



# A new model to estimate daytime net surface radiation under all sky conditions

Inmaculada Foyo-Moreno<sup>a,b</sup>, Ismael L. Lozano<sup>c,d,\*</sup>, Inmaculada Alados<sup>a,e</sup>,  
Juan Luis Guerrero-Rascado<sup>a,b</sup>

<sup>a</sup> Andalusian Institute for Earth System Research, Granada 18006, Spain

<sup>b</sup> Department of Applied Physics, University of Granada, Granada 18071, Spain

<sup>c</sup> Department of Forest Sciences, University of Helsinki, Helsinki 00014, Finland

<sup>d</sup> Institute for Atmospheric and Earth System Research, University of Helsinki, Helsinki 00014, Finland

<sup>e</sup> Applied Physics II Department, University of Málaga, Málaga 29071, Spain

## ARTICLE INFO

### Keywords:

Net radiation  
Shortwave radiation  
Longwave radiation

## ABSTRACT

Net surface radiation is a crucial parameter across various fields, as it represents the available energy for the energy exchange between the surface and the atmosphere. This work presents a new model for estimating instantaneous daytime net surface radiation ( $R_n$ ) under all sky conditions, using solar position via  $\cos \Theta_z$  and the clearness index ( $k_t$ ) as predictors. Global solar radiation ( $G_T$ ) is the primary factor influencing  $R_n$  and is extensively measured at numerous radiometric stations. Consequently, this model takes advantage of using a single input ( $G_T$ ). The model was validated against other empirical models at various sites with diverse climatological characteristics. Two types of models were evaluated, one including reflected global solar irradiance ( $G_{\uparrow}$ ) as an additional input variable alongside  $G_T$ . The best results were obtained when incorporating  $G_{\uparrow}$ . However, this poses a challenge as  $G_{\uparrow}$  is not measured at most radiometric stations. Nevertheless, in both types, the simplest model consistently outperformed the others, revealing no significant improvements with the addition of extra variables. Overall, the proposed model demonstrated good fit with the experimental data, although with some overestimation. The coefficient of determination ( $R^2$ ) is over 0.94, except at sites with extreme surface albedo conditions ( $\alpha > 0.55$ ). Mean bias error values ranged from  $4 \text{ Wm}^{-2}$  to  $44 \text{ Wm}^{-2}$ , while root mean square error values varied from  $25 \text{ Wm}^{-2}$  to  $62 \text{ Wm}^{-2}$ . Additional assessments across different seasons and sky conditions revealed improved performance during colder seasons and under cloudy conditions. Finally, the statistical analysis of the proposed model falls within the range of other more sophisticated models that involve additional input variables.

## 1. Introduction

Net radiation ( $R_n$ ) is the available energy provided by the Sun at the surface, since  $R_n$  results from the net balance of incoming and outgoing radiation at the surface, encompassing both shortwave and longwave ranges.  $R_n$  is a key parameter for studying surface processes, as this energy is used to heat the soil and the atmosphere, as well as to evaporate soil water. Consequently,  $R_n$  regulates several biological and environmental processes, including evapotranspiration (Lu et al., 2013, 2014; Wang and Liang, 2008), which is a critical component of agricultural, hydrological, and ecological research (Jiang et al., 2015), photosynthesis, turbulent and conductive heat fluxes. In the net balance,

the incoming radiation for shortwave range strongly depends on the site's latitude, solar position and sky conditions. On the other hand, the incoming longwave radiation directly depends on atmospheric properties such as air temperature (through Stefan Boltzmann's law) and water vapour pressure, and indirectly through changes in the atmospheric emissivity.

However, despite the importance of  $R_n$ , there are relatively few sites worldwide where it is measured. In fact, in contrast to the land surface,  $R_n$  is not routinely measured for the ocean surface. When  $R_n$  measurements are unavailable, various estimation methods are proposed as alternatives. These methods utilize either satellite data, a combination of satellite and surface data, or stand-alone surface data. When surface data

\* Corresponding author at: Department of Forest Sciences, University of Helsinki, Helsinki 00014, Finland, Institute for Atmospheric and Earth System Research, University of Helsinki.

<https://doi.org/10.1016/j.atmosres.2024.107886>

Received 8 August 2024; Received in revised form 30 November 2024; Accepted 19 December 2024

Available online 24 December 2024

0169-8095/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

is used, the scientific literature suggests several methods to estimate  $R_n$  based on meteorological measurements, which can be categorized into two methodological types. The simplest methods use only air temperature as an input variable to estimate downward longwave radiation of the atmosphere (Wright and Jensen, 1972; Brutsaert, 1975; Doorenbos and Pruitt, 1977; Weiss, 1982; Jensen et al., 1990), since air temperature influences the estimation of longwave radiation emitted by the atmosphere via the Stefan-Boltzmann law and subsequently impacts net radiation. However, this single parameter is insufficient to encompass the full variability of  $R_n$ . The other category includes methods that, in addition to air temperature, incorporate surface temperature to estimate longwave radiation emitted by the surface (Brutsaert, 1982; Ortega-Farías et al., 2000; Bisht et al., 2005; Saito et al., 2006; Cui et al., 2010; Hemmati et al., 2012). Irmak et al. (2003) also related  $R_n$  to a set of meteorological variables. However, Kjaersgaard et al. (2007a, 2007b) demonstrated that certain variables, such as daily maximum and minimum air temperatures, are inter-correlated, which may lead to multicollinearity and reduce the stability of prediction models (Jiang et al., 2015). Conversely, Wang and Liang (2009) were the first to consider the inclusion of surface elevation in  $R_n$  estimation. However, they showed that the original model, without this variable, provided accurate estimates.

Furthermore, many commonly used methods are based on the relationship between  $R_n$  and incoming shortwave radiation (often referred to as global solar radiation,  $G_1$ ) (Jiang et al., 2015; Alados et al., 2003), since  $G_1$  plays a critical role in the climate system and global energy balance (Wang et al., 2022). This approach is practical because  $G_1$  is commonly measured at many radiometric stations. There are two main types of empirical methods according to previous studies. The first type, which is widely used, estimates  $R_n$  from  $G_1$  and other meteorological variables through simple linear regressions. The second type frequently involves creating hybrid models that combine empirical and physical sub-models, allowing for the separate estimation of each individual component of  $R_n$ . Different authors have evaluated these models (e.g. Iziomon et al., 2000; Alados et al., 2003; Kjaersgaard et al., 2007a, 2007b; Sentelhas and Gillespie, 2008). Specifically, Jiang et al. (2015) evaluated seven models based on incident and/or net shortwave radiation and other meteorological variables suitable for global application. Based on these evaluations, they proposed a new empirical model that incorporates multiple variables, including the normalized difference vegetation index (NDVI).

Regarding  $G_1$  various forecasting methodologies are employed to estimate it, including empirical models, machine learning and physical and hybrid models. Lu et al. (2023) developed hybrid models and quantified the uncertainty in estimates caused by uncertainty in the measurement and atmospheric parameters for clear skies. They concluded that, under clear skies conditions, aerosol optical depth contributed the most to the accuracy of  $G_1$  estimates, with an average contribution of 58 %. Empirical models do not consider the effects of physical processes or topographic or climatic features (Lu et al., 2023), but these models are preferred by its simplicity and the advantage is that with only a few input parameters is possible the estimation of interest variable, the disadvantage is the lack of the universality. Wang et al. (2024) established and compared four hybrid models by coupling a physical model with machine learning model to estimate  $G_1$ . Uncertainties in the atmospheric parameters greatly limit the performance of the hybrid model.

Based on our previous models and works, the main goal of this study is to propose a new, simplified model that requires only global irradiance measurements, which are available at many radiometric stations worldwide, and the solar position. The fundamental idea underlying this model is to consider a single input variable, which is the product of the clearness index ( $k_t$ ) and the cosine of solar zenith angle ( $\cos \Theta_z$ ). The model was developed using data from one station in Payerne (PAY), Switzerland during 2020 and 2021 and validated using data from seven stations: Barrows (BAR) in Alaska (USA); Gobabeb (GOB) in the Arabian

desert; Izaña (IZA) in the Canary Islands, Spain; Budapest (BUD) in Hungary, Tateno (TAT) in Japan for the year 2022; and Toravere (TOR) in Estonia for the year 2019, each with a complete year of data; and data from Payerne during 2022. The structure of this article is organized as follows. Section 2 describes the experimental sites and instrumentation, Section 3 details the methodology, Section 4 presents the results and discussions, and Section 5 provides the conclusions.

## 2. Experimental sites and instrumentation

The measurements used in this study were collected at seven radiometric stations (Fig. 1) belonging to Baseline Surface Radiation Network (BSRN, <http://www.bsrn.awi.de>) (Driemel et al., 2018; Ohmura et al., 1998). The selection of sites has been delimited to these sites, considering that not all variables are measured in the majority of sites of BSRN and in order to validate the empirical models and calculate their own coefficients, it is necessary to provide two full years. Thus, the sites are shown in Table 1.

The aim of the BSRN is to provide high quality measurements of longwave and shortwave radiation fluxes, employing a high sampling rate of 1 Hz, and a short retrieval interval of 1 min (Ohmura et al., 1998). The BSRN states accuracy requirements of  $5 \text{ Wm}^{-2}$  for shortwave radiation and  $20 \text{ Wm}^{-2}$  for longwave radiation (Ohmura et al., 1998). The specifications for each station included in this study are provided in Table 1. Climate specifications are determined according to the Köppen classification (Köppen and Geiger, 1936). The study sites are distributed worldwide to cover a wide range of locations and climates: dry arid low latitude (Bwd; GOB), temperate Mediterranean climate (Csb; IZA), humid continental mild summer (Dfb; BUD, PAY and TOR) and polar tundra (ET; BAR) and temperate no dry season, hot summer (Cfa; TAT).

Thus, the differences among the sites become evident when considering various factors that influence net radiation. All sites, except TAT, are rural, each exhibiting different climatic conditions, with BAR and GOB representing extreme environments. PAY, BUD and TOR experience cold climates with warm summers, while IZA and TOR have temperate climates, with IZA characterised by warm summers and TAT experiences hot summers. Regarding surface type, three sites are predominantly grass-covered (BUD, TAT and TOR), BAR is characterised by tundra, GOB by desert, IZA by rocky terrain, and PAY by cultivated land. Latitude is another important factor, from  $23,5^\circ\text{N}$  at GOB to  $70^\circ\text{N}$  at BAR, and altitude varies significantly, from 8 m at BAR and 2373 m at IZA.

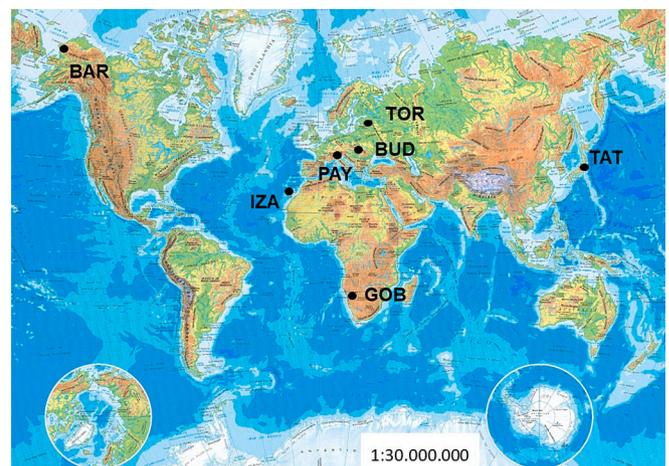


Fig. 1. Location of the Baseline Surface Radiation Network (BSRN) sites: Barrow (BAR), Budapest (BUD), Gobabeb (GOB), Izaña (IZA), Payerne (PAY), Tateno (TAT), Toravere (TOR). Scale: 1:30,000,000. Information obtained from Geographic Institute, Ministry of Public Works, Spain Government (<http://www.ign.es/web/catalogo-cartoteca/resources/html/031459>.)

**Table 1**  
Information related to the seven observation sites.

Station	Years and Datasets	Location	Surface type	Topography type	Rural/urban	Köppen-Geiger classification (*)	Instruments
Barrow (BAR) (71,32 N, 156,01 W, 8 m a.s.l.)	2021 / 2022	USA, Alaska	Tundra	Flat	Rural	ET (Polar, tundra, low evapotranspiration, no summer)	Pyranometers: Eppley (two PSP). Pyrgeometers: Eppley (two PIR).
Budapest-Lorinc (BUD) (47,43 N, 19,18E, 139,1 m a.s.l.)	2021 / 2022	Hungary, Budapest	Grass	Flat	Rural	Dfb (Cold, no dry season, warm summer)	Pyranometers: Kipp & Zonen (CMP11 and CMP6). Pyrgeometers: Kipp & Zonen (CGR4 and CGR3)
Gobabeb (GOB) (23,56S, 15,04E, 407 m a.s.l.)	2021 / 2022	Namibia, Namib Desert	Desert gravel	Flat	Rural	BWh (Arid, desert, hot)	Pyranometers: Kipp & Zonen (two CMP22). Pyrgeometers: Kipp & Zonen (two CGR4).
Izaña (IZA) (28,31 N,16,50 W, 2372,9 m a.s.l.)	2021 / 2022	Spain, Tenerife	Rock	Mountain top	Rural	Csb (Temperate, dry summer, warm summer)	Pyranometers: EKO (MS-802F and MR-60). Pyrgeometer: Kipp and Zonen (CGR4).
Payerne (PAY) (46,81 N, 6,94E, 491 m a.s.l.)	2020/ 2021 / 2022	Switzerland	Cultivated	Flat	Rural	Dfb (Cold, no dry season, warm summer)	Pyranometers: Kipp & Zonen (CMP22 and CMP21). Pyrgeometers: Kipp & Zonen (CG4) and Eppley (PIR).
Tateno (TAT) 36,06 N,140,16E, 25 m a.s.l.)	2021 / 2022	Japan	Grass	Flat	Urban	Cfa (Temperate, no dry season, hot summer)	Pyranometer, Kipp & Zonen, (two CMP21). Pyrgeometer, Kipp & Zonen, (two CGR4)
Torovere (TOR) (58,26 N, 26,46E, 70 m a.s.l.)	2018 / 2019	Estonia	Grass	Flat	Rural	Dfb (Cold, no dry season, warm summer)	Pyranometers: Kipp & Zonen (SMP11 and CMP21), Pyrgeometers: Kipp & Zonen (CGR4) and Eppley (PIR).

(\*) Köppen, W & Geiger, R., 1936. Das geographische System der Klimate, Belen.

Each station is equipped with two pyranometers to measure downwelling ( $G_d$ ) and upwelling (reflected,  $G_r$ ) shortwave radiation, as well as two pyrgeometers to measure downwelling ( $L_d$ ) and upwelling (emitted by the surface;  $L_r$ ) longwave radiation.

To maximize the representativeness and significance of this study, an in-depth data quality control was performed for each site, applying the three tests suggested by Lozano et al. (2023) for the photosynthetic active radiation range of solar radiation to the shortwave radiation range. The first test aims to exclude errors due to the cosine response of the instruments; therefore, only data with a solar zenith angle ( $\theta_z$ ) below  $80^\circ$  have been considered. The second test involves the clearness index ( $k_t$ ), defined as the ratio between the solar global radiation ( $G_t$ ) and the extraterrestrial global irradiance ( $G_{ext}$ ). This test limits  $k_t$  values between 0 and 1. The third test prevents anomalies, outliers, and extreme values by limiting the maximum and minimum value of the solar radiation. Thus, values exceeding those expected under clear or overcast skies are eliminated. Finally, to avoid anomalous data resulting from equipment malfunctioning (e.g. voltage issues) and outliers in longwave radiation data, a visual inspection was performed.

### 3. Methodology

In this work, instantaneous data recorded under all sky conditions for solar elevations greater than  $10^\circ$  were used to avoid cosine effects. Only daytime measurements, defined as the time period between sunrise and sunset, were considered.  $R_n$  is the sum of different net terms: the difference between downward ( $\downarrow$ ) and upward ( $\uparrow$ ) components in both the shortwave and longwave ranges ( $G$  and  $L$ , respectively). Therefore, the corresponding equation for  $R_n$  is:

$$R_n = G_n + L_n = (G_d - G_r) + (L_d - L_r) = G_d(1 - \alpha) + (L_d - L_r) \quad (1)$$

The variables included in eq. (1) are:

$G_n$ : Net shortwave radiation.

$L_n$ : Net longwave radiation.

$G_d$ : Downwelling shortwave radiation.

$G_r$ : Reflected shortwave radiation.

$L_d$ : Downwelling longwave radiation.

$L_r$ : Upwelling longwave radiation.

$\alpha$ : Surface albedo.

Table 2 shows information about the variables that will be used.

It is widely known that the primary driver of net daytime radiation is its shortwave component, a concept highlighted in following section 4.1. In fact, the correlation observed between  $R_n$  and  $G_d$  for the selected sites is very high, with determination coefficients ranging from 0,94 at TAT to 0,97 at PAY, with the exception of BAR. This reinforces the conclusion that, in general,  $L_n$  shows negligible variation compared to  $G_d$ . On the other hand, previous research (Foyo-Moreno et al. (1999, 2007, 2017) has consistently demonstrated a linear relationship between the maximum values of solar radiation, occurring under clear-sky conditions and  $\theta_z$  in different spectral ranges, including ultraviolet and visible wavelengths. To refine peak radiation estimates to account for varying sky conditions, the clearness index ( $k_t$ ) has been introduced to this relationship as an additional factor. This parameter is defined as the ratio between global solar radiation and extraterrestrial global solar radiation, both on a horizontal surface. The corresponding expression is:

$$k_t = \frac{G_t}{G_{ext}} \quad (2)$$

**Table 2**  
Information about variables used in the study.

Abbreviation	Variable	Unit	Source
$G_d$	Downwelling shortwave radiation	$Wm^{-2}$	In situ
$G_r$	Upwelling shortwave radiation	$Wm^{-2}$	In situ
$L_d$	Downwelling longwave radiation	$Wm^{-2}$	In situ
$L_r$	Upwelling longwave radiation	$Wm^{-2}$	In situ
$\alpha$	Albedo		Calculated
$E_0$	Inverse relative Earth-Sun distance		Calculated
$G_{ext}$	Extraterrestrial global irradiance	$Wm^{-2}$	Calculated
$k_t$	Clearness index		Calculated
$T_a$	Air temperature	$^\circ C$ or $K$	In situ
$\cos \theta_z$	Cosine solar zenith angle		Calculated

where  $G_{\text{ext}}$  is extraterrestrial global solar irradiance.  $G_{\text{ext}}$  is calculated from the expression:

$$G_{\text{ext}} = I_{\text{sc}} E_o \cos \theta_z \quad (3)$$

where  $I_{\text{sc}}$  is the solar constant ( $1361,1 \text{ Wm}^{-2}$ , Gueymard, 2018),  $E_o$  is the inverse relative Earth–Sun distance and  $\theta_z$  is the solar zenith angle. The expression of  $\cos \theta_z$  is:

$$\cos \theta_z = \sin \delta \sin \varphi + \cos \delta \cos \varphi \cos \omega \quad (4)$$

Thus, Fig. 2 shows the relationship between  $R_n$  and this combined metric (the product of  $\cos \theta_z$  and  $k_t$ ) at the PAY station over the course of two years (2020 and 2021). The following relationship has been found with a determination coefficient ( $R^2$ ) of 0,981 at PAY:

$$R_n = (-16,7 \pm 0,2) + (716 \pm 1)k_t \cos \theta_z + (241 \pm 1)(k_t \cos \theta_z)^2 \quad (5)$$

Considering the analysis presented in section 4.1, the choice of PAY to generate the model was motivated by its intermediate conditions compared to the other study sites.

Considering the dependence of  $R_n$  on  $G_1$  using an empirical linear regression model expressed as  $R_n = A G_1 + B$ , where  $A$  and  $B$  are regression constants, this relationship can be reformulated as:

$$R_n = X(1 - \alpha)G_1 + Y = X G_n + Y \quad (6)$$

where  $X$  and  $Y$  are the regression constants. This negates the assumption that surface albedo can be considered constant. By combining eq. (1) and eq. (6),  $L_n$  can be expressed as a linear function of  $G_n$  as follows:

$$L_n = (X - 1)G_n + Y \quad (7)$$

Defining  $d$   $L_n/dG_n$  as  $\lambda$  (the longwave exchange coefficient), eq. (7) can be expressed as follows:

$$\lambda = dL_n/dG_n = (X - 1) = (L_n - Y) / G_n \quad (8)$$

with  $G_1 = 0$ ,  $L_n = L_0$  such that  $Y = L_0$  in eq. (7). Adding  $G_n$  to both sides of eq. (7) results in:

$$R_n = (1 + \lambda)G_n + L_0 \quad (9)$$

$\lambda$  can be considered as an index of surface thermal response. Thus,  $\lambda = 0$  would imply that an increase in  $G_n$  is entirely allocated to evapotranspiration, resulting in no change in surface temperature and no alteration in  $L_n$  with an increase in  $G_n$ . The longwave exchange coefficient is observed to be negative for all sites, with values of  $-0,10$  at BAR and  $-0,21$  at TOR. For PAY and TAT, the value is  $-0,15$ , while IZA has a

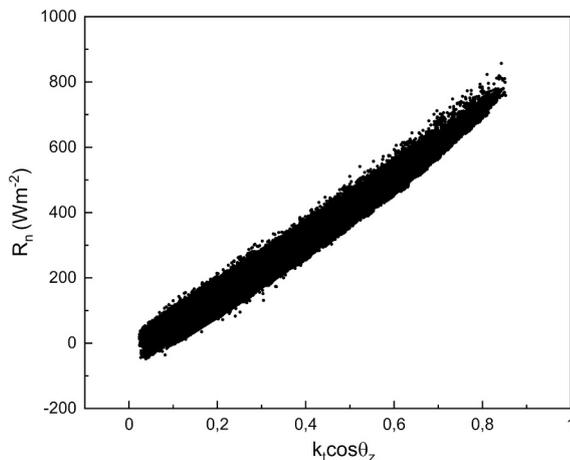


Fig. 2. Relationship between net radiation  $R_n$  and the product of  $\cos \theta_z$  per  $k_t$  in Payerne for 2020 and 2021.

value of  $-0,18$ , and GOB and BUD have values of  $-0,19$  and  $-0,20$  respectively. These results are consistent with those found by Iziomon et al. (2000), who reported an average value of  $-0,20$  for three sites at different altitudes in the southern Upper Rhine valley. The negative value of  $\lambda$  implies that  $G_n$  increases more rapidly than does the energy flux allocated to evapotranspiration at the surface. As surface temperature rises more quickly,  $L_n$  becomes more negative (Iziomon et al., 2000). The different values obtained at the seven sites show the different capacity to the excess of sensible heat flux that is made available at the surface in this way subsequently gets transformed into convection and change in storage (Iziomon et al., 2000). Thus, BAR presents a low surface thermal response compared to the other sites.

To separately compare the variables used in this work, including the product of  $k_t \cos \theta_z$  as a new proposed variable, an additional Spearman correlation analysis was conducted for each site. Table 3 presents the Spearman coefficient values. For four sites,  $k_t \cos \theta_z$  displays the highest coefficient, followed by  $G_1$  and  $\cos \theta_z$ , while for the remain sites, the order is  $G_1$ ,  $k_t \cos \theta_z$  and  $\cos \theta_z$ . Thus, the Spearman correlation analysis between  $R_n$  and the all variables used in this work reveals a high correlation overall, with three primary variables ranked as  $k_t \cos \theta_z$ ,  $G_1$  and  $\cos \theta_z$ , with values exceeding 0,90 for the first two. Also, Peng et al. (2021) demonstrated that  $R_n$  is primarily influenced by downward shortwave radiation ( $G_1$ ). They used data collected from 66 globally distributed moored buoy sites across five networks/projects.  $G_1$  was found to be the most important variable, followed by the clearness index.

## 4. Results and discussion

### 4.1. Global characterization

#### 4.1.1. General context

Table 4 provides a characterization of the study variables, including  $R_n$  and its components ( $G_1$ ,  $G_t$ ,  $L_1$  and  $L_t$ ),  $\alpha$ , and  $k_t$  for all stations. This basic descriptive statistic includes parameters such as the arithmetic mean (Ave), standard deviation (SD), first and third quartile (Q1 and Q3, respectively), median (Md), and 10th and 90th percentiles (P10 and P90, respectively).

Based on sky conditions evaluated using the  $k_t$  index, it is observed that TAT exhibits the lowest mean values, recording  $0,23 \pm 0,05$  followed by BAR and TOR with values  $0,48 \pm 0,18$  and  $0,49 \pm 0,24$ , respectively. The relatively high standard deviation value for these two last sites suggest high variability. The highest mean

value are registered at IZA ( $0,78 \pm 0,10$ ) and GOB ( $0,72 \pm 0,09$ ). Notably, these two sites present low variability concerning this index. This result indicates, on average, a greater presence of clear sky conditions at IZA and GOB, while TAT, BAR and TOR experience a greater presence of clouds and/or aerosols, TOR showing the greatest range of values between the extreme values compared to the other stations. TAT and IZA exhibit contrasting patterns with respect to sky conditions as indicated by the  $k_t$  index. Although both sites have a temperate climate, TAT does not experience a dry season, which favors the formation of the convective clouds. Conversely, the albedo behaves oppositely. The highest values are observed at TAT ( $0,50 \pm 0,20$ ), BAR ( $0,48 \pm 0,33$ ) and also GOB ( $0,37 \pm 0,05$ ), while the lowest is at IZA ( $0,16 \pm 0,02$ ). Additionally, BAR exhibits a higher range of albedo variation compared to the other sites, ranging between 0,16 (P10) and 0,89 (P90), followed by TAT with values of 0,20 and 0,75 for P10 and P90, respectively. It is also noteworthy that GOB has high albedo values, with P10 of 0,33 and P90 of 0,41, although the range of values is smaller than at BAR and TAT. The high albedo values for at BAR are attributed to the tundra surface type and the cold climate, which includes snow cover for half of the year, in contrast to IZA, which exhibits a lower range of albedos.

For all sites,  $L_t$  is always higher than  $L_1$ , resulting in a negative  $L_n$ . On the other hand, the net balance is positive during the day, indicating that the net shortwave radiation is higher than the net longwave radiation.

**Table 3**  
Spearman's correlation for every site between  $R_n$  and the all variables used in this work.

Station	$G_i$	$\alpha$	$k_t$	$\cos\theta_z$	$E_o$	$T_a$	$k_t \cos\theta_z$
BAR	0,936**	0,054**	0,673**	0,827**	-0,339**	0,265**	0,937**
BUD	0,981**	-0,524**	0,762**	0,907**	-0,339**	0,392**	0,979**
GOB	0,978**	-0,896**	0,790**	0,941**	0,435**	0,157**	0,974**
IZA	0,981**	0,367**	0,754**	0,946**	-0,101**	0,164**	0,975**
PAY	0,985**	-0,539**	0,779**	0,882**	-0,434**		0,985**
TAT	0,950**	-0,255**	0,641**	0,685**	-0,344**	0,476**	0,957**
TOR	0,938**	-0,043**	0,748**	0,800**	-0,482**		0,940**

\*\* The correlation is significant at the 0.01 level.

**Table 4**  
Statistical parameters for  $k_t$ ,  $\alpha$ ,  $R_n$ ,  $G_i$ ,  $L_i$ ,  $G_t$ ,  $L_t$ . Includes arithmetic mean (Ave), standard deviation (SD), first quartile (Q1), third quartile (Q3), median (Md), and the 10th (P10) and 90th (P90) percentiles.

	Ave $\pm$ SD	Q1	Md	Q3	P10	P90		Ave $\pm$ SD	Q1	Md	Q3	P10	P90
<b>BAR (2022)</b>							<b>IZA (2022)</b>						
$R_n(Wm^{-2})$	90 $\pm$ 100	15	62	146	-15	251	$R_n(Wm^{-2})$	350 $\pm$ 180	201	357	493	93	593
$G_i(Wm^{-2})$	270 $\pm$ 150	149	240	379	97	502	$G_i(Wm^{-2})$	650 $\pm$ 260	435	663	863	280	1008
$L_i(Wm^{-2})$	280 $\pm$ 50	267	294	317	221	328	$L_i(Wm^{-2})$	250 $\pm$ 30	232	250	274	218	299
$G_t(Wm^{-2})$	140 $\pm$ 130	35	86	229	19	369	$G_t(Wm^{-2})$	110 $\pm$ 50	69	108	142	42	169
$L_t(Wm^{-2})$	320 $\pm$ 40	305	331	345	270	361	$L_t(Wm^{-2})$	450 $\pm$ 60	404	446	493	371	536
$\alpha$	0,48 $\pm$ 0,33	0,18	0,24	0,84	0,16	0,89	$\alpha$	0,16 $\pm$ 0,02	0,15	0,16	0,17	0,14	0,17
$k_t$	0,48 $\pm$ 0,18	0,33	0,48	0,63	0,23	0,72	$k_t$	0,78 $\pm$ 0,10	0,76	0,81	0,84	0,68	0,86
<b>BUD (2022)</b>							<b>PAY (2022)</b>						
$R_n(Wm^{-2})$	250 $\pm$ 160	119	235	371	54	480	$R_n(Wm^{-2})$	270 $\pm$ 180	115	233	403	58	538
$G_i(Wm^{-2})$	490 $\pm$ 240	285	458	677	182	843	$G_i(Wm^{-2})$	470 $\pm$ 250	246	430	675	157	848
$L_i(Wm^{-2})$	330 $\pm$ 50	284	331	367	248	392	$L_i(Wm^{-2})$	320 $\pm$ 40	293	329	357	255	376
$G_t(Wm^{-2})$	100 $\pm$ 50	66	98	136	43	168	$G_t(Wm^{-2})$	100 $\pm$ 50	57	96	138	36	163
$L_t(Wm^{-2})$	460 $\pm$ 80	397	451	515	352	566	$L_t(Wm^{-2})$	430 $\pm$ 60	385	424	464	352	498
$\alpha$	0,22 $\pm$ 0,03	0,19	0,21	0,24	0,18	0,26	$\alpha$	0,22 $\pm$ 0,03	0,20	0,22	0,24	0,18	0,26
$k_t$	0,65 $\pm$ 0,14	0,57	0,69	0,75	0,42	0,79	$k_t$	0,61 $\pm$ 0,19	0,48	0,69	0,76	0,31	0,79
<b>GOB (2022)</b>							<b>TAT (2022)</b>						
$R_n(Wm^{-2})$	270 $\pm$ 150	149	270	386	49	479	$R_n(Wm^{-2})$	230 $\pm$ 170	100	183	321	54	501
$G_i(Wm^{-2})$	670 $\pm$ 270	456	688	882	270	1033	$G_i(Wm^{-2})$	460 $\pm$ 240	214	352	563	136	799
$L_i(Wm^{-2})$	350 $\pm$ 30	322	346	367	303	391	$L_i(Wm^{-2})$	360 $\pm$ 60	309	367	415	266	434
$G_t(Wm^{-2})$	240 $\pm$ 90	176	254	308	110	348	$G_t(Wm^{-2})$	90 $\pm$ 50	49	87	135	27	166
$L_t(Wm^{-2})$	510 $\pm$ 60	467	510	552	423	581	$L_t(Wm^{-2})$	450 $\pm$ 60	407	451	486	371	517
$\alpha$	0,37 $\pm$ 0,05	0,35	0,37	0,39	0,33	0,41	$\alpha$	0,50 $\pm$ 0,2	0,31	0,53	0,70	0,20	0,75
$k_t$	0,72 $\pm$ 0,09	0,70	0,75	0,78	0,61	0,80	$k_t$	0,23 $\pm$ 0,05	0,20	0,22	0,26	0,18	0,29
<b>TOR (2019)</b>													
$R_n(Wm^{-2})$	160 $\pm$ 140	42	111	250	11	376							
$G_i(Wm^{-2})$	320 $\pm$ 230	128	251	481	71	670							
$L_i(Wm^{-2})$	310 $\pm$ 50	284	317	349	252	370							
$G_t(Wm^{-2})$	80 $\pm$ 50	32	67	121	17	161							
$L_t(Wm^{-2})$	400 $\pm$ 50	354	398	435	325	463							
$\alpha$	0,27 $\pm$ 0,14	0,21	0,23	0,25	0,19	0,31							
$k_t$	0,49 $\pm$ 0,24	0,26	0,50	0,72	0,16	0,78							

The component of  $R_n$  with the highest values is  $G_i$ , except at BAR and TOR, where the highest values are observed for  $L_t$  and at TAT the found value for  $L_t$  is close to  $G_i$ . Conversely, the variable with the lowest values for all sites is  $G_t$ , as expected. According to the  $k_t$  values, BAR and TOR present the lowest values of solar radiation, with a mean value for  $G_i$  of (270  $\pm$  150)  $Wm^{-2}$  and (320  $\pm$  230)  $Wm^{-2}$ , respectively. In contrast, GOB and IZA present the highest values (670  $\pm$  270)  $Wm^{-2}$  and (650  $\pm$  260)  $Wm^{-2}$  respectively, clear skies predominate in both locations and, additionally, IZA's higher altitude results in higher levels of received radiation. This indicates 148 % more solar radiation at GOB compared to BAR. Consequently, these results are extrapolated to  $G_t$ . Regarding the  $R_n$  values, IZA also presents the maximum values with a mean of (350  $\pm$  180)  $Wm^{-2}$ , while BAR records the lowest values, averaging (90  $\pm$  100)  $Wm^{-2}$ . This represents 289 % higher  $R_n$  at IZA compared to BAR. In terms of P90, the highest value is found at IZA (593  $Wm^{-2}$ ), and the lowest one is recorded at BAR (251  $Wm^{-2}$ ). Considering the median values, the most substantial difference among stations occurs between

IZA and BAR for the variable  $R_n$  (476 %) and between GOB and TOR for  $G_t$  (279 %).

Regarding  $L_i$ , the highest mean value is also found at TAT (360  $\pm$  60)  $Wm^{-2}$  and GOB (350  $\pm$  30)  $Wm^{-2}$ , and the lowest at IZA (250  $\pm$  30)  $Wm^{-2}$ . Thus, the difference between sites is minimal, with TAT having 44 % higher  $L_i$  compared to IZA. This difference is associated with the greater cloud coverage at TAT, as indicated by the  $k_t$  index, given the role of clouds in longwave radiation emission. This difference increases for  $L_t$ , with a 59 % difference between GOB and BAR, having a mean value of (510  $\pm$  60)  $Wm^{-2}$  and (320  $\pm$  40)  $Wm^{-2}$ , respectively, associated with higher surface temperatures at BAR, which has an arid climate, in contrast to BAR, which has a polar climate and a surface covered by snow for extended periods. Among all variables,  $R_n$  presents the highest variability across all sites, generally followed by  $G_i$ , except at BAR, where high variability is observed not for  $G_i$  but for  $G_t$ . In contrast  $L_i$  and  $L_t$  present the lowest variability for all sites, indicating that the main contribution to  $R_n$ , including its high variability, is from shortwave

radiation.

Considering variability as the ratio between the difference P90 and P10 and the median value, BAR exhibits the highest variability for  $R_n$  (429 %), while IZA shows the lowest (140 %). For  $G_{\downarrow}$ , TOR presents the highest variability (239 %), while PAY demonstrates the lowest (111 %). It is noteworthy to highlight the substantial variability found at BAR for  $G_{\uparrow}$  (407 %) in line with the high variability detected in  $\alpha$ . Conversely, GOB and IZA show the lowest variability for  $G_{\uparrow}$  (94 % and 99 %, respectively). For  $L_{\downarrow}$ , the highest variability is found at TAT (46 %) and BUD (43 %) and the lowest at GOB (27 %). Finally, for  $L_{\uparrow}$ , also the highest is found at BUD (47 %) and the lowest at BAR (27 %). This initial analysis of the data for each site reveals significant differences among them, reflecting their different climatic characteristics. BAR and GOB particularly stand out in terms of extreme characteristics. Regardless, the primary factor controlling  $R_n$  values for all sites during the day is the downwelling shortwave radiation, which forms the main basis of the model proposed in this work.

4.1.2. Seasonal characterization

To assess potential seasonal dependence, Fig. 3 presents boxplots for all stations across the four seasons: winter (January, February and March), spring (April, May and June), summer (July, August and September) and autumn (October, November and December). A clear seasonal pattern emerges for all stations and variables ( $R_n$  and its components), with maximum values occurring in the warmer seasons and minimum values in the colder seasons. Notably,  $R_n$  and shortwave radiation generally peak in spring or summer and reach its lowest values

in autumn or winter. Interestingly, the pattern is reversed at GOB. Seasonal variations in shortwave radiation result from the longer path of the radiation with greater solar zenith angles. Meanwhile, upwelling longwave radiation exhibits higher values during the hottest seasons when the surface temperatures are elevated. In the case of downwelling longwave radiation, clouds play a fundamental and complex role.

The seasonal variation in median  $R_n$  values ranges between  $83 \text{ Wm}^{-2}$  at IZA and  $212 \text{ Wm}^{-2}$  at PAY, with  $165 \text{ Wm}^{-2}$  at BAR, where the value is negative in winter ( $-40 \text{ Wm}^{-2}$ ). The highest absolute differences occur at BAR and TOR, and the smallest at IZA, with seasonal variability of 413 %, 408 % and 26 % for BAR, TOR and IZA respectively. For  $G_{\downarrow}$ , the differences vary between  $125 \text{ Wm}^{-2}$  and  $316 \text{ Wm}^{-2}$  (TAT and PAY, respectively), with higher variability at TOR (302 %) and lower at IZA (27 %). Similarly, for  $G_{\uparrow}$ , differences range between  $10 \text{ Wm}^{-2}$  at TAT and  $189 \text{ Wm}^{-2}$  at BAR, with a higher variability in BAR and TOR (511 % and 400 %, respectively), and lower in GOB, also at IZA and TAT (33 %, 35 % and 12 %, respectively). Consequently, seasonal differences are consistently more pronounced at BAR and TOR and less pronounced at TAT and GOB for both  $R_n$  and shortwave radiation. Similar to shortwave radiation, a seasonal pattern emerges for longwave radiation, with higher values in summer (rather than spring) and lower in winter (rather than autumn), except at GOB, which shows the inverse pattern, similar to shortwave radiation. The seasonal differences in median values for  $L_{\downarrow}$  vary from  $50 \text{ Wm}^{-2}$  at GOB and  $52 \text{ Wm}^{-2}$  at TOR to  $150 \text{ Wm}^{-2}$  at BAR and TAT, presenting BAR relative high variability in winter (56 %), with values lower than those at other sites, which is associated with low air temperatures. Similarly, TAT shows higher variability in winter and

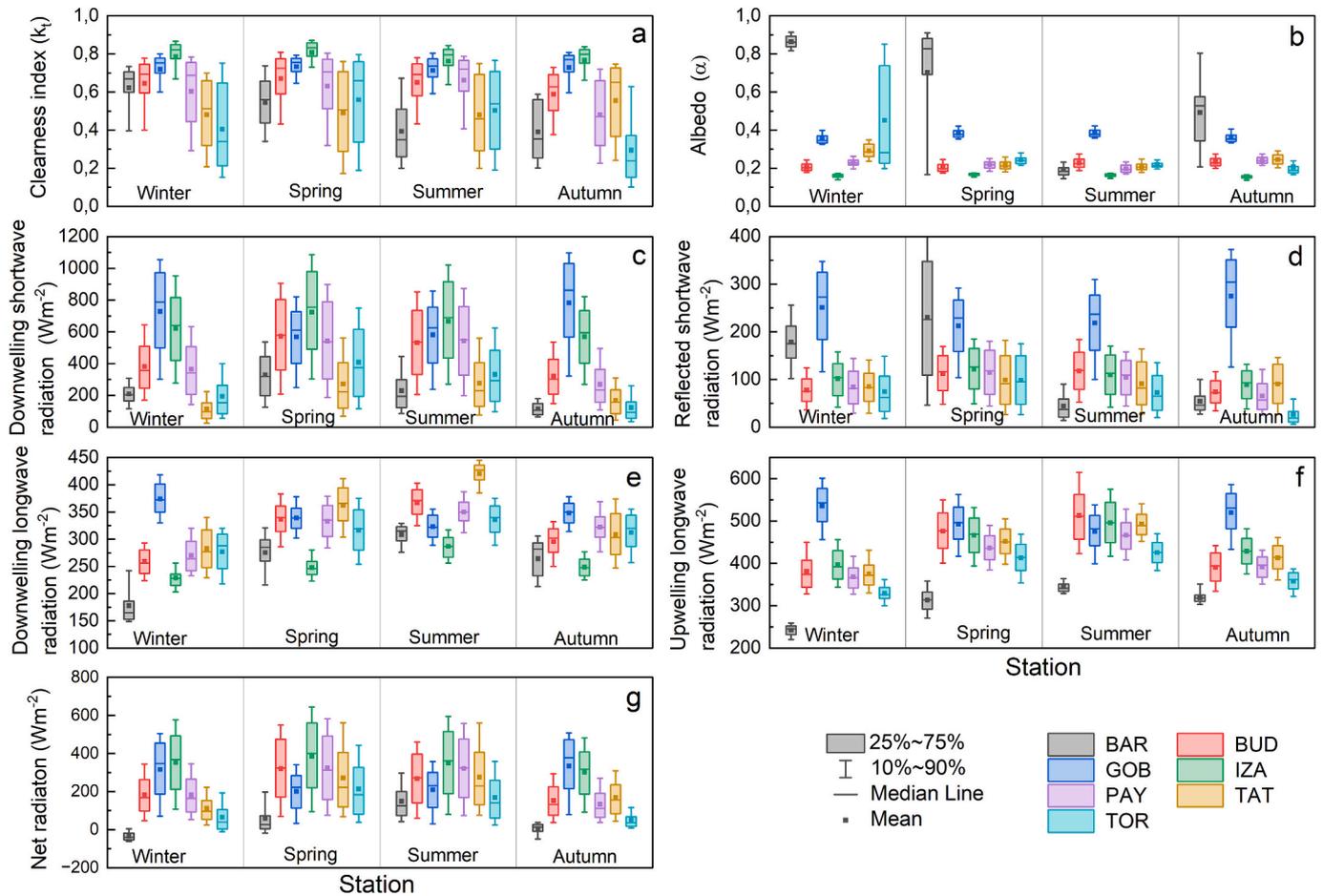


Fig. 3. Boxplot for all seasons across each station: a) clearness index; b) albedo; c) downwelling shortwave radiation; d) reflected shortwave radiation; e) downwelling longwave radiation; f) upwelling longwave radiation; g) net radiation. The whiskers represent the P10 and P90 percentiles, the box edges correspond to the P25 and P75 percentiles, the midline represents the median and the dot indicates the mean value.

autumn (42 %), a pattern linked to the increased cloud cover during these seasons, while for  $L_{\uparrow}$ , they range from  $61 \text{ Wm}^{-2}$  at GOB to  $137 \text{ Wm}^{-2}$  at BUD. Seasonal variability for  $L_{\downarrow}$  ranges from 16 % at GOB and 91 % at BAR, and for  $L_{\uparrow}$ , between 13 % at GOB and 39 % at BAR. This component ( $L_{\downarrow}$ ) exhibits the lowest variability, followed by  $L_{\uparrow}$ . Thus, the highest seasonal differences are found in shortwave radiation rather than longwave radiation, as mentioned in the global analysis. Additionally, for  $R_n$  and shortwave radiation, the differences between sites are more pronounced in colder seasons, with high variability values reaching 2593 % for  $R_n$  between GOB and BAR and 1500 % for  $G_{\uparrow}$  between GOB and TOR in autumn. In winter the differences are 125 % and 122 % for  $L_{\downarrow}$  and  $L_{\uparrow}$ , respectively. Minor differences are observed in hotter seasons for all variables.

4.1.3. Sky conditions characterization

The presence of clouds significantly influences net radiation because clouds affect both shortwave and longwave net radiation (Alados et al., 2003). Therefore, when direct cloud measurements are unavailable, an alternative method to account for cloud effects is to use the clearness index as an indicator of the overall atmospheric transmittance, which relates to the presence of clouds and also aerosols. This index also depends on the Sun's position (Foyo-Moreno et al., 2023), exhibiting certain seasonal variations. In this work, we have classified the data into categories of  $k_t$  according to the following thresholds:  $k_{t1}$  ( $0,0 < k_t \leq 0,35$ ),  $k_{t2}$  ( $0,35 < k_t \leq 0,7$ ) and  $k_{t3}$  ( $0,70 < k_t < 1,0$ ). The most frequent category for IZA, GOB and PAY is  $k_{t3}$  (associated with clear skies), with relative frequencies of 88 %, 74 % and 47 % respectively. At BAR, the  $k_{t2}$  category is the most frequent (59 %) and the  $k_{t1}$  is the least frequent. At TAT, also the  $k_{t2}$  category is also most frequent (43 %) and the less frequent is  $k_{t3}$  (26 %). At BUD, both  $k_{t2}$  and  $k_{t3}$  categories present similar values (48 % and 47 %, respectively), while at TOR, the values are also similar, with  $k_{t1}$  being the most frequent category (38 %).

Fig. 4 is analogous to Fig. 3, now distinguishing between the three considered categories of  $k_t$ . Overall, a clear pattern emerges across all sites. The highest values for  $R_n$  and its components consistently appear in the  $k_{t3}$  category (clear skies), while the lowest values typically occur in the  $k_{t1}$  category (cloudy skies). An exception to this is  $L_{\downarrow}$ , which shows an inverse relationship, with higher values in the  $k_{t1}$  category and lower values in the  $k_{t3}$  category, as expected due to clouds emitting longwave radiation. However, there are some exceptions. For instance, at BUD,  $L_{\downarrow}$  shows an inverse pattern compared to the other sites, although the differences are minimal, with a variability of only 4 % between categories. Similarly, at BAR, for  $L_{\uparrow}$  display low variability (5 %). Additionally, the lowest value for  $R_n$  and  $G_{\uparrow}$  are found in category  $k_{t2}$  rather than  $k_{t1}$ , at BAR and GOB, respectively.

For every site, the differences between  $k_t$  categories range from 200 % at BAR to close to 600 % at TOR (IZA also presents a high variability with a value of 577 %) and the difference between sites is more high for intermediate conditions (377 %) and low for cloudy conditions (144 %). The median value for  $R_n$  in the  $k_{t3}$  category is  $117 \text{ Wm}^{-2}$  at BAR and  $411 \text{ Wm}^{-2}$  at TAT. For this category, the differences between sites are 251 %. For  $G_{\uparrow}$ , median values range between  $482 \text{ Wm}^{-2}$  and  $773 \text{ Wm}^{-2}$  at BAR and GOB, respectively, in the  $k_{t3}$  category, indicating that, for clear skies, the differences observed between GOB and BAR are primarily explained by latitude, i.e. the high-latitude site (BAR) exhibits the lowest values, while the low-latitude site (GOB) shows the highest values. The values range between  $105 \text{ Wm}^{-2}$  and  $294 \text{ Wm}^{-2}$  at TOR and GOB, respectively, in the  $k_{t1}$  category. The differences between sites are 60 % in the  $k_{t3}$  category and 180 % for the  $k_{t1}$  category, indicating higher variability under cloudy skies. In general, for every site there is a higher variability for cloudy skies and lower for clear skies. The differences between  $k_t$  categories for each site are generally similar, with values close to 300 %. However, the greatest difference is observed at IZA (541 %), which is attributed to the high altitude of this mountain site, while

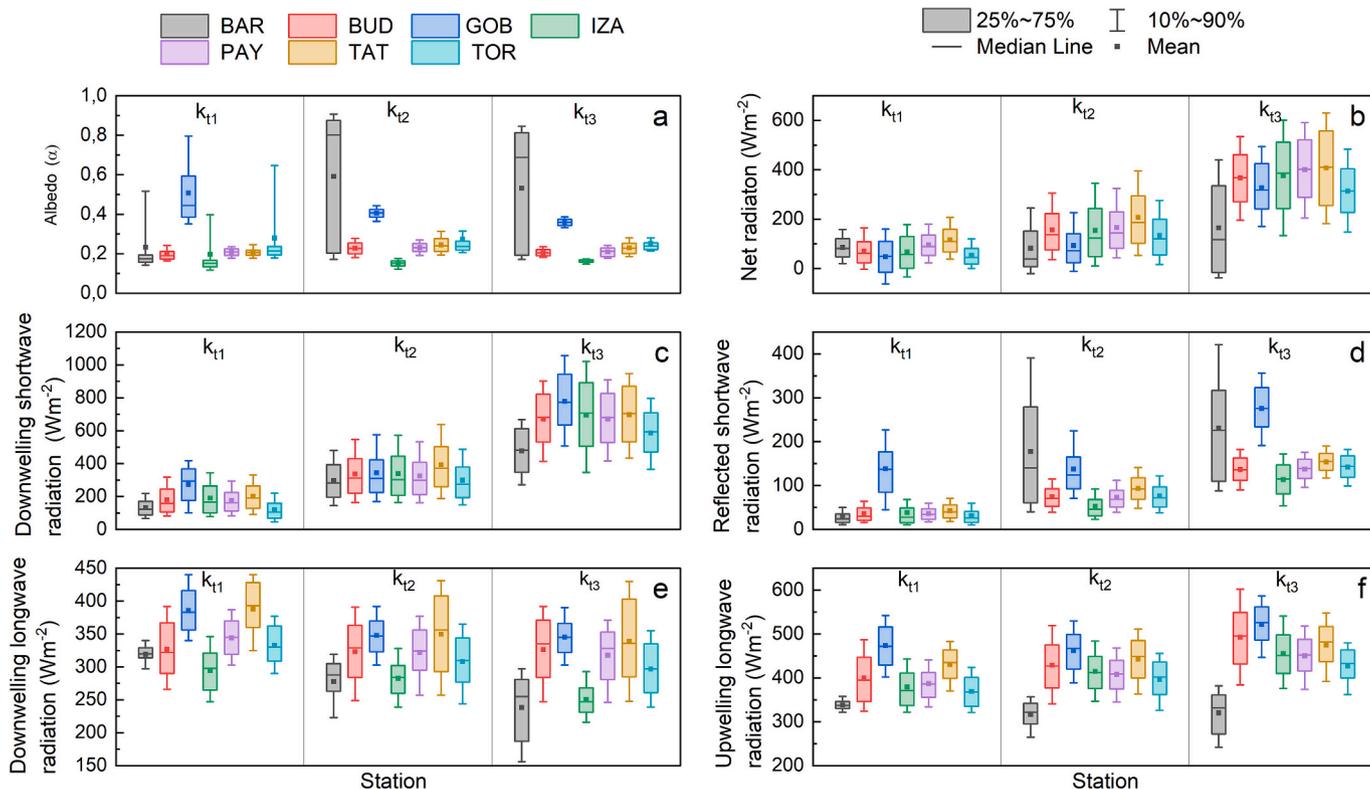


Fig. 4. Boxplot for the three categories of the clearness index  $k_t$  ( $k_{t1}$ :  $0,0 < k_t \leq 0,35$ ,  $k_{t2}$ :  $0,35 < k_t \leq 0,7$ , and  $k_{t3}$ :  $0,70 < k_t < 1,0$ ) at each station: a) albedo; b) downwelling shortwave radiation; c) reflected shortwave radiation; d) downwelling longwave radiation; e) upwelling longwave radiation; f) net radiation. The whiskers represent the P10 and P90 percentiles, the box edges correspond to the P25 and P75 percentiles, the midline represents the median and the dot indicates the mean value.

the smallest difference is observed at GOB (163 %), where the high values of  $G_l$  are primarily due to the low altitude. Regarding the  $G_l$  variable, as for  $G_t$ , at every site the variability is low under clear skies, except at BAR with a higher variability for intermediate conditions. There is significant variability distinguishing between these  $k_t$  categories at BAR (842 %) and low variability at GOB (123 %). The high variability at BAR is associated with the high range of values of albedo against the low range at GOB. As expected, the differences between sites are higher for the  $k_{t1}$  category (471 %) compared to the  $k_{t3}$  category (140 %). Considering the longwave radiation, the pattern followed for  $L_l$ , as previously mentioned, is inverse, with high values in the  $k_{t1}$  category (except BUD) and low values in the  $k_{t3}$  category (except BUD, although there is a minor difference and PAY), an expected result because the clouds emit longwave radiation. The values range from 298  $Wm^{-2}$  at IZA to 393  $Wm^{-2}$  at TAT. For all categories, the highest values are observed at TAT and the lowest at IZA. That is, under the same cloudy conditions, the values of  $L_l$  are primarily influenced by air temperature, with lower values observed at higher-altitude sites. The variability between categories is minor at BUD (4 %) and major at BAR (25 %). Now, for  $L_t$  the pattern is opposite to that  $L_l$ , with values highest in the  $k_{t3}$  category, and lowest in the  $k_{t1}$  category, except at BAR with values very similar between categories, this result is due to higher surface temperature for clear skies, presenting BAR the lowest variability between categories (5 %), in contrast to BUD (25 %), which is attributed to the low surface temperature of the high-latitude site. These differences are primarily due to air temperature, with lower values at mountain sites. The median values range from 332  $Wm^{-2}$  at BAR to 526  $Wm^{-2}$  at GOB, in fact, for all categories the lowest values are found at BAR and the highest are found at GOB, due to the values of surface temperature, this result indicates that under the same sky condition, the main factor affecting to  $L_t$  is surface temperature. The differences between sites for  $L_l$  is higher for clear skies (40 %) compared to cloudy skies (32 %) or skies with intermediate conditions (25 %). Similarly, for  $L_t$ , the differences between sites are also higher under clear skies (58 %) compared to cloudy skies (40 %).

Under clear skies, both shortwave radiation and  $L_t$  exhibit higher values, which are associated with the solar position and elevated surface temperatures, respectively. The maximum values observed under clear skies and the minimum values under cloudy conditions for shortwave radiation highlights the significant attenuation caused by clouds. However, the highest values for  $L_l$  are found under cloudy conditions, which can be explained by the high emissivity of clouds (depending on the clouds properties). In general, clouds increase downwelling longwave radiation, as they are one of the main factors affecting  $L_l$ .

For  $R_n$  and its shortwave components, the percentage differences distinguishing these three categories are generally high across all sites, with BAR exhibiting particularly high variability in  $G_t$  (848 %), and GOB showing low variability for the same variable (123 %). However, for the longwave components, these differences are minimal. In general, the differences between sites are greater under clear skies for  $R_n$  and longwave radiation, whereas for shortwave radiation, the differences are more pronounced under cloudy skies.

As verified in the seasonal study described in section 4.1.b, the greatest differences across all the stations are consistently observed for  $R_n$  and its shortwave radiation components, while smaller differences are found in its longwave components. This indicates that the primary contribution to  $R_n$  is from shortwave radiation. The effects of clouds on shortwave and longwave radiation are opposite, attenuating shortwave radiation while increasing downwelling longwave radiation. However, in terms of net balance, shortwave radiation is the primary driver. It is worthy to highlight two sites with very different extreme characteristics: BAR and GOB. GOB exhibits the highest values for all components in every  $k_t$  category, while BAR presents the lowest values for all categories for  $L_t$ , as expected due to its low temperatures, with minimum values of  $R_n$  under clear and intermediate conditions. Conversely, the minimum values for  $L_l$  have been found at IZA, which has a low relative frequency

of cloudy skies.

An analysis of relative frequency for each season at the six stations reveals that for GOB and IZA the most frequent category is always  $k_{t3}$ , with values consistently above 70 %, and exceeding 83 % at IZA. However, at TOR, the most frequent category for all seasons is  $k_{t1}$ , with a value of 72 % in autumn. At BUD, the percentage of  $k_{t2}$  and  $k_{t3}$  is very similar for all seasons except autumn, where the frequency of  $k_{t1}$  increases. At BAR, the most frequent category is  $k_{t2}$ , except in summer where  $k_{t1}$  predominates with 50 %, and in autumn where  $k_{t1}$  and  $k_{t2}$  categories are nearly equal in frequency (49 %). Finally, at TAT, the most frequent category is also  $k_{t2}$  with values close to 40 % and with higher predominance in winter (61 %).

Summarising, in the seasonal analysis, the major differences between sites are found during the colder seasons for  $R_n$  and all its components, while minor differences are observed in the warmer seasons. In the analysis by  $k_t$  categories, the major differences are found under cloudy conditions for shortwave radiation, whereas for longwave radiation, the differences (although much smaller) are also found under cloudy conditions for  $L_l$  but for  $L_t$  the differences are higher under clear conditions, by the difference between the surface temperatures for the sites with extreme climate (GOB and BAR). This highlights the different roles of clouds depending on the range of wavelength. Furthermore, when comparing differences for a given site, it is interesting to note that the differences for the shortwave components are always higher for sky conditions than for seasons. However, for the longwave components, although much smaller, the differences are greater for seasons than for  $k_t$  categories, except at BAR for  $L_t$ . Additionally, it is noteworthy that at TOR, the differences for  $R_n$  and its shortwave components exhibit similar values for both seasons and sky conditions. The differences between sites for all variables are always lower for categories of  $k_t$  than distinguishing between seasons. Thus, the variability of  $R_n$ , which is mainly determined by the variability in shortwave radiation, is captured by the type of sky using the  $k_t$  index, showing minor variability between sites, and then a model using the index  $k_t$  could be more universal. This result also corroborates the basis of the model proposed in this work and presented in the next section.

#### 4.2. Model validation

Model validation employed Mean Bias Error (MBE), Root Mean Squared Error (RMSE) and  $R^2$ , with definitions and results presented in Table 5 and Table 6, respectively. The later Table includes validation results for all stations. In general, the model presents good results across all the sites (except at BAR), with  $R^2$  exceeding 0,942. The poorer performance of the model at BAR will be evaluated later. The model tends to overestimate  $R_n$  across the sites, as indicated by MBE values ranging from 3  $Wm^{-2}$  to 132  $Wm^{-2}$  and RMSE values from 25  $Wm^{-2}$  to 141  $Wm^{-2}$  at PAY and GOB, respectively. Excluding BAR and GOB, MBE values range from 3  $Wm^{-2}$  to 44  $Wm^{-2}$  and RMSE values from 36  $Wm^{-2}$  to 62  $Wm^{-2}$ .

**Table 5**  
Summary of statistical metrics used in the study.

Metrics	Equation
Determination coefficient, $R^2$	$\frac{(\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})])^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}$
Mean bias error, MBE	$\frac{1}{n} \sum_{i=1}^n (y_i - x_i)$
Root mean squared error, RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2}$
	$x_i$ experimental values $y_i$ estimated values $\bar{x}$ mean of experimental values $\bar{y}$ mean of estimated values $n$ number of values

**Table 6**

Statistical results for the model using the dataset from each site. Includes mean bias error (MBE), root mean squared error (RMSE), coefficient of determination ( $R^2$ ).

Station	$R_n$ $Wm^{-2}$	MBE $Wm^{-2}$	RMSE $Wm^{-2}$	$R^2$
BAR	93	51	118	0,249
BUD	251	31	48	0,950
GOB	267	132	141	0,942
IZA	348	44	62	0,943
PAY	369	4	25	0,981
TAT	230	3	36	0,954
TOR	156	16	36	0,953

Jiang et al. (2015) developed a multivariate linear regression model to estimate  $R_n$  using shortwave radiation, achieving a global model RMSE of  $40 Wm^{-2}$  and a conditional mode ranging between  $18 Wm^{-2}$  and  $54 Wm^{-2}$ . Their findings suggest that incorporating net shortwave radiation, which accounts for surface albedo effects, yields better results compared to using incident shortwave radiation in nearly all eight evaluated models. However, the complexity of Jiang et al.'s model is notable, as it integrates solar shortwave radiation with conventional meteorological observations (air temperature and relative humidity) and satellite-derived measurements (NDVI and albedo). Similarly, Ferreira et al. (2020) utilized satellite data and found MBE values ranging from  $39 Wm^{-2}$  to  $85 Wm^{-2}$  and for RMSE from  $45 Wm^{-2}$  to  $89 Wm^{-2}$ . Their work highlights the potential of combining remote sensing data with more accessible in-situ weather data to provide spatiotemporal estimates of  $R_n$ , particularly in regions where direct net radiation data are unavailable.

An evaluation of the model's performance, distinguished by seasons, is presented in Table 7. Notably, in this seasonal analysis, the model performs well at BAR during the summer, with a  $R^2$  of 0,922. Consistent with the earlier global analysis, the largest errors are found at GOB. With

**Table 7**

Statistical results for the model, categorized by seasons (Winter (Wi), Spring (Sp), Summer (Su), and Autumn (Au)), using the dataset from each site. Includes mean bias error (MBE), root mean squared error (RMSE), coefficient of determination ( $R^2$ ).

Station		$R_n$ $Wm^{-2}$	MBE $Wm^{-2}$	RMSE $Wm^{-2}$	$R^2$
BAR	Wi	-34	98	141	0,018
	Sp	60	119	156	0,241
	Su	149	-32	43	0,922
	Au	7	39	57	0,0001
BUD	Wi	182	20	32	0,950
	Sp	120	25	37	0,975
	Su	268	49	68	0,918
	Au	153	13	23	0,968
GOB	Wi	316	118	128	0,944
	Sp	202	135	142	0,942
	Su	210	139	147	0,944
	Au	334	137	148	0,949
IZA	Wi	353	7	41	0,944
	Sp	384	69	78	0,971
	Su	349	66	78	0,956
	Au	301	23	34	0,971
PAY	Wi	182	10	29	0,946
	Sp	323	1	23	0,985
	Su	322	3	25	0,982
	Au	134	1	25	0,926
TAT	Wi	113	34	49	0,845
	Sp	271	-1	31	0,972
	Su	276	-20	32	0,980
	Au	168	23	40	0,913
TOR	Wi	66	25	52	0,730
	Sp	213	20	36	0,964
	Su	169	13	28	0,970
	Au	51	-1	22	0,845

the exception of BAR,  $R^2$  values are consistently over 0,90 across most sites, although TOR shows lower  $R^2$  values of 0,730 in winter and 0,845 in autumn. In general, the model tends to overestimate  $R_n$  in every season, except for summer at BAR and summer at TAT and autumn at TAT and TOR. Regarding MBE and RMSE, for all sites, the smallest errors are observed in the colder seasons at most sites, with the exception at PAY. Specifically, MBE values range from  $-1 Wm^{-2}$  at TOR to  $13 Wm^{-2}$  at BUD in autumn, and RMSE values in autumn range from  $22 Wm^{-2}$  at TOR to  $34 Wm^{-2}$  at IZA (excluding BAR and GOB). Conversely, the worst results occur during the warmer seasons, except at PAY, TAT and TOR, where winter shows better performance. MBE values range from  $-1$  at TAT to  $69 Wm^{-2}$  at IZA, and RMSE values range from  $23 Wm^{-2}$  at PAY to  $78 Wm^{-2}$  at IZA.

The evaluation was also conducted based on  $k_t$  categories (Table 8). The  $R^2$  values are consistently above 0,770 for all sites and categories, except at BAR for the  $k_{t2}$  and  $k_{t3}$  categories. Overall, the model tends to overestimate net radiation across all categories, except for the  $k_{t1}$  category at BAR, PAY, TAT and TOR. Notably, the  $k_{t1}$  category consistently yields the best results, while  $k_{t3}$  generally produces poorer outcomes. However, exceptions include PAY, which exhibits better results in the  $k_{t3}$  category and poorer results in the  $k_{t2}$  category. In general, the best model performance is observed for skies with low transparency, i.e., corresponding to the presence of clouds and/or aerosols.

For the  $k_{t1}$  category, MBE values range from  $-27 Wm^{-2}$  to  $89 Wm^{-2}$  at BAR and GOB, respectively. Excluding these sites, MBE values range from  $-15 Wm^{-2}$  to  $22 Wm^{-2}$  at PAY and IZA, respectively, while RMSE from  $23 Wm^{-2}$  to  $45 Wm^{-2}$  at TOR and IZA, respectively. Similarly, for the  $k_{t3}$  category, MBE ranges from  $7 Wm^{-2}$  to  $147 Wm^{-2}$ , whilst RMSE between  $21 Wm^{-2}$  and  $153 Wm^{-2}$  at PAY and GOB, respectively. These findings suggest that the model performs less satisfactorily under very clear conditions, possibly due to its simplicity, which relies solely on predictors related to shortwave radiation and does not account for longwave radiation. In contrast, the model demonstrates better performance under cloudy conditions and/or high aerosol loads. Thus, the model performs satisfactorily across all the scenarios, leveraging instantaneous measurements rather than daily values for enhanced precision. The simplicity of the model is its primary advantage, as it requires only one measured variable ( $G_1$ ), which is typically available at most radiometric stations.

**Table 8**

Statistical results for the model, categorized by  $k_t$  intervals ( $k_{t1}$ :  $0,0 < k_t \leq 0,35$ ,  $k_{t2}$ :  $0,35 < k_t \leq 0,70$ , and  $k_{t3}$ :  $0,70 < k_t < 1,0$ ), using the dataset from each site. Includes mean bias error (MBE), root mean squared error (RMSE), coefficient of determination ( $R^2$ ).

Station		$R_n$ $Wm^{-2}$	MBE $Wm^{-2}$	RMSE $Wm^{-2}$	$R^2$
BAR	$k_{t1}$	85	-27	38	0,807
	$k_{t2}$	81	77	129	0,166
	$k_{t3}$	164	111	170	0,543
BUD	$k_{t1}$	70	13	33	0,762
	$k_{t2}$	156	24	39	0,912
	$k_{t3}$	367	40	57	0,908
GOB	$k_{t1}$	48	89	116	0,283
	$k_{t2}$	93	91	101	0,829
	$k_{t3}$	327	147	153	0,915
IZA	$k_{t1}$	67	22	45	0,776
	$k_{t2}$	154	26	58	0,866
	$k_{t3}$	376	46	64	0,939
PAY	$k_{t1}$	96	-15	29	0,852
	$k_{t2}$	166	7	28	0,950
	$k_{t3}$	401	7	21	0,984
TAT	$k_{t1}$	116	-19	31	0,870
	$k_{t2}$	207	10	38	0,932
	$k_{t3}$	408	17	39	0,973
TOR	$k_{t1}$	53	-3	23	0,786
	$k_{t2}$	134	25	40	0,911
	$k_{t3}$	313	34	44	0,960

### 4.3. Comparison with other models

This section aims to compare the proposed model with other empirical models. Our evaluation is based on the previous work of Jiang et al. (2015), which analyzed models for estimating  $R_n$  using incident and/or net shortwave radiation, along with other meteorological variables, for global application. They found that all empirical net radiation fitting models that included shortwave radiation could be used for net radiation estimation in most situations because their fitting accuracy was acceptable despite some differences from each other. The comparison distinguishes between models that incorporate albedo as an input and those that do not, with a particular focus on models using reflected shortwave radiation as an input variable. The models not using albedo data are as follows:

$$\text{Model 1 (Mod 1): } R_n = a_1 G_1 + b_1.$$

$$\text{Model 2 (Mod 2): } R_n = a_2 G_1 + b_2 k_t + c_2$$

$$\text{Model 3 (Mod 3): } R_n = a_3 G_1 + b_3 T_{a,\circ C} + c_3 E_0 + d_3$$

The models including albedo are:

$$\text{Model 4 (Mod 4): } R_n = a_4 G_1 (1 - \alpha) + b_5$$

$$\text{Model 5 (Mod 5): } R_n = a_5 G_1 (1 - \alpha) + b_5 \sigma T_{a,K}^4 + c_6$$

$$\text{Model 6 (Mod 6): } R_n = a_6 \left[ G_1 (1 - \alpha) + D_1 T_{a,K}^6 - \sigma T_{a,K}^4 \right] + b_6 k_t + c_6$$

where  $T_{a,\circ C}$  is the mean air temperature ( $^{\circ}C$ ),  $E_0$  is the inverse relative Earth–Sun distance,  $T_{a,K}$  is the absolute air temperature,  $\sigma$  is the Stefan–Boltzmann constant ( $5,67 \times 10^{-8} \text{ W K}^{-4} \text{ m}^{-2}$ ),  $D_1$  is an empirical constant ( $5,31 \times 10^{-13} \text{ W K}^{-6} \text{ m}^{-2}$  (Swinbank, 1963)) and  $a_i, b_i, c_i$  and  $d_i$  are the coefficients specified in Table 9. While these coefficients differ from those originally used by the authors, they have been derived by Jiang et al. (2015). The decision to use these coefficients is based on Jiang et al.'s extensive use of a large and diverse database spanning several years. This broader dataset enhances the universality and applicability of the empirical models.

Table 10 shows the results of this comparison. Among the models of the first type (Mod 1, Mod 2, Mod 3), excluding our model, the simplest model (Mod 1), which only utilizes  $G_1$  as an input variable, exhibits the lowest MBE. Excluding BAR due to its very low  $R^2$  value, and notably GOB, which presents large errors with an MBE of  $149 \text{ Wm}^{-2}$  and a RMSE of  $155 \text{ Wm}^{-2}$ , the MBE values for Mod 1 range from  $18 \text{ Wm}^{-2}$  to  $57 \text{ Wm}^{-2}$ , with corresponding RMSE values ranging from  $35 \text{ Wm}^{-2}$  to  $68 \text{ Wm}^{-2}$  at PAY and IZA, respectively. However, including our proposed model results improved performance, with reduced MBE and RMSE values. Specifically, compared to Mod 1, our model reduces errors at all sites. At BUD, MBE decreases from  $46 \text{ Wm}^{-2}$  to  $31 \text{ Wm}^{-2}$ , and RMSE decreases from  $57 \text{ Wm}^{-2}$  to  $48 \text{ Wm}^{-2}$ . At GOB, MBE decreases from  $149 \text{ Wm}^{-2}$  to  $132 \text{ Wm}^{-2}$ , and RMSE decreases from  $155 \text{ Wm}^{-2}$  to  $141 \text{ Wm}^{-2}$ . At IZA, MBE decreases from  $57 \text{ Wm}^{-2}$  to  $44 \text{ Wm}^{-2}$ , and RMSE decreases from  $68 \text{ Wm}^{-2}$  to  $62 \text{ Wm}^{-2}$ . At PAY, MBE decreases from  $18 \text{ Wm}^{-2}$  to  $4 \text{ Wm}^{-2}$ , and RMSE decreases from  $35 \text{ Wm}^{-2}$  to  $25 \text{ Wm}^{-2}$ . At TAT, MBE decreases from  $18 \text{ Wm}^{-2}$  to  $3 \text{ Wm}^{-2}$  and RMSE from  $47 \text{ Wm}^{-2}$  to  $36 \text{ Wm}^{-2}$ . Finally, at TOR, MBE decreases from  $30 \text{ Wm}^{-2}$  to  $16 \text{ Wm}^{-2}$ , and RMSE decreases from  $46 \text{ Wm}^{-2}$  to  $36 \text{ Wm}^{-2}$ . Thus, the model proposed performs well at every site except at BAR. It is worth noting that Mod 2, which includes  $G_1$  and  $k_t$  as input variables, performs poorly. This could be attributed to the fact that the variable  $k_t$  already incorporates  $G_1$ .

The analysis of the second type of models also reveals that simplicity yields the best results, with the exception of IZA, where Mod 5, which

**Table 9**  
Fitting parameters for each model.

Model	$a_i$	$b_i$	$c_i$	$d_i$
Mod 1	0,654	- 20,317		
Mod 2	0,867	- 81,483	6310	
Mod 3	0,721	0,777	- 301,420	296,842
Mod 4	0,781	- 13,596		
Mod 5	0,724	0,211	- 77,253	
Mod 6	0,863	- 90,491	87,219	

includes air temperature as an additional variable, performs best. Notably, models incorporating albedo present high  $R^2$  values at BAR, unlike the first type of models, and also demonstrate significantly reduced errors at GOB. Specifically, MBE values range from  $-4 \text{ Wm}^{-2}$  at BAR to  $63 \text{ Wm}^{-2}$  at IZA, with RMSE values from  $29 \text{ Wm}^{-2}$  at PAY to  $73 \text{ Wm}^{-2}$  at IZA. Despite minimal differences among models including the albedo, it is evident that adding more variables does not necessarily improve the results. Generally, except at BAR and GOB, the proposed model significantly outperforms all other analyzed models, demonstrating substantially reduced MBE values even compared to the best-performing model (Mod 4), which includes  $G_1$  and  $G_1$ . Our model consistently shows the lowest MBE and RMSE values when compared with Mod 4, the simplest model including albedo. Specifically, at BUD, the MBE value for our model is  $31 \text{ Wm}^{-2}$ , compared to  $35 \text{ Wm}^{-2}$  for Mod 4; at IZA, it is  $44 \text{ Wm}^{-2}$  versus  $63 \text{ Wm}^{-2}$ ; at PAY, it is  $4 \text{ Wm}^{-2}$  versus  $8 \text{ Wm}^{-2}$ , at TAT  $3 \text{ Wm}^{-2}$  versus  $5 \text{ Wm}^{-2}$  and at TOR,  $6 \text{ Wm}^{-2}$  versus  $30 \text{ Wm}^{-2}$ .

In summary, the comparison between the two types of models highlights an improvement when incorporating reflected shortwave radiation, though this variable is not commonly measured at most radiometric stations. Besides, the inclusion of additional variables does not necessarily enhance the model's performance. Notably, for the first type of models, all models perform poorly at both BAR and GOB, with MBE and RMSE values exceeding  $100 \text{ Wm}^{-2}$ . This analysis highlights two specific sites with distinct characteristics. The first one is BAR, where no model performs well, except those including  $G_1$ . This site exhibits highly variable albedo values ranging from 0,16 to 0,89, indicating the need of a model that incorporates reflected solar radiation. The second site is GOB, characterised by high values of albedo exceeding 0,30. GOB underscores the importance of including albedo information in the model. Therefore, the optimal model for both sites should include either  $G_1$  or albedo as input variable to improve accuracy. An additional evaluation was conducted at BAR, stratifying the analysis based on data for low and high albedo (Table 11). For data with low albedo ( $\leq 0,55$ ), all models perform very well, with  $R^2$  exceeding 0,90 for each model. Specifically, for the first type of models, MBE values range from  $-30 \text{ Wm}^{-2}$  (for the model here proposed) to  $27 \text{ Wm}^{-2}$  (for Mod 3), and RMSE values range from  $40 \text{ Wm}^{-2}$  (for Mod 1) to  $46 \text{ Wm}^{-2}$  (for our model). Similarly, for the second type of models, MBE values range between  $-18 \text{ Wm}^{-2}$  and  $-22 \text{ Wm}^{-2}$ , with RMSE values varying between 31 and  $39 \text{ Wm}^{-2}$ . These results indicate that the models are effective with data characterised by low albedo values. This poor performance is attributed to the challenges in estimating net radiation over surface characterised by sparse or no vegetation and high albedo. In such cases, the physically-based longwave radiation parameterization models or non-linear models should be considered (Jiang et al., 2015).

Among the models, Mod 6 presents the best results, though it differs only slightly from Mod 3, which includes only two input variables. The latter model requires knowledge of both  $G_1$  and albedo, adding to its complexity.

An additional analysis was conducted by calculating adjusted coefficients for our model specific to each site using one year of measurements (2021 for all sites except TOR). For TOR, the evaluation was performed using data for the year 2018. Table 12 shows the coefficients for each site, while Table 13 presents the results of this evaluation. As anticipated, the model shows improved performance with site-specific coefficients. However, substantial improvements are primarily observed at the two sites where models previously performed poorly: BAR and GOB. For example, for the models that do not include the albedo, at BAR for the model with  $G_1$  and  $k_t$  the RMSE is reduced from  $161 \text{ Wm}^{-2}$  to  $100 \text{ Wm}^{-2}$  and, at GOB, the same model shows a reduction from  $273 \text{ Wm}^{-2}$  to  $37 \text{ Wm}^{-2}$ . Despite these improvements, the proposed model, when using site-specific coefficients, shows RMSE values comparable to other models, ranging from  $25 \text{ Wm}^{-2}$  at PAY to  $38 \text{ Wm}^{-2}$  at TOR. This demonstrates the robust performance of the proposed model across different sites.

**Table 10**

Statistical results for the models (Mod1, Mod2, Mod3, Mod4, Mod5 and Mod6) using the dataset from each site. Includes mean bias error (MBE), root mean squared error (RMSE), coefficient of determination ( $R^2$ ).

Station	Model	$R_n$ $Wm^{-2}$	MBE $Wm^{-2}$	RMSE $Wm^{-2}$	$R^2$	Station	Model	$R_n$ $Wm^{-2}$	MBE $Wm^{-2}$	RMSE $Wm^{-2}$	$R^2$
<b>BAR</b>	Mod 1	93	66	127	0,234	<b>IZA</b>	Mod 1	348	57	68	0,960
	Mod 2		112	161	0,263		Mod 2		158	168	0,963
	Mod 3		107	154	0,274		Mod 3		128	135	0,947
	Mod 4		-4	32	0,937		Mod 4		63	73	0,963
	Mod 5		-10	33	0,944		Mod 5		50	66	0,954
	Mod 6		-17	30	0,959		Mod 6		62	73	0,961
<b>BUD</b>	Mod 1	251	46	57	0,956	<b>PAY</b>	Mod 1	326	18	35	0,974
	Mod 2		123	135	0,958		Mod 2		94	102	0,987
	Mod 3		112	121	0,944		Mod 3		-	-	-
	Mod 4		35	44	0,970		Mod 4		8	29	0,982
	Mod 5		36	49	0,959		Mod 5		-	-	-
	Mod 6		39	50	0,962		Mod 6		-	-	-
<b>GOB</b>	Mod 1	267	149	155	0,949	<b>TAT</b>	Mod 1	230	18	47	0,937
	Mod 2		259	273	0,950		Mod 2		92	105	0,953
	Mod 3		229	236	0,940		Mod 3		77	86	0,951
	Mod 4		53	59	0,973		Mod 4		5	34	0,965
	Mod 5		60	69	0,961		Mod 5		12	37	0,971
	Mod 6		60	67	0,965		Mod 6		17	29	0,983
<b>TOR</b>	Mod 1	156	30	46	0,943						
	Mod 2		85	98	0,959						
	Mod 3		-	-	-						
	Mod 4		16	30	0,967						
	Mod 5		-	-	-						
	Mod 6		-	-	-						

**Table 11**

Statistical results for the proposed model (Model) and for the existing models (Mod1, Mod2, Mod3, Mod4, Mod5 and Mod6) using the dataset from BAR, categorized by surface albedo. Includes mean bias error (MBE), root mean squared error (RMSE), coefficient of determination ( $R^2$ ).

$\alpha$	Model	$R_n$ $Wm^{-2}$	MBE $Wm^{-2}$	RMSE $Wm^{-2}$	$R^2$
$\leq 0.55$	Model	157	-30	46	0,900
	Mod 1		-18	40	0,901
	Mod 2		27	45	0,919
	Mod 3		25	44	0,901
	Mod 4		-18	33	0,944
	Mod 5		-22	39	0,938
	Mod 6		-19	31	0,964
$> 0.55$	Model	14	152	168	0,467
	Mod 1		167	185	0,441
	Mod 2		215	236	0,481
	Mod 3		204	221	0,485
	Mod 4		13	31	0,587
	Mod 5		6	26	0,669
	Mod 6		-15	26	0,719

**Table 12**

Fitting parameters for the proposed model for each site. (Model:  $A_i + B_i k_t \cos \theta_z + C_i (k_t \cos \theta_z)^2$ ).

Station	$A_i$	$B_i$	$C_i$
<b>BAR</b>	$45,7 \pm 0,6$	$-83 \pm 6$	$1466 \pm 11$
<b>BUD</b>	$-49,4 \pm 0,3$	$799 \pm 2$	$109 \pm 2$
<b>GOB</b>	$-70,6 \pm 0,3$	$560 \pm 2$	$246 \pm 2$
<b>IZA</b>	$-102,0 \pm 0,4$	$950 \pm 2$	$-54 \pm 2$
<b>PAY</b>	$-16,7 \pm 0,2$	$716 \pm 1$	$241 \pm 1$
<b>TAT</b>	$4,0 \pm 0,3$	$541 \pm 2$	$449 \pm 2$
<b>TOR</b>	$-19,9 \pm 0,2$	$590 \pm 2$	$390 \pm 3$

**5. Conclusions**

This study introduces a new, straightforward model for estimating instantaneous daytime net radiation ( $R_n$ ) under all sky conditions. The model is based on the premise that the diurnal variability of  $R_n$  mirrors the sinusoidal pattern of global solar radiation ( $G_I$ ), being the primary

**Table 13**

Statistical results for the proposed model (Model) with site-specific coefficients. Includes mean bias error (MBE), root mean squared error (RMSE), coefficient of determination ( $R^2$ ).

Station	$R_n$ $Wm^{-2}$	MBE $Wm^{-2}$	RMSE $Wm^{-2}$	$R^2$
<b>BAR</b>	93	18	100	0,267
<b>BUD</b>	251	7	36	0,952
<b>GOB</b>	267	4	37	0,943
<b>IZA</b>	348	9	42	0,948
<b>PAY</b>	369	4	25	0,981
<b>TAT</b>	230	3	26	0,954
<b>TOR</b>	156	7	38	0,947

contributor to the net energy balance. Building on previous works that developed models for different spectral ranges using only two parameters, i.e. the solar position (through  $\cos \theta_z$ ) and the clearness index ( $k_t$ ), both on a horizontal surface, this work proposes a novel model that leverages the relationship between  $R_n$  and the product  $k_t \cos \theta_z$ . The model reveals a quadratic relationship with a  $R^2$  of 0,981.

Six empirical models have been evaluated, categorized into two types based on whether they included reflected global irradiance ( $G_T$ ) in addition to downward global irradiance ( $G_I$ ) as an input variable. These models were validated at seven sites with diverse climate characteristics: Barrows in Alaska (USA), Gobabeb in the Namib desert, Izaña in the Canary Islands in Spain, Budapest in Hungary, Payerne in Switzerland, Tateno in Japan and Toravere in Estonia. Overall, incorporating  $G_T$  improved model performance, particularly at sites with high albedo values. Despite this, the simplest model, i.e. using only  $G_I$  as the input variable, consistently yielded the best results. The addition of more variables did not significantly enhance performance. In fact, a high correlation has been observed for all sites (except for BAR, which exhibits high values of surface albedo) between  $R_n$  and  $G_I$ , with correlation coefficients exceeding 0,90. Furthermore, the Spearman correlation analysis revealed that the most effective variables for estimating  $R_n$ , in order of importance, are the product  $k_t \cos \theta_z$  and  $G_I$ .

In general, the proposed model demonstrates strong performance, with  $R^2$  values exceeding 0,94, though it tends to overestimate in most cases. The Mean Bias Error (MBE) ranges from  $3 Wm^{-2}$  to  $44 Wm^{-2}$ , and

the Root Mean Squared Error (RMSE) varies from  $25 \text{ Wm}^{-2}$  to  $62 \text{ Wm}^{-2}$ . Notably, restricting albedo values below 0,55 yields good results at Barrows, comparable to those at other stations. Further analysis by season and sky conditions (based on  $k_t$  categories) reveals improved performance during cold seasons and under cloudy skies. This may be attributed to the reduced absolute magnitude of incident solar radiation under cloudy conditions. The statistical performance of the proposed model is comparable to that of more complex models using additional input variables. Additionally, an evaluation of models using site-specific coefficients showed improved results, with RMSE values being similar for both types of models (those including or excluding albedo as an input variable).

The primary advantage of the proposed model is its reliance on a single input variable: global solar radiation ( $G_{\downarrow}$ ). This variable is commonly available at most radiometric stations. However, the model's performance at sites with a wide range of albedo values may be limited, similar to other empirical models. In such cases, incorporating reflected global irradiance ( $G_{\uparrow}$ ) alongside  $G_{\downarrow}$  could improve accuracy. Unfortunately,  $G_{\uparrow}$  is not measured at many radiometric stations, which poses a challenge.

Regarding the global applicability of the empirical model proposed in this study, its limitations are noted, particularly for aquatic surfaces. The sites selected for this study are terrestrial stations, and, unlike land surfaces,  $R_n$  is not routinely measured over aquatic surfaces. It is also important to emphasize that simple models based on the relationship between  $R_n$  and  $G_{\downarrow}$  are inadequate for sites with high albedo. In such cases, physically-based longwave radiation parameterization models or nonlinear models should be considered. The selection of sites in this work was limited because not all variables included in  $R_n$  were available for a two-year period. Therefore, extending the database to include more sites with a broader range of altitudes, latitudes, and different surface types could enhance the global applicability of the model.

#### CRedit authorship contribution statement

**Inmaculada Foyo-Moreno:** Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Ismael L. Lozano:** Writing – review & editing, Supervision. **Inmaculada Alados:** Writing – review & editing, Supervision. **Juan Luis Guerrero-Rascado:** Writing – review & editing, Supervision.

#### 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.

#### Acknowledgements

This work was supported by the Spanish Ministry of Economy and Competitiveness through projects CGL2017-90884-REDT, PID2020-120015RB-I00, PID2020-117825GB-C21, PID2020-117825GB-C22 and PID2021-128008OB-I00, by the Andalusia Regional Government, University of Granada and FEDER funds through project C-EXP-363-UGR23. This research was partially supported by the Scientific Units of Excellence Program (grant no. UCE-PP2017-02). The authors thankfully acknowledge the FEDER program for the instrumentation used in this work. The authors would like to thank the BSRN network for providing the databases employed in this work.

#### Data availability

Data will be made available on request.

#### References

- Alados, I., Foyo-Moreno, I., Olmo, F.J., Alados-Arboledas, L., 2003. Relationship between net radiation and solar radiation for semi-arid shrub-land. *Agric. For. Meteorol.* 116, 221–227. [https://doi.org/10.1016/S0168-1923\(03\)00038-8](https://doi.org/10.1016/S0168-1923(03)00038-8).
- Bisht, G., Venturini, V., Islama, S., Jiang, L., 2005. Estimation of the net radiation using MODIS (Moderate Resolution Imaging Spectroradiometer) data for clear sky days. *Remote Sens. Environ.* 97, 52–67. <https://doi.org/10.1016/j.rse.2005.03.014>.
- Brutsaert, W.H., 1975. On a derivable formula for long-wave radiation from clear skies. *Water Resour. Res.* 11, 742–744. <https://doi.org/10.1029/WR011i005p00742>.
- Brutsaert, W.H., 1982. *Evaporation into the Atmosphere: Theory, History, and Applications*. Springer.
- Doorenbos, J., Pruitt, W.O., 1977. *Guidelines for predicting crop water requirements*. In: *Irrigation and Drainage Paper 24*. Food and Agriculture Organization of the United Nations, Rome, Italy.
- 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., König-Langlo, G., 2018. Baseline Surface Radiation Network (BSRN): Structure and data description (1992–2017). *Earth Syst. Sci. Data* 10 (3), 1491–1501. <https://doi.org/10.5194/essd-10-1491-2018>.
- Ferreira, T.R., Da Silva, B.B., De Moura, M.S.B., Verhoef, A., Nóbrega, R.L.B., 2020. The use of remote sensing for reliable estimation of net radiation and its components: a case study for contrasting land covers in an agricultural hotspot of the Brazilian semiarid region Thomás. *Agric. For. Meteorol.* 291, 108052. <https://doi.org/10.1016/j.agrformet.2020.108052>.
- Foyo-Moreno, I., Vida, J., Alados-Arboledas, L., 1999. A simple all weather model to estimate ultraviolet solar radiation (290–385 nm). *J. Appl. Meteorol. Climatol.* 38 (7), 1020–1026. [https://doi.org/10.1175/1520-0450\(1999\)038<1020:ASAWMT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1999)038<1020:ASAWMT>2.0.CO;2).
- Foyo-Moreno, I., Alados, I., Alados-Arboledas, L., 2007. Adaptation of an empirical model for erythemal ultraviolet irradiance. *Ann. Geophys.* 25 (7), 1499–1508. <https://doi.org/10.5194/angeo-25-1499-2007>.
- Foyo-Moreno, I., Alados, I., Alados-Arboledas, L., 2017. A new conventional regression model to estimate hourly photosynthetic photon flux density under all sky conditions. *Int. J. Climatol.* 37, 1067–1075. <https://doi.org/10.1002/joc.5063>.
- Foyo-Moreno, I., Lozano, I.L., Alados, I., Guerrero-Rascado, J.L., 2023. A new method to estimate aerosol radiative forcing on photosynthetically active radiation. *Atmos. Res.* 291, 106819. <https://doi.org/10.1016/j.atmosres.2023.106819>.
- Gueymard, Ch.A., 2018. Revised composite extraterrestrial spectrum based on recent solar irradiance observations. <https://doi.org/10.1016/j.solener.2018.04.067>.
- Hemmati, S., Gatmiri, B., Cui, Y.J., Vincent, M., 2012. Thermo-hydro-mechanical modelling of soil settlements induced by soil-vegetation-atmosphere interactions. *Eng. Geol.* 139–140, 1–16.
- Irmak, S., Irmak, A., Jones, J., Howell, T., Jacobs, J., Allen, R., Hoogenboom, G., 2003. Predicting daily net radiation using minimum climatological data. *J. Irrig. Drain. Eng.* 129 (4), 256–269.
- Iziomon, M.G., Mayer, H., Matzarakis, A., 2000. Empirical models for estimating net radiative flux: a case study for three mid-latitude sites with orographic variability. *Astrophys. Space Sci.* 273, 313–330. <https://doi.org/10.1023/A:1002787922933>.
- Jensen, M.E., Burman, R.D., Allen, R.G., 1990. *Evapotranspiration and Irrigation Water Requirements*. ASCE, New York.
- Jiang, B., Zhang, Y., Liang, S., Wohlfahrt, G., Arain, A., Cescatti, A., Georgiadis, T., Jia, K., Kiely, G., Lund, M., Montagani, L., Magliulo, V., Serrano-Ortiz, P., Oechel, W., Vaccari, F.P., Yao, Y., Zhang, X., 2015. Empirical estimation of daytime net radiation from shortwave radiation and ancillary information. *Agric. For. Meteorol.* 211–212, 23–26. <https://doi.org/10.1016/j.agrformet.2015.05.003>.
- Kjaersgaard, J., Cuenca, R., Plauborg, F., Hansen, S., 2007a. Long-term comparisons of net radiation calculation schemes. *Boundary-Layer Meteorol.* 123 (3), 417–431.
- Kjaersgaard, J.H., Plauborg, F.L., Hansen, S., 2007b. Comparison of models for calculating daytime long-wave irradiance using long term data set. *Agric. For. Meteorol.* 143, 49–63. <https://doi.org/10.1016/j.agrformet.2006.11.007>.
- Köppen, W., Geiger, R., 1936. *Das geographische System der Klimate*. Berlin.
- Lozano, I., Alados, I., Sánchez-Hernández, G., Guerrero-Rascado, J.L., Foyo-Moreno, I., 2023. Improving the estimation of the diffuse component of photosynthetically active radiation (PAR). *J. Geo. Res. Atmos.* 128, e2023JD039256. <https://doi.org/10.1029/2023JD039256>.
- Lu, J., Tang, R.L., Tang, H.J., Li, Z.L., 2013. Derivation of daily evaporative fraction based on temporal variations in surface temperature, air temperature, and net radiation. *Remote Sens.* 5 (10), 5369–5396.
- Lu, J., Tang, R.L., Tang, H.J., Li, Z.L., Zhou, G.Q., Shao, K., Bi, Y.Y., Labeled, J., 2014. Daily evaporative fraction parameterization scheme driven by day–night differences in surface parameters: improvement and validation. *Remote Sens.* 6 (5), 4369–4390.
- Lu, Y., Wang, L., Zhu, C., Zou, L., Zhang, M., Feng, L., 2023. Predicting surface solar radiation using a hybrid radiative Transfer–Machine learning model. *Renew. Sust. Energ. Rev.* 173, 113105. <https://doi.org/10.1016/j.rser.2022.113105>.
- Ohmura, A., Dutton, E.G., Forgan, B., Fröhlich, C., Gilgen, H., Hegner, H., Heimo, A., König-Langlo, G., McArthur, B., Müller, G., Philipona, R., Pinker, R., Whitlock, C.H., Dehne, K., Wild, M., 1998. Baseline surface radiation network (BSRN/WCRP): new precision radiometry for climate research. *Bull. Am. Meteorol. Soc.* 79 (10), 2115–2136. [https://doi.org/10.1175/1520-0477\(1998\)079<2115:BSRNBW>2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079<2115:BSRNBW>2.0.CO;2).
- Ortega-Farias, S., Antonioletti, R., Olioso, A., 2000. Net radiation model evaluation at an hourly time step for Mediterranean conditions. *Agronomie* 20 (2), 157–164.

- Peng, J., Jiang, B., Chen, H., Liang, S., Liang, H., Li, S., Han, J., Liu, Q., Cheng, J., Yao, Y., Jia, K., Zhang, X., 2021. A New Empirical Estimation Scheme for Daily Net Radiation at the Ocean Surface. *Remote Sens.* 13, 4170. <https://doi.org/10.3390/rs13204170>.
- Saito, H., Simunek, J., Mohanty, B.P., 2006. Numerical analysis of coupled water, vapor, and heat transport in the vadose zone. *Vadose Zone J.* 5 (2), 784–800.
- Sentelhas, P.C., Gillespie, T.J., 2008. Estimating hourly net radiation for leaf wetness duration using the Penman–Monteith equation. *Theor. Appl. Climatol.* 91 (1), 205–215.
- Swinbank, W.C., 1963. Longwave radiation from clear skies. *Q.J.R. Meteorol. Soc.* 89, 339–348.
- Wang, K.C., Liang, S.L., 2008. An improved method for estimating global evapotranspiration based on satellite determination of surface net radiation, vegetation index, temperature, and soil moisture. *J. Hydrometeorol.* 9 (4), 712–727.
- Wang, K., Liang, S., 2009. Estimation of daytime net radiation from shortwave radiation measurements and meteorological observations. *J. Appl. Meteorol. Climatol.* 48 (3), 634–643.
- Wang, Z., Zhang, M., Wang, L., Qin, W., 2022. A comprehensive research on the global all-sky surface solar radiation and its driving factors during 1980–2019. *Atmos. Res.* 265, 105870. <https://doi.org/10.1016/j.atmosres.2021.105870>.
- Wang, L., Lu, Y., Wang, Z., Li, H., Zhang, M., 2024. Hourly solar radiation estimation and uncertainty quantification using hybrid models. *Renew. Sust. Energ. Rev.* 202, 114727. <https://doi.org/10.1016/j.rser.2024.114727>.
- Weiss, A., 1982. An experimental study of net radiation, its components and prediction. *Agron. J.* 74(5):871e4.
- Wright, J.L., Jensen, M.E., 1972. Peak water requirements of crops in Southern Idaho. In: *Proceedings of the American Society of Civil Engineers. Journal of Irrigation and Drainage Engineering*, 96, pp. 193–201.