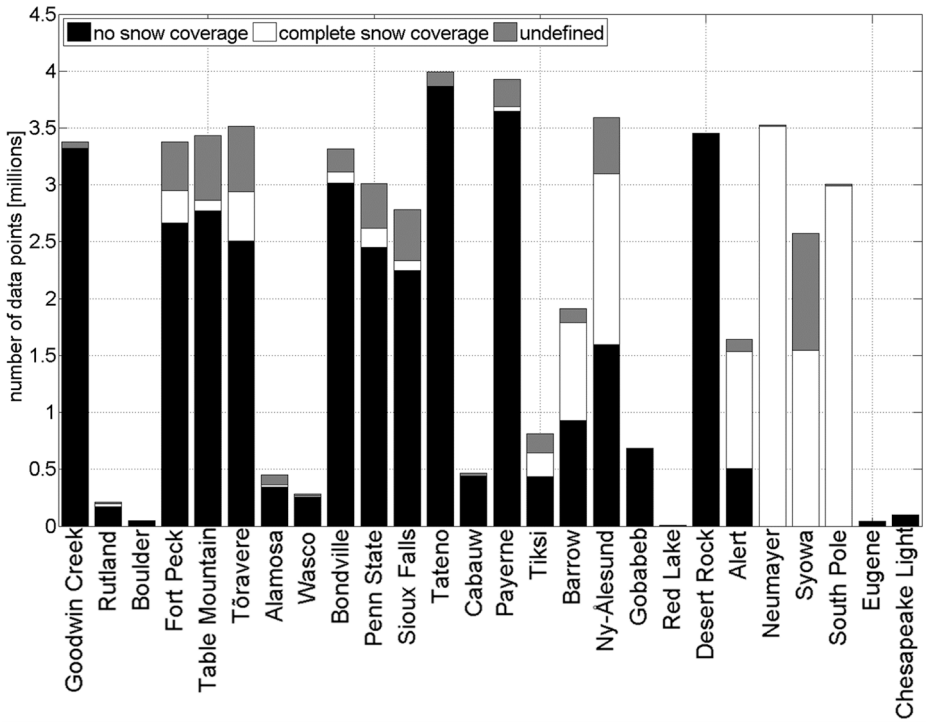

Fig. 6. Site-specific numbers of quality-assured daytime data points divided into records with no snow coverage, complete snow coverage, and undefined level of coverage.

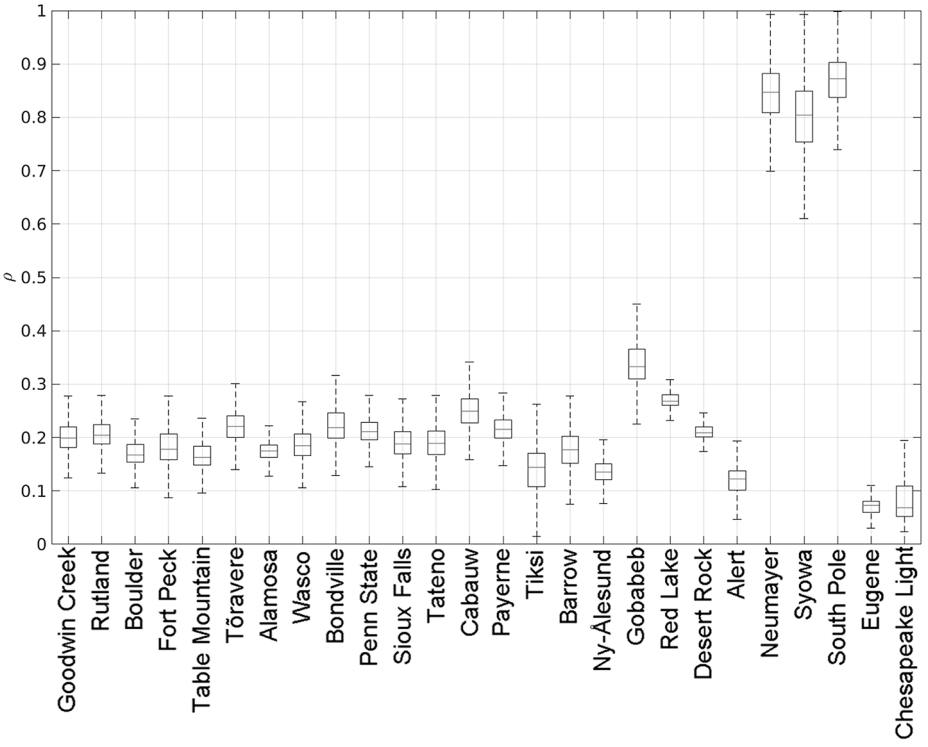

Fig. 7. Box plot of the ground reflectance (ρ) measurements considered in the analysis. (The top and bottom edges of the boxes indicate the 75th and 25th percentiles, respectively, and the red lines show the corresponding median values. The whiskers extend 2.7 standard deviations (σ) out of the mean values. The data points beyond these boundaries are not plotted.).

4.3. Deviations in Empirical Distribution Functions

As implied by Fig. 12, Gobabeb is an example of a site where d measurements may not be useful for improving the accuracy of ρ estimation. In fact, MAE decreases there when switching from the best-performing d-dependent model (M2_20) to the best d-independent one (Mz_9). Interestingly, while M2_20 provides the lowest RMSE, it is Mz_9 that outperforms the other models in terms of MAE. However, Mz_9 is

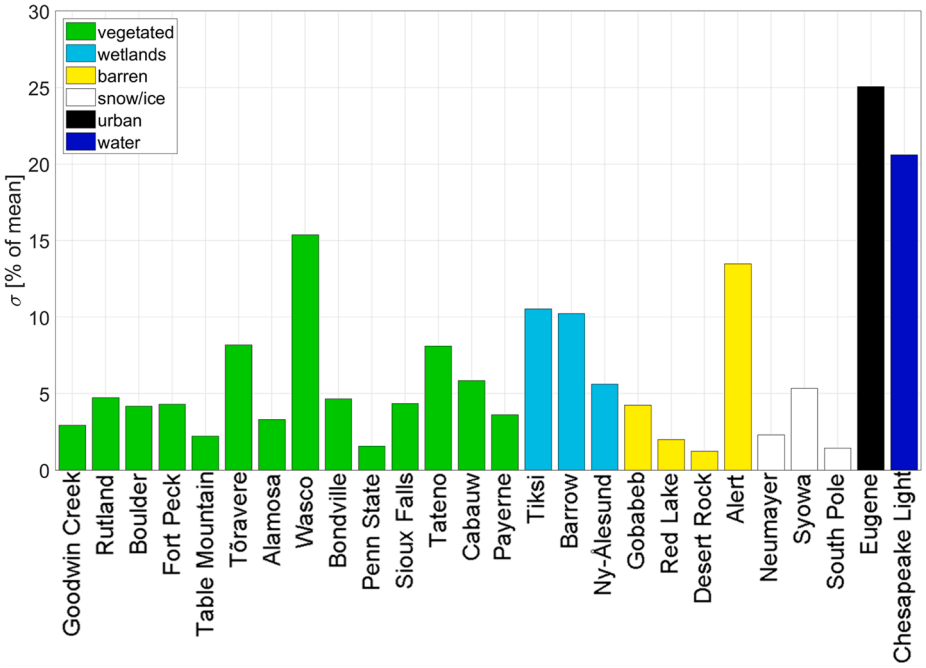

Fig. 8. Mean-normalised standard deviation (σ) of monthly mean surface reflectance (ρ).

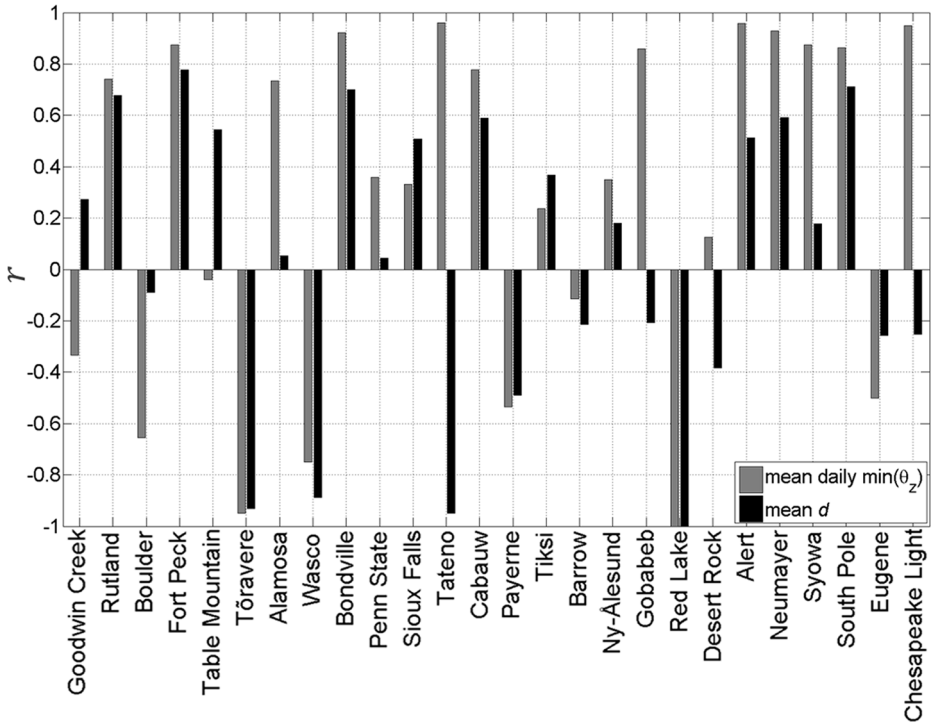

Fig. 9. Pearson correlation coefficients (r) of monthly mean ρ with the corresponding mean daily minimum solar zenith angle (θ_z) and mean diffuse irradiance fraction (d).

the only evaluated model with completely unconstrained coefficients, which may limit its generalisability. In terms of MAE, M2_20 is only at the fourth rank with a 6% higher MAE than that of Mz_9. The reason becomes clear with the help of Fig. 13 that shows the deviations of the cumulative distributions of the ρ estimates from the distribution of the observations. Mz_9 captures the overall distribution of the observations slightly better than M2_20. However, M2_20 provides the lowest RMSE due to its better accuracy at high ρ levels. The figure also indicates D for the selected models. In terms of D, the best-performing model is Mz_8. The distribution of its estimates follows that of the observations particularly well at low ρ levels. For this reason, with Mz_8, the obtained MAE is still lower than with M2_20. Hence, the rather small difference in MAE between Mz_8 and Mz_9 suggests that the best choice for modelling ρ at desert sites such as Gobabeb is Mz_8.

4.4. Overall Performance

Table 5 lists the models recommended for each site by category. The

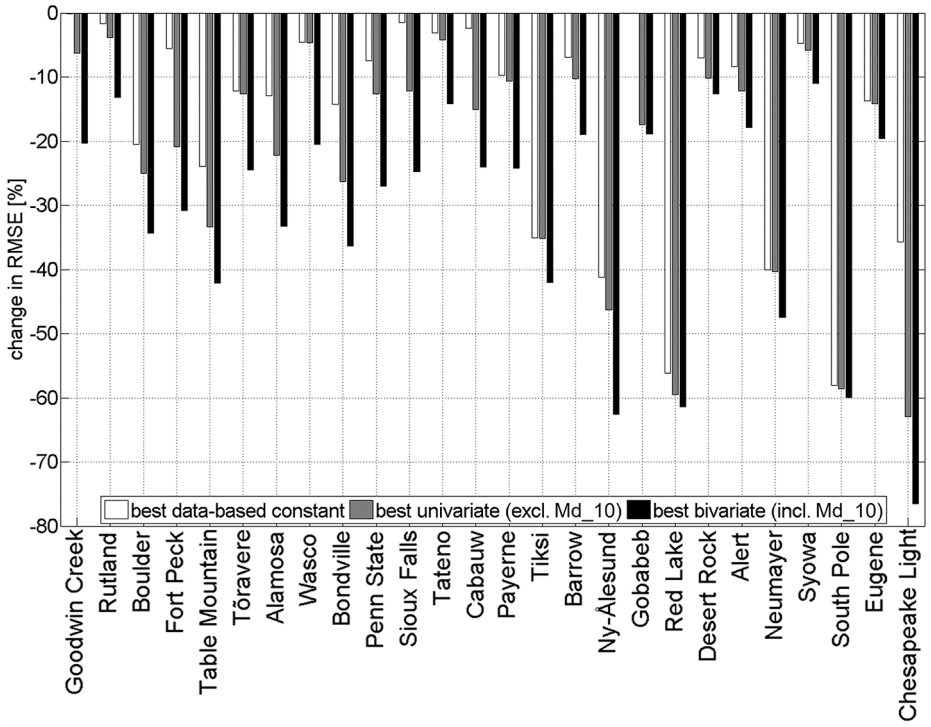

Fig. 10. Reduction in root-mean-square error (RMSE) due to estimating ground reflectance based on in-situ-measured data instead of literature-derived constants.

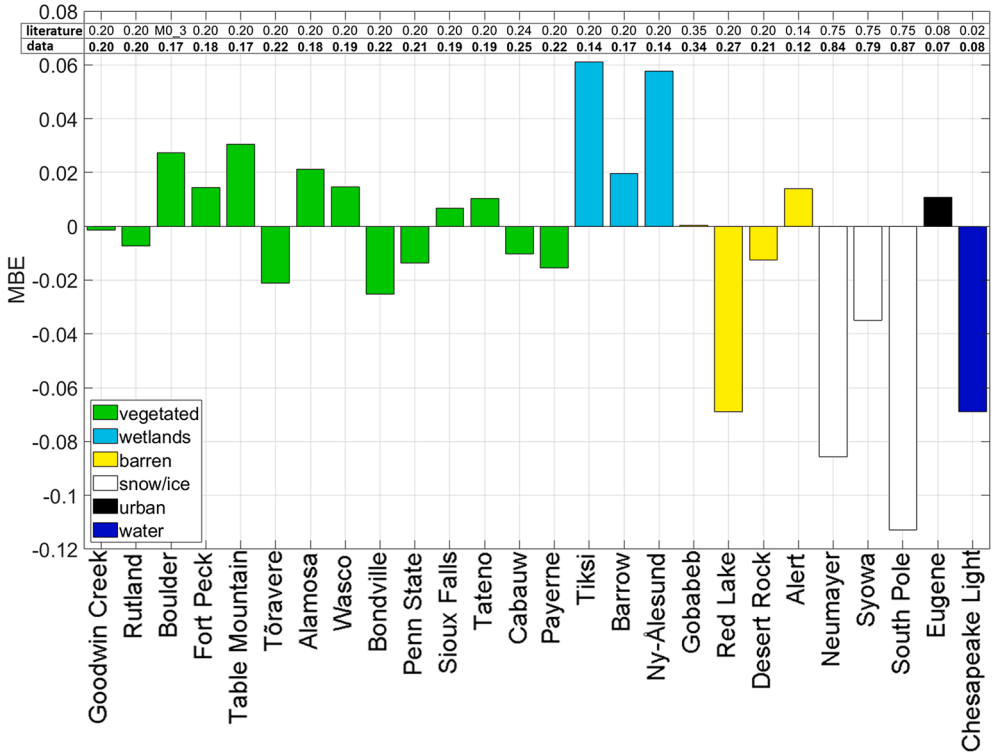

Fig. 11. Mean bias error (MBE) resulting from the usage of literature-derived ground reflectance i.e., models M0_1, M0_2, or M0_3 (varying monthly) with the respective reflectance levels indicated on the top panel.

models are identified based on a comparative analysis integrating all the considered performance indicators. As an entirely empirical formulation with three fitting coefficients, Mz_9 is listed as the best option only for the sites where it clearly improves estimation accuracy when switching from the other *d*-independent models.

The MAE levels specified in Table 5 give the average error in absolute terms. The table indicates that the permanent-wetland and snow-and-ice sites suffer from the largest errors in estimated surface reflectance (MAE of up to 0.069 in Syowa) while the urban and water sites feature the smallest errors with MAE of 0.013 and 0.011, respectively. Part of this

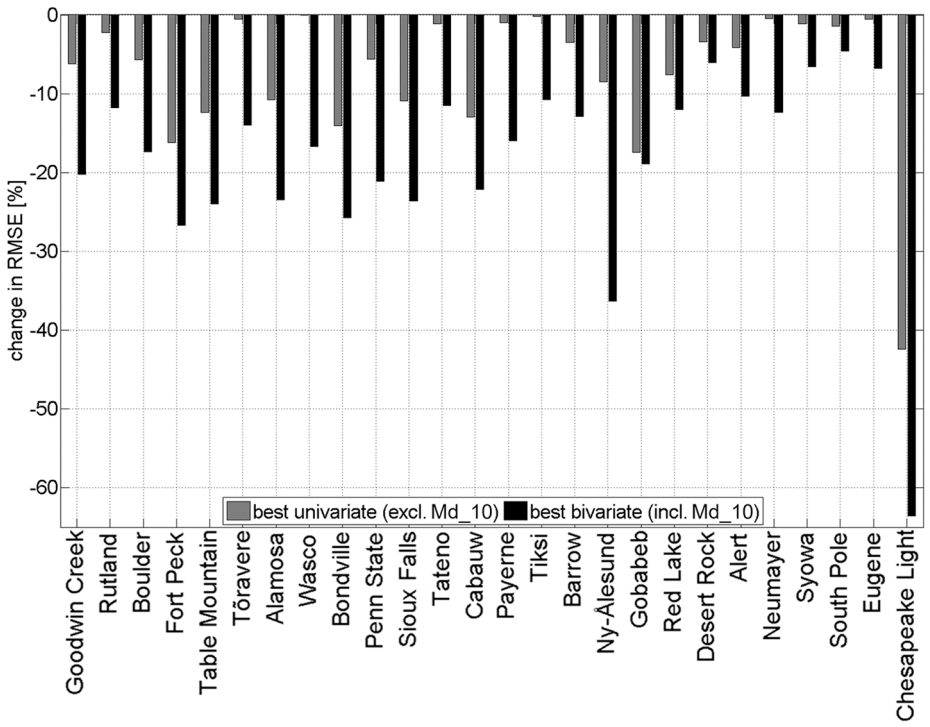

Fig. 12. Reduction in root-mean-square error (RMSE) due to estimating ground reflectance using time-variant models instead of the minimum-RMSE constant estimator (M0_4).

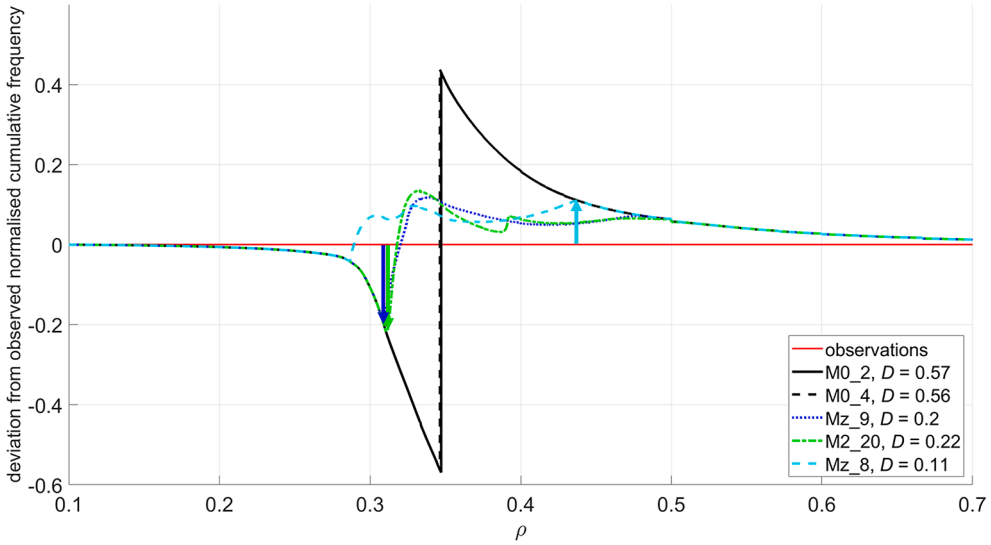

Fig. 13. Deviations of the ground reflectance (ρ) estimates' empirical cumulative distribution functions from the observations' distribution function in Gobabeb along with the corresponding Kolmogorov–Smirnov statistics (D) depicted by vertical arrows. (With the constant estimators, M0_2 and M0_4, D represents the proportion of ρ measurements below 0.35 and 0.34, respectively).

difference can be attributed to the variations in average reflectance between sites by up to a factor 4. Since most model formulations include the value of a site-specific instance of reflectance, the relative error should eliminate that effect. Fig. 14 then shows that, relatively speaking, it is actually the snow-and-ice sites whose ρ estimates are the most accurate. The tundra and permanent-wetland, urban, and water sites stand out with their low accuracy — in Tiksi, the relative, mean-normalised MAE (rMAE) of the best-performing model reaches 25%. The poor accuracy at the tundra and permanent-wetland sites may be due to the strong seasonal variability in ρ (see Fig. 8 in Section 3.3) and/or frequent occurrence of water body or snowmelt patches in the field of the sensor's view. The average minimum rMAE across all the sites is 11%. The average minimum mean-normalised RMSE is 17% with a maximum of 34% in Tiksi.

4.4.1. *Models Independent of Diffuse Fraction*

Mz_7 and Mz_23 are the best-performing d-independent models for most of the studied vegetated sites. Among the temperate and humid vegetated sites, Mz_7 appears to be favoured by higher latitudes (44° on average) than Mz_23 (39° on average).

The arid and semiarid sites tend to favour Mz_23 with the most desert-like site, Gobabeb, as a notable exception. In Gobabeb, Mz_8 is the

Table 5
Best-performing ground reflectance models for each site and model category. (The abbreviation of a model is in bold in the table when the model notably improves estimation accuracy compared to the best option of a simpler category.)

Site	constant			*d*-independent			*d*-dependent		
	Model	MAE	*D*	Model	MAE	*D*	Model	MAE	*D*
Goodwin Creek	M0_4	0.025	0.53	**Mz_23**	0.022	0.24	**M2_20**	0.018	0.18
Rutland	M0_4	0.026	0.53	**Mz_7**	0.024	0.33	**M2_20**	0.020	0.22
Boulder	M0_22	0.025	0.51	**Mz_8**	0.022	0.26	**M2_20**	0.019	0.23
Fort Peck	M0_4	0.031	0.51	**Mz_23**	0.024	0.16	**M2_14**	0.020	0.17
Table Mountain	M0_4	0.025	0.52	**Mz_23**	0.021	0.16	**M2_20**	0.018	0.18
Tōravere	M0_4	0.028	0.56	**Mz_8**	0.028	0.44	**M2_20**	0.023	0.23
Alamosa	M0_4	0.021	0.53	**Mz_23**	0.019	0.10	**M2_20**	0.017	0.33
Wasco	**M0_4**	0.031	0.57	Mz_7	0.031	0.52	**M2_20**	0.025	0.15
Bondville	M0_4	0.030	0.50	**Mz_23**	0.024	0.15	**M2_20**	0.019	0.18
Penn State	M0_4	0.023	0.51	**Mz_7**	0.022	0.24	**M2_20**	0.018	0.23
Sioux Falls	M0_4	0.028	0.50	**Mz_23**	0.024	0.17	**M2_20**	0.020	0.19
Tateno	M0_4	0.029	0.57	**Mz_7**	0.028	0.43	**M2_20**	0.024	0.23
Cabauw	M0_4	0.032	0.51	**Mz_7**	0.028	0.17	**M2_26**	0.024	0.24
Payerne	M0_21	0.023	0.56	**Mz_7**	0.023	0.42	**M2_20**	0.019	0.22
Tiksi	**M0_21**	0.039	0.64	Mz_9	0.039	0.64	**M2_20**	0.035	0.35
Barrow	M0_4	0.034	0.56	**Mz_23**	0.034	0.28	**M2_20**	0.029	0.29
Ny-Ålesund	M0_4	0.026	0.54	**Mz_23**	0.026	0.16	**M2_20**	0.018	0.29
Gobabeb	M0_4	0.046	0.56	**Mz_8**	0.031	0.11	M2_20	0.031	0.22
Red Lake	M0_21	0.021	0.55	**Mz_23**	0.021	0.19	**M2_20**	0.019	0.25
Desert Rock	M0_21	0.017	0.54	**Mz_23**	0.017	0.17	**M2_20**	0.016	0.16
Alert	M0_21	0.022	0.59	**Mz_8**	0.021	0.31	**M2_12**	0.018	0.22
Neumayer	**M0_21**	0.047	0.53	Mz_9	0.048	0.50	**M2_20**	0.040	0.41
Syowa	**M0_21**	0.071	0.55	Mz_9	0.073	0.63	**M2_20**	0.069	0.36
South Pole	M0_21	0.040	0.52	**Mz_7**	0.039	0.38	**M2_20**	0.038	0.43
Eugene	**M0_21**	0.013	0.59	Mz_9	0.014	0.68	**M2_20**	0.013	0.43
Chesapeake Light	M0_4	0.044	0.58	**Mz_7**	0.021	0.20	**M2_20**	0.011	0.29

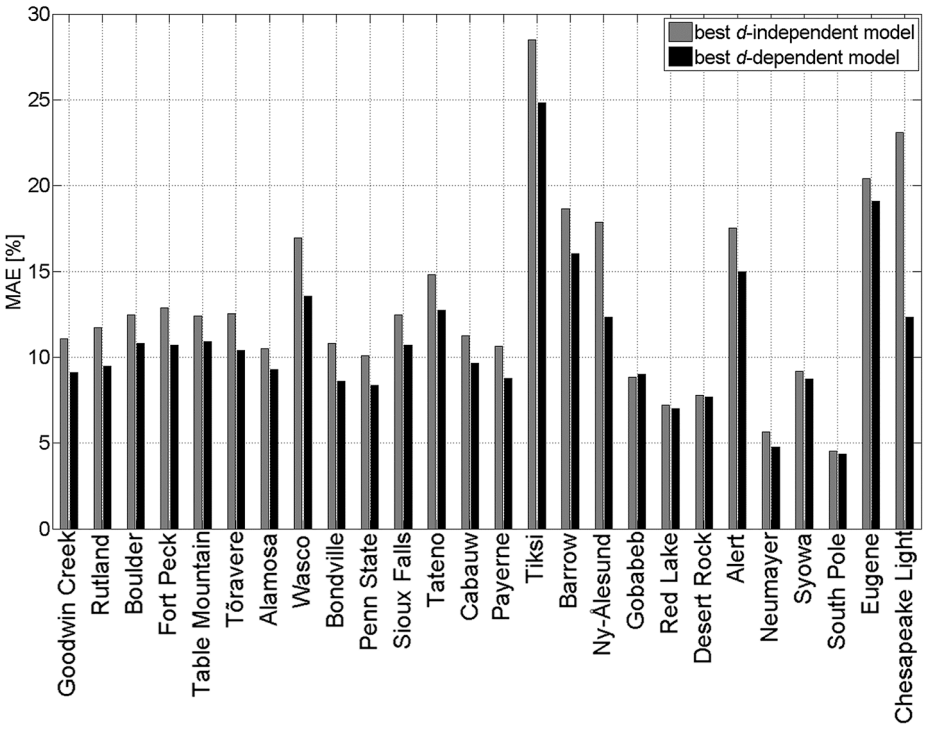

Fig. 14. Relative mean absolute error (MAE) of the ground reflectance estimates made by the best-performing option out of the models independent (grey) and dependent (black) on diffuse irradiance fraction (*d*).

best-performing option. Three models, Mz_7,8,23, perform very similarly in Red Lake and Desert Rock with slightly better indicators obtained for Mz_23. At the barren tundra site, Alert, Mz_8 is again the most suitable.

As already shown in Fig. 12, the urban and snow-and-ice sites are characterised by approximately Lambertian surface behaviour with the univariate models providing no or low gain in estimation accuracy. When considering the data points with snow coverage from all the considered sites together, no univariate model can outperform M0_4 ($\rho = 0.83$) in terms of RMSE or M0_21 ($\rho = 0.84$) in terms of MAE. Finally, the only considered water surface site, Chesapeake Light, also favours Mz_7 among the univariate models with the following parametrisation verifying the expected high level of specularity: $\rho_n = 0.014$ and $b = 1.731$.

4.4.2. Models Dependent on Diffuse Fraction

The consideration of d improves the accuracy of ρ estimation everywhere except for the desert sites. Thanks to its six coefficients, M2_20 provides the best accuracy at almost all the sites. For assessing its capability to capture the underlying physics, one can compare the coefficient levels obtained for different sites and attempt to relate them to physical site properties. Fig. 15 shows the comparison for the forward (f_{fs}) and backscatter (f_{bs}) coefficients. While both coefficients affect the estimated difference between ρ_n and ρ_d, f_{fs} also functions as the weight of the exponential term quantifying the non-Lambertian behaviour of the surface.

The dispersion of the coefficient values presented in the figure does not suggest any clear link to the estimated ρ_n level or to the surface and climate types of the site. The presented values were obtained with the exponent's three fitting coefficients (b_0, b_1, and b_2) varying almost freely, which naturally also affects the obtained levels for f_{fs} and f_{bs}. Gueymard (1987) specifies tentative values for the two assuming a fixed set of b_0,b_1, and b_2. It results in $\{f_{fs},f_{bs}\} = 1$ for green grass, to which the f_{fs} and f_{bs} of other surface types are relative. When fitting the model with this set to the data used in the present study, most of the resulting f_{bs} values saturate at zero i.e., its lower bound. f_{fs}, in turn, mainly takes on values between 0.8 and 1.2. Hence, while f_{fs} mostly behaves as expected, the obtained f_{bs} values do not agree with the findings of Gueymard (1987). Consequently, the results of this study do not support a general M2_20 parametrisation that would be suitable for an entire surface type category.

As M2_20 appears to suffer from overfitting, it is interesting to look into the performance of other models that also perform well but have fewer coefficients. The novel formulation M2_26 is one of the best-performing models at 19 sites. It performs better or comparably with M2_20 at 15 out of the 17 vegetation-covered sites. M2_14 and M2_18 also perform well at most vegetated sites.

As shown in Fig. 12, the three considered desert sites do not behave uniformly when it comes to the gain in ρ estimation accuracy due to using d as a predictor. In Gobabeb, the gain does not exist. In Red Lake, only M2_20 outperforms d-independent models in terms of MAE — that, too, only by 0.002 and with a higher D. In Desert Rock, also M2_26 and M2_18 slightly improve accuracy compared to Mz_23, the top-performing univariate model. However, the Desert Rock site is actually partially vegetated at least for part of the year (National Oceanic and Atmospheric Administration, 2019). These findings suggest a hypothesis that the completely barren desert sites such as Gobabeb and Red Lake favour d-independent models. This could be explained by the fact that desert sites mostly experience clear skies, which turns d into a weak indicator of solar elevation rather than one of atmospheric turbidity. Another rationale could be the spectral albedo of barren ground, which tends to be flatter than that of, e.g., green grass (Baldridge et al., 2009). As a result, the broadband reflectance of barren ground can be expected to be less sensitive to d. The hypothesis would have to be tested based on measurements from other barren desert sites.

As discussed above, the snow-and-ice sites show Lambertian behaviour. However, the estimation of ρ for snow seems to slightly improve when considering d. M2_20 performs consistently well at the three snow-and-ice sites as well as with the dataset that combines the records with complete snow coverage from all sites. The model specifically developed for snow, M2_18, clearly outperforms other models only in Neumayer. It is noteworthy that with snow, M2_18 provides a very similar accuracy to that of the other bivariate models with three coefficients (M2_17,19,26). When considering all the 13 million data points identified to certainly feature snow coverage, M2_20 provides modest reductions of 3% and 2% in RMSE and MAE, respectively, when compared to the M0_4-estimated $\rho = 0.83$. The single urban site, Eugene, is the only site where Md_10 performs comparatively well. However, Md_10 with $\rho_b = 0.066$ and $\rho_d = 0.076$ only reaches the same MAE as the M0_21-estimated $\rho = 0.073$. As shown in Figs. 8 and 9, the site's relatively high seasonal variability is negatively correlated with θ_z, which suggests physical changes in the land cover during the measurement period. In Chesapeake Light, M2_20 provides the best performance clearly surpassing all the other models, including the ones developed for semi-specular surfaces.

In conclusion, we can recommend M2_20 for the single water site and — to obtain more generalisable bivariate parametrisations — M2_26, M2_18, or M2_14 for vegetated sites and M2_18 for snow-and-ice sites.

5. Conclusion

The emergence of bifacial photovoltaics has made the accuracy of surface reflectance (ρ) estimation more relevant than ever before. Yet, only two previous papers have compared the performance of different ρ models, and they have done so on a limited dataset. In the present study, we addressed this gap by evaluating the performance of 20 models proposed in the preceding literature and of six novel ρ formulations. The evaluation was based on measurements from 26 different sites representing a broad range of climate and ground surface types.

On a global average, the best-performing ρ models result in a relative mean absolute error (rMAE) of 11% and a relative root-mean-square error of 17%. The variability between sites is large, however. Most of the vegetated sites feature rMAE of approximately 10%. Here, the tundra sites are a notable exception where the rMAE levels of the best-performing models reach values of up to 25%. Snow-covered and

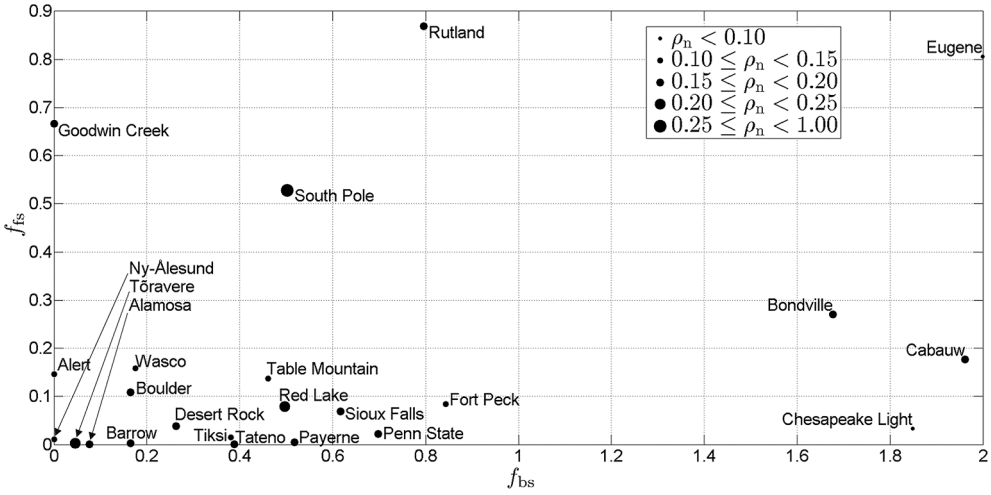

Fig. 15. Model M2_20 (see Table 1) coefficients quantifying forward (f_{fs}) and backscatter (f_{bs}) obtained for the studied sites.

barren sites are the most predictable. rMAE can go down to 8% at desert sites and 5% at the sites with permanent snow coverage. However, the *absolute* error at the snowy sites is high because of their high ρ. Moreover, the generalisability of constant ρ estimates between different sites is particularly difficult with snow-covered and barren sites due to the wide spectrum of existing surface colours within these surface type categories.

Contrary to previous research but in line with intuition, we found that data-based estimation notably improves the accuracy of ρ estimates at all the sites. On average, this gain in terms of MAE is 22%, 29%, and 39% with constant, univariate, and bivariate models, respectively. The data-based estimation is particularly important for surface types whose colour varies a lot between sites, in particular snow-covered and barren sites. The primary time-variant factor of ρ is solar elevation. Time-variant estimation significantly improves accuracy at all the considered sites except for a single urban site and the snow-covered sites. The less Lambertian the surface type can be expected to be, the more advisable time-variant estimation is. Prime examples are the surfaces covered either partially or entirely by water. The consideration of the impact of diffuse irradiance fraction clearly adds to this gain everywhere except for the desert sites.

Thanks to its comparatively large number of coefficients, the ρ model of Gueymard (1987) (M2_20) provides the best accuracy at almost all the sites. However, the novel bivariate model, M2_26, reaches a performance comparable to that of M2_20 at 19 out of the 26 considered sites with only three parameters. Another novel adaptation, Mz_23, is the best-performing univariate model at most arid and semi-arid sites as well as humid vegetated sites located at lower latitudes.

This paper, therefore, provides robust insights into the intrinsic quality of different surface reflectance models. Since practitioners rarely have the luxury to use extensive measurement data, further research shall investigate how the duration and timing of measurement campaigns affect the effective performance of these models for the design of solar energy systems.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Augustine, J.A., DeLuisi, J.J., Long, C.N., 2000. SURFRAD–A National Surface Radiation Budget Network for Atmospheric Research. Bull. Am. Meteorol. Soc. 81, 2341–2358. https://doi.org/10.1175/1520-0477(2000)081<2341:SANSRB>2.3.CO;2.

Augustine, J.A., Hodges, G.B., Cornwall, C.R., Michalsky, J.J., Medina, C.I., 2005. An Update on SURFRAD—The GCOS Surface Radiation Budget Network for the Continental United States. Journal of Atmospheric and Oceanic Technology 22, 1460–1472. https://doi.org/10.1175/JTECH1806.1.

Baldridge, A.M., Hook, S.J., Grove, C.I., Rivera, G., 2009. The ASTER spectral library version 2.0. Remote Sens. Environ. 113, 711–715. https://doi.org/10.1016/j.rse.2008.11.007.

Blanc, P., Wald, L., 2012. The SG2 algorithm for a fast and accurate computation of the position of the Sun for multi-decadal time period. Sol. Energy 86, 3072–3083. https://doi.org/10.1016/j.solener.2012.07.018.

Cebecauer, T., Šúri, M., 2015. Typical Meteorological Year Data: SolarGIS Approach. Energy Procedia 69, 1958–1969. https://doi.org/10.1016/j.egypro.2015.03.195.

Chiodetti, M., Lindsay, A., Dupeyrat, P., Binesti, D., Lutun, E., Radouane, K., Mousel, S., 2016. PV Bifacial Yield Simulation with a Variable Albedo Model, in. In: Proceedings of 32nd European Photovoltaic Solar Energy Conference and Exhibition, Munich, Germany, pp. 1449–1455. https://doi.org/10.4229/EUPVSEC20162016-5DP.1.4.

Cox, C.J., Halliwell, D., 2019. Basic and other measurements of radiation at station Alert (ALE), reference list of 115 datasets (2004-08 to 2014-02). doi:10.1594/PANGAEA.903206.

Denn, F.M., Schuster, G.L., Smith Jr., W.L., 2019. Basic measurements of radiation at station Chesapeake Light (CLH), reference list of 183 datasets (2000–05 to 2015–08). doi:10.1594/PANGAEA.903254.

Dickinson, R.E., 1983. Land Surface Processes and Climate—Surface Albedos and Energy Balance, in: Saltzman, B. (Ed.), Advances in Geophysics. Elsevier. volume 25 of Theory of Climate, pp. 305–353. doi:10.1016/S0065-2687(08)60176-4.

Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., Denn, F.M., Duprat, T., Fukuda, M., Grobe, H., Haeffelin, M., Hodges, G., Hyett, N., Ijima, O., Kallis, A., Knap, W., Kustov, V., Long, C.N., Longenecker, D., Lupi, A., Maturilli, M., Mimouni, M., Ntsangwane, L., Ogihara, H., Olano, X., Olefs, M., Omori, M., Passamani, L., Pereira, E.B., Schmithüsen, H., Schumacher, S., Sieger, R., Tamlyn, J., Vogt, R., Vuilleumier, L., Xia, X., Ohmura, A., König-Langlo, G., 2018. Baseline Surface Radiation Network (BSRN): Structure and data description (1992–2017). Earth System Science Data 10, 1491–1501. https://doi.org/10.5194/essd-10-1491-2018.

Dutton, E.G., 2019. Basic and other measurements of radiation at station Barrow (BAR), reference list of 218 datasets (1992–01 to 2010–02). doi:10.1594/PANGAEA.903249.

Espinar, B., Wald, L., Blanc, P., Hoyer-Klick, C., Schroedter-Homscheidt, M., Wanderer, T., 2011. Report on the Harmonization and Qualification of Meteorological Data. Technical Report D3.2. Project ENDORSE, Energy Downstream Service.

Fabbri, B., Schuster, G., Denn, F.M., Arduini, R., Madigan, J., 2016. 15-Year Climatology of BSRN Measurements at Chesapeake Light Station (CLH).

Fukuda, M., Doi, M., Kawashima, K., Yamanouchi, T., 2019. Basic and other measurements of radiation at station Syowa (SYO), reference list of 241 datasets (1994–01 to 2014–01). doi:10.1594/PANGAEA.903345.

Gardner, C.L., Nadeau, C.A., 1988. Estimating south slope irradiance in the Arctic—A comparison of experimental and modeled values. Sol. Energy 41, 227–240. https://doi.org/10.1016/0038-092X(88)90140-5.

Gilman, P., Dobos, A., DiOrio, N., Freeman, J., Janzou, S., Ryberg, D., 2018. SAM Photovoltaic Model Technical Reference Update. Technical Report NREL/TP-6A20-67399. National Renewable Energy Laboratory. Denver, CO, USA.

Gueymard, C.A., 1987. An anisotropic solar irradiance model for tilted surfaces and its comparison with selected engineering algorithms. Sol. Energy 38, 367–386. https://doi.org/10.1016/0038-092X(87)90009-0.

Gueymard, C.A., 1993. Mathematically integrable parameterization of clear-sky beam and global irradiances and its use in daily irradiation applications. Sol. Energy 50, 385–397. https://doi.org/10.1016/0038-092X(93)90059-W.

Gueymard, C.A., 1995. SMARTS2, A Simple Model of the Atmospheric Radiative Transfer of Sunshine: Algorithms and Performance Assessment. Professional Paper FSEC-PF-270-95. Florida Solar Energy Center. Cocoa, Florida, USA.

Gueymard, C.A., 2005. SMARTS Code, Version 2.9.5. USA.

Gueymard, C.A., Lara-Fanego, V., Sengupta, M., Xie, Y., 2019. Surface albedo and reflectance: Review of definitions, angular and spectral effects, and intercomparison of major data sources in support of advanced solar irradiance modeling over the Americas. Sol. Energy 182, 194–212. https://doi.org/10.1016/j.solener.2019.02.040.

Gueymard, C.A., Ruiz-Arias, J.A., 2015. Validation of direct normal irradiance predictions under arid conditions: A review of radiative models and their turbidity-dependent performance. Renew. Sustain. Energy Rev. 45, 379–396. https://doi.org/10.1016/j.rser.2015.01.065.

Hall, D., 1985. Remote Sensing of Ice and Snow. Remote Sensing Applications. Springer, Netherlands.

Hansen, C.W., Gooding, R.L., Guay, N.G., Kallickal, J.J., Ellibee, D.E., Riley, D.M., Asgharzadeh, A., Marion, B., Toor, F., Stein, J.S., 2017. A Detailed Model of Rear-Side Irradiance for Bifacial PV Modules, in: 44th IEEE Photovoltaic Specialists Conference (PVSC), Washington, D.C., USA. pp. 1543–1548. doi:10.1109/PVSC.2017.8366707.

Hou, Y.T., Moorthi, S., Campana, K., 2002. Parameterization of Solar Radiation Transfer in the NCEP Models. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service National Centers for Environmental Prediction.

Ijima, O., Hirose, Y., Naganuma, H., Ohkawara, N., 2019. Basic and other measurements of radiation at station Tateno (TAT), reference list of 233 datasets (1996–02 to 2015–06). doi:10.1594/PANGAEA.903355.

Ineichen, P., Guisan, O., Perez, R., 1990. Ground-reflected radiation and albedo. Sol. Energy 44, 207–214. https://doi.org/10.1016/0038-092X(90)90149-7.

Kallis, A., 2019. Basic and other measurements of radiation at station Toravere (2019–03).

Knap, W., 2019. Basic and other measurements of radiation at station Cabauw (CAB), reference list of 127 datasets (2005-02 to 2015-08). doi:10.1594/PANGAEA.903251.

König-Langlo, G., 2019. Basic and other measurements of radiation at station Neumayer (GVN), reference list of 274 datasets (1992–04 to 2015–01). doi:10.1594/PANGAEA.903327.

König-Langlo, G., Driemel, A., Raffel, B., Sieger, R., 2015. BSRN snapshot 2015–09, links to zip archives. doi:10.1594/PANGAEA.852720.

Kustov, V., 2019. Basic and other measurements of radiation at station Tiksi (TIK), reference list of 53 datasets (2006–10 to 2014–11). doi:10.1594/PANGAEA.903358.

Larsen, J.C., Barkstrom, B., 1977. Effects of Realistic Angular Reflection Laws for the Earth's Surface upon Calculations of the Earth-Atmosphere Albedo, in: International Symposium on Radiation in the Atmosphere, Garmisch-Partenkirchen, Germany.

Liu, B.Y.H., Jordan, R.C., 1960. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Sol. Energy 4, 1–19. https://doi.org/10.1016/0038-092X(60)90062-1.

Liu, B.Y.H., Jordan, R.C., 1963. The long-term average performance of flat-plate solar-energy collectors: With design data for the U.S., its outlying possessions and Canada. Sol. Energy 7, 53–74. https://doi.org/10.1016/0038-092X(63)90006-9.

Long, C.E., Michalsky, J., 2019. Basic and other measurements of radiation at station South Pole (SPO), reference list of 281 datasets (1992–01 to 2015–07). doi:10.1594/PANGAEA.903343.

Maturilli, M., Herber, A., 2019. Basic and other measurements of radiation at station Ny-Alesund (NYA), reference list of 269 datasets (1992–08 to 2014–12). doi:10.1594/PANGAEA.903328.

Maturilli, M., Herber, A., König-Langlo, G., 2015. Surface radiation climatology for Ny-Ålesund, Svalbard (78.9 N), basic observations for trend detection. Theoret. Appl. Climatol. 120, 331–339. https://doi.org/10.1007/s00704-014-1173-4.
Mermoud, A., Wittmer, B., 2016. PVsyst's new framework to simulate bifacial systems.
Mieslinger, T., Ament, F., Chhatbar, K., Meyer, R., 2014. A New Method for Fusion of Measured and Model-derived Solar Radiation Time-series. Energy Procedia 48, 1617–1626. https://doi.org/10.1016/j.egypro.2014.02.182.
Moon, P., Spencer, D.E., 1942. Illumination from a Non-Uniform Sky. Transactions of the Illuminating Engineering Society 37, 707–725.
NASA, 2012. MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V051.
National Oceanic and Atmospheric Administration, 2019. SURFRAD (Surface Radiation Budget) Network. https://www.esrl.noaa.gov/gmd/grad/surfrad/.
Nkemdirim, L.C., 1972. A Note on the Albedo of Surfaces. J. Appl. Meteorol. 11, 867–874. https://doi.org/10.1175/1520-0450(1972)011<0867:ANOTAO>2.0.CO;2.
Psiloglou, B.E., Kambezidis, H.D., 2009. Estimation of the ground albedo for the Athens area, Greece. J. Atmos. Solar Terr. Phys. 71, 943–954. https://doi.org/10.1016/j.jastp.2009.03.017.
Ross, J., 1981. The Radiation Regime and Architecture of Plant Stands. Tasks for Vegetation Science. Springer, Netherlands.
Schlick, C., 1994. An Inexpensive BRDF Model for Physically-based Rendering. Computer Graphics Forum 13, 233–246. https://doi.org/10.1111/1467-8659.1330233.
Temps, R.C., Coulson, K.L., 1977. Solar radiation incident upon slopes of different orientations. Sol. Energy 19, 179–184. https://doi.org/10.1016/0038-092X(77)90056-1.
Vogt, R., 2019. Basic and other measurements of radiation at station Gobabeb (GOB), reference list of 40 datasets (2012–05 to 2015–08). doi:10.1594/PANGAEA.903313.
Vuilleumier, L., Hauser, M., Félix, C., Vignola, F., Blanc, P., Kazantzidis, A., Calpini, B., 2014. Accuracy of ground surface broadband shortwave radiation monitoring: Shortwave radiation monitoring accuracy. Journal of Geophysical Research: Atmospheres 119, 13838–13860. https://doi.org/10.1002/2014JD022335.
Vuilleumier, L., Heimo, A., 2019. Basic and other measurements of radiation at station Payerne (PAY), reference list of 224 datasets (1992–10 to 2011–05). doi:10.1594/PANGAEA.903332.
Wang, Z., Zeng, X., Barlage, M., 2007. Moderate Resolution Imaging Spectroradiometer bidirectional reflectance distribution function–based albedo parameterization for weather and climate models. Journal of Geophysical Research: Atmospheres 112. https://doi.org/10.1029/2005JD006736.
Yang, F., Mitchell, K., Hou, Y.T., Dai, Y., Zeng, X., Wang, Z., Liang, X.Z., 2008. Dependence of Land Surface Albedo on Solar Zenith Angle: Observations and Model Parameterization. Journal of Applied Meteorology and Climatology 47, 2963–2982. https://doi.org/10.1175/2008JAMC1843.1.
Yusufoglu, U.A., Pletzer, T.M., Koduvelikulathu, L.J., Comparotto, C., Kopecek, R., Kurz, H., 2015. Analysis of the Annual Performance of Bifacial Modules and Optimization Methods. IEEE Journal of Photovoltaics 5, 320–328. https://doi.org/10.1109/JPHOTOV.2014.2364406.
Ziar, H., Sönmez, F.F., Isabella, O., Zeman, M., 2019. A comprehensive albedo model for solar energy applications: Geometric spectral albedo. Appl. Energy 255, 113867. https://doi.org/10.1016/j.apenergy.2019.113867.