sid.inpe.br/mtc-m21c/2021/06.14.17.32-TDI

ESTIMATION OF SUNSHINE DURATION OVER BRAZIL BASED ON GEOSTATIONARY SATELLITE DATA: CPTEC/INPE MODEL VALIDATION AND IMPROVEMENTS

Maria Lívia Lins Mattos Gava

Master's Dissertation of the Graduate Course in Meteorology, guided by Drs. Simone Marilene Sievert da Costa Coelho, and Anthony Carlos Silva Porfirio, approved in May 24, 2021.

URL of the original document:
<http://urlib.net/8JMKD3MGP3W34R/44SAA3P>

INPE
São José dos Campos
2021

PUBLISHED BY:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

sid.inpe.br/mtc-m21c/2021/06.14.17.32-TDI

ESTIMATION OF SUNSHINE DURATION OVER BRAZIL BASED ON GEOSTATIONARY SATELLITE DATA: CPTEC/INPE MODEL VALIDATION AND IMPROVEMENTS

Maria Lívia Lins Mattos Gava

Master's Dissertation of the Graduate Course in Meteorology, guided by Drs. Simone Marilene Sievert da Costa Coelho, and Anthony Carlos Silva Porfirio, approved in May 24, 2021.

URL of the original document:
<http://urlib.net/8JMKD3MGP3W34R/44SAA3P>

INPE
São José dos Campos
2021

Cataloging in Publication Data

Gava, Maria Lívia Lins Mattos.

Ga24e Estimation of Sunshine Duration over Brazil based on geostationary satellite data: CPTEC/INPE model validation and improvements / Maria Lívia Lins Mattos Gava. – São José dos Campos : INPE, 2021.

xxiv + 100 p. ; (sid.inpe.br/mtc-m21c/2021/06.14.17.32-TDI)

Dissertation (Master in Meteorology) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2021.

Guiding : Drs. Simone Marilene Sievert da Costa Coelho, and Anthony Carlos Silva Porfirio.

1. Gridded SDU dataset. 2. Campbell-Stokes recorders. 3. Inter-comparison. 4. Satellitebased modeling. 5. GOES satellite. I.Title.

CDU 551.501.8

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Pós Graduação

Meteorologia

ATA DE REUNIÃO

DEFESA FINAL DE DISSERTAÇÃO: MARIA LÍVIA LINS MATTOS GAVA- REGISTRO 373782/2019

BANCA: 146/2021

No dia 24 de maio de 2021, ás 14h00 por videoconferência, a aluna mencionada acima defendeu seu trabalho final, intitulado **: "Estimation of Sunshine Duration over Brazil based on geostationary satellite data: CPTEC/INPE model validation and improvements."** (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. A aluna foi **APROVADA** pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Mestre em Meteorologia.

A banca sugere incorporar as observações dos membros na dissertação. O texto final deverá ser conferido pelos orientadores e ser entregue no prazo estipulado pelo regimento do curso.

Eu, Chou Sin Chan como Coordenadora do curso, assino esta ATA em nome de todos os membros, com o consentimento dos mesmos.

Membros da Banca

Dr. Juan Carlos Ceballos - **Presidente** - CPTEC/INPE

Dra. Simone Marilene Sievert da Costa - **Orientadora** - CPTEC/INPE

Dr. Anthony Carlos Silva Porfirio - **Orientador** - CPTEC/INPE

Dr. Rodrigo Santos Costa - **Membro da banca** - INPE/CGCT/DIIAV

Prof. Dr. Leonardo de Faria Peres - **Membro da banca externo** - UFRJ

Prof. Dr. Gustavo Bastos Lyra - **Membro da banca externo** - UFRRJ

Documento assinado eletronicamente por **Chou Sin Chan**, **Pesquisadora**, em 31/05/2021, às 11:31 (horário oficial de Brasília), com fundamento no art. 6º do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site http://sei.mctic.gov.br/verifica.html, informando o código verificador **7469452** e o código CRC **B96926FE**.

"Whether you think you can or whether you think you can't, you're right."

Henry Ford

I dedicate this dissertation to my mother, that has taught me by example how to be strong.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my spiritual guides, that always give me strength to pursuit my objectives.

The completion of this work could not been possible without the many suggestions and critical thoughts of Dr. Simone Sievert and Dr. Anthony Porfírio, to whom I am deeply thankful.

I appreciate the DSA's staff for the development of the satellite product that was analyzed in this work and for the support during this project.

I would like to acknowledge the recognized effort of the National Institute of Meteorology (INMET) and the Centre for Earth System Science at Brazilian Institute for Space Research (CCST/INPE) to provide ground measurements, a key component for many research applications, including this one.

I thank the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Finance Code 001, for the financial support.

And finally, my family and friends that kindly supported me through this adventure.

ABSTRACT

The main objective of this work was to validate the satellite-based sunshine duration (SDU) product developed at the Satellite and Environmental System Division on the Center for Weather Forecast and Climate Studies/National Institute for Space Research (DSA/CPTEC/INPE) for the period from from September 2013 to December 2017. In addition, it was sought to propose and assess potential improvements in the model in question. To achieve these objectives, this study had three steps. Firstly, data from Campbell-Stokes recorders and automatic sensors, provided by the National Institute of Meteorology (INMET) and the Brazilian Environmental Data Organization System (SONDA), respectively, were collected. Afterward, the quality analysis of these databases in the period of interest was performed. Given the results, the INMET data were selected to be the reference SDU data for the validation process. Later, the CPTEC daily SDU estimates were validated, through comparisons to the "ground truth" and the satellite-derived SDU estimates from the Satellite Application Facility on Climate Monitoring (CMSAF). To provide a condensed and more robust analysis, the stations were grouped by climate zones, and the statistical parameters were evaluated in terms of these regions for each month. Both products demonstrated an overall good performance. The best results were obtained for the stations south of 15°S, encompassed by the Tropical Central Brazil (Warm and Mesothermic) and Humid Temperate regions. The satellites products exhibited small overestimation tendencies for these locations (approximately 0.5h), whereupon the CPTEC product showed smaller bias for the former region and the CMSAF dataset for the latter. For the Equatorial region, overestimation was also found. Concerning this location the bias of the CPTEC product was on average 1 hour higher than the aforementioned regions, while the results for the CMSAF estimates lie close to the previous ones. The Northeast Brazil showed a seasonal underestimation tendency for the CPTEC dataset, that was mainly attributed to the cloudiness parameterization. For this region, the CMSAF presented high positive bias that might be related to misrepresentation of low-warm cloud fields. The root mean squared error (RMSE) and correlation coefficient (r) for these products varied within regions and evaluated month. The RMSE values lied between 1.01 and 2.55h and the r values were in general over 0.7, indicating strong positive correlation. These results are in agreement with previous reports on satellite-derived SDU estimates presented in literature. Lastly, some aspects of the CPTEC model were analyzed, aiming to provide meaningful insights on potential improvements: the treatment regarding the beginning and end of the day; the impact of the amount of available images on the estimation quality, and the importance of a proper clear sky reflectance assessment. It was suggested considerations on the sunrise and sunset to prevent estimates greater than the maximum day length; quality flags concerning the number of available images for the SDU estimate should be provided to the data user as ancillary data; and the implementation of a variable clear sky reflectance field along daytime, season and location to better represent this parameter. The results obtained showed that the CPTEC SDU estimates present a good performance over Brazil, being a reliable alternative to provide information for many applications.

Keywords: Gridded SDU dataset. Campbell-Stokes recorders. Inter-comparison. Satellite-based modeling. GOES satellite. Brazilian radiometric network.

ESTIMATIVA DA INSOLAÇÃO SOBRE O BRASIL ATRAVÉS DE DADOS DE SATÉLITE GEOESTACIONÁRIO: VALIDAÇÃO E APERFEIÇOAMENTO DO MODELO DO CPTEC/INPE

RESUMO

O presente trabalho teve como objetivo avaliar o desempenho das estimativas de insolação (n) por satélite desenvolvido pelo INPE sobre o Brasil. Adicionalmente, procurou-se propor e testar potencias melhorias no modelo em questão. O período analisado compreende de setembro de 2013 a dezembro de 2017. Para alcançar esses objetivos a pesquisa teve três etapas. Primeiramente, dados de heliógrados e sensores automáticos, fornecidos pelo Instituto Nacional de Meteorologia (INMET) e pelo Sistema de Organização Nacional de Dados Ambientais (SONDA), respectivamente, foram coletados. A qualidade dos dados no período de interesse foram investigadas. Após análise da qualidade dos dados, a base de dados do INMET foi selecionada para servir como dado de referência para as etapas de validação do presente trabalho. Na segunda parte, produto de n do CPTEC foi validado, através da comparação com dados de superfície e estimativas de n derivadas de dados de satélite da Instalação de Aplicação de Satélites no Monitoramento Climático (CMSAF). Para proporcionar uma análise condensada e robusta, as estações foram agrupadas em zonas climáticas, e os parâmetros estatísticos foram obtidos em termos desta regionalização para cada mês. Ambos os produtos demonstraram um bom desempenho no geral. Os melhores resultados foram obtidos para as estações ao sul de 15°S, compreendidas pelas regiões Tropical Brasil Central (Quente e Mesotérmico) e Temperado Úmido. Os produtos exibiram pequenas tendências de superestimativa para esses locais (aproximadamente 0.5h), enquanto o CPTEC apresentou menor bias para a primeira região e o CMSAF para a última. Para a região Equatorial, também foi encontrada tendência de superestimativa. O bias do produto do CPTEC foi em média 1 hora maior do que para as regiões anteriormente citadas, enquanto os resultados para as estimativas do CMSAF foram similares às regiões ao sul. O Nordeste do Brasil apresentou uma tendência de subestimativa sazonal para os dados do CPTEC, que foi atribuída principalmente à parametrização da nebulosidade. Para esta região, o CMSAF apresentou altos valores de bias positivo, que podem estar relacionados à representação inadequada de cobertura de nuvens quentes e baixas. A raiz do erro quadrático médio (REMQ) e o coeficiente de correlação (r) para esses produtos variaram de acordo com a região e mês avaliado. Os valores de REMQ variaram entre 1.01 e 2.55h e os valores de r estiverem em geral acima de 0.7, indicando forte correlação positiva. Os resultados obtidos estiveram em conformidade com os relatados anteriormente na literatura de estimativas de n por satélite. Na última etapa, alguns aspectos do modelo do CPTEC foram investigados: o tratamento no início e no final do dia; o impacto da quantidade de imagens disponíveis na qualidade da estimativa e a importância da avaliação adequada da refletância de céu claro. Foram sugeridas possíveis considerações acerca do nascer e pôr-do-sol para evitar estimativas maiores que a duração máxima do dia; o fornecimento de sinalizadores de qualidade relativos ao número de imagens disponíveis para a estimativa da n ao usuário; e a implementação de um campo variável de reflectância de céu claro ao longo do dia, estação e local para melhor representar este parâmetro. Os resultados obtidos mostraram que as estimativas de n do CPTEC apresentam bom desempenho para o Brasil, sendo uma alternativa confiável para fornecer informações para diversas aplicações.

Palavras-chave: Dados de insolação em grade. Heliógrafos. Inter-comparação. Modelagem baseada em dados satelitais. Satélite GOES. Rede radiométrica Brasileira.

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF ABBREVIATIONS

BSRN	–	Baseline Surface Radiation Network
CMSAF	–	Satellite Application Facility on Climate Monitoring
CPTEC	–	Center for Weather Forecast and Climate Studies
CS	–	Campbell-Stokes
DIIAV	–	Impacts, Adaptation and Vulnerabilities Division
DISSM	–	Satellite and Meteorological Sensors Divison
DNI	–	Direct normalized irradiance
DSA	–	Satellite and Environmental System Division
EUMETSAT	–	European Organization for the Exploitation of Meteorological Satellites
GCOS	–	Global Climate Observing System
GL	–	Global radiation 1.2 satellite-based model
GOES	–	Geostationary Operational Environmental Satellite
HYREX	–	Hidrology Radar Experiment
IAG	–	Institute of Astronomy, Geophysics and Atmospheric Sciences
IBGE	–	Brazilian Institute of Geography and Statistics
INMET	–	National Institute of Meteorology
INPE	–	National Institute for Space Research
KDE	–	Kernel density estimation
KED	–	Kriging with external drift
MAD	–	Mean absolute difference
MAE	–	Mean absolute error
MBE	–	Mean biased error
MD	–	Mean difference
METEOSAT	–	Meteorological Satellite
MODIS	–	Moderate Resolution Imaging Spectroradiometer
MSG	–	Meteosat Second Generation
MVIRI	–	Meteosat Visible and InfraRed Imager
RMSD	–	Root mean squared difference
RMSE	–	Root mean squared error
SARAH	–	Surface Radiation Data Set - Heliosat
SCE-UA	–	Shuffled Complex Evolution Algorithm
SDU	–	Sunshine duration
SEA	–	Sun elevation angle
SEVIRI	–	Spinning Enhanced Visible and InfraRed Imager
SID	–	Solar Incoming Direct radiation
SIS	–	Surface Incoming Shortwave radiation
SONDA	–	Brazilian Environmental Data Organization System
SZA	–	Solar zenith angle
WMO	–	World Meteorological Organization

CONTENTS

Page

1 INTRODUCTION

This study aims to explore the performance of the sunshine duration model ran operationally at the Satellite and Environmental System Division on the Center for Weather Forecast and Climate Studies/National Institute for Space Research (DSA/CPTEC/INPE). This chapter describes briefly the background, and the goals of this research, followed by an outline of the dissertation.

1.1 Background and motivation

Conceptually, sunshine duration (SDU) is the total of hours that sunlight reaches the Earth's surface directly from the sun. With the advance of the technology and the measurement instruments, it was formally defined as the sum of the periods in which direct solar irradiance reaches or exceeds 120 W/m^2 (WORLD METEOROLOGICAL ORGANIZATION - WMO, 2008). Along with precipitation and surface air temperature, it is one of the most important parameters in climate monitoring (KOTHE et al., 2013). In a given area, the amount of sunshine received is the major factor determining the local climate (BERTRAND et al., 2013).

The SDU importance has been known for a long time and its first measurements date back to the 19th century. In fact, there are time series as long as 100 years of SDU measurements accumulated at networks all over the world. SDU data is relevant for a number applications, such as yield planning in agriculture (RAO et al., 1998; XUE et al., 2011; HUANG et al., 2012; WANG et al., 2015), analysis of the thermal loads and sunshine duration on buildings (SHAO, 1990), input parameter in soil water balance models (WARNANT et al., 1994), and even research on human's health (MCGRATH et al., 2002; NASTOS; MATZARAKIS, 2006; KELLER et al., 2019). Akinoglu (2008) stated that SDU data is the best long term, trustworthy, and readily available measurement to estimate the surface solar irradiation, due to the linear relationship between these variables, described by Angstrom (1924). These data, achieved by means of Campbell-Stokes (CS) type recorders, the most common instrument used to measure SDU, are simple to collect. Although some issues, e.g. overburning due to intermittent sunshine, non-standard paper cards use, could reduce its reliability, which would interfere with the quality of solar irradiation estimates. Some of this aspects on SDU measurements are described on Chapter 3.

Based on that, it is clear the necessity of SDU records. However, there is a relatively small number of stations that measure it. Overall, networks of SDU are sparse and insufficient to cover large areas (KANDIRMAZ; KABA, 2014). In Brazil, the

National Institute of Meteorology (INMET) currently operates a network with 245 stations, besides that, it provides SDU time series from other 85 sites, that are no longer operational. The effort to maintain this network is essential to provide reliable SDU records for climate studies. Nonetheless, comparatively to the temperature or precipitation networks with thousand stations, the SDU current network seems inadequate to cover the Brazilian large extent territory.

Along with that, once that meteorological station measurements are point-based observations (ZHU et al., 2020), SDU in the vicinity has to be obtained through interpolation techiniques. Therefore the accuracy of the method strongly relies primarily upon the number and spatial distribution of meteorological stations. Generally, the station's distribution is heterogeneous, with most of them near cities, and for several reasons, extensive areas remain without records. For instance, it can be observed in Brazil, where some regions as the Northern have very few stations, while in the South and Southeast regions present a denser network. Consequently, the resulting interpolated field is usually poor for representing the temporal and spatial SDU variability characteristics (WU et al., 2016).

Geostationary satellites perform measurements with high spatial and temporal resolution and cover large areas, so their data can be used as an alternative to estimate SDU. Literature on the subject provides some proposed methods to accomplishes it. Given the fact that clouds are the primary responsible for SDU changes, several methods to derive SDU rely on it. Different techniques have been considered, such as SDU estimate based on cloud cover index (KANDIRMAZ, 2006; CEBALLOS; RODRIGUES, 2008; SHAMIM et al., 2012) and from satellite-derived cloud-type products (GOOD et al., 2010; WU et al., 2016; ZHU et al., 2020). Another approach is given by Kothe et al. (2017), in which SDU is calculated based on direct normal radiation threshold, using data from the Meteorological Satellite (Meteosat) series operated by European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). This approach has been considered one of the most advanced remote sensing tool to estimate the SDU, being operational at the Satellite Application Facility on Climate Monitoring (CMSAF).

Ceballos and Rodrigues (2008) proposed a SDU method based on Geostationary Operational Environmental Satellite (GOES) data. It is operational at the DSA/CPTEC/INPE. This method (hereafter CPTEC method) was validated by the authors using in situ measurements from São Paulo and Fortaleza, and their results indicated a good agreement between the estimates and the observed data.

Thereafter, Porfirio (2012) extended the validation of CPTEC method for the Northeast Brazil, using records from 53 stations for 2008, showing a good performance of the method. Furthermore, the author investigated the model's deficiencies and showed that by adjusting the parameter related to the clear sky reflectance for the region, better estimates were achieved. Notwithstanding, the climatic characteristics of that region are not representative of the whole country.

In addition, the satellite-derived SDU data set from CMSAF also provides estimates over Brazilian territory. However, the accuracy of the product was only evaluated for the monthly sums, and regarding few stations. Therefore, it is still needed a deep validation of daily SDU estimates based on satellite data over Brazil, in order to provide reliable SDU data.

1.2 Objectives

Given the great importance of good, trustworthy SDU data for many applications, the main purpose of this study was to validate and provide improvements to the CPTEC method for SDU estimate over Brazil for the period of September 2013 to December 2017.

To achieve that, the following specific objectives were proposed:

- Survey the available in situ SDU observations and propose a simplified quality control methodology for heliographic data;
- Evaluate the SDU datasets to select the most appropriate one to be used as reference in the validation process;
- Assess the quality of the CPTEC method for the Brazilian territory through the comparison of the estimates with the reference data set. Statistical analyses will be performed for daily SDU.
- Perform an intercomparison of the CPTEC and the CMSAF product (Kothe et al. (2017) method). The main aim here is twofold: first to verify if the CPTEC product is coherent with other satellite-derived estimates and second to seek the satellite-based estimates limitations and advantages.
- Investigate the considerations used in the CPTEC method, in order to provide suggestions on potential improvements.

1.3 Dissertation outline

This chapter described briefly the main topics associated to this dissertation, other chapters will give more details on the subject, mainly in terms of the SDU satellite-based methods and measurements.

The dissertation is organized as follows:

Chapter 2 provides a literature review focused on physical concepts of SDU, the instruments commonly employed for its records and previous remote sensing approaches for assessing it. At the end of the chapter, an overview of the subject over Brazil, regarding SDU's networks, and efforts to estimate and characterize it is performed.

Chapter 3 presents the readily available ground measurements sources for SDU studies, the employed quality control measures and the analysis on this data that led to the reference data set selection.

Chapter 4 introduces the satellite-derived SDU estimates algorithms evaluated in this research, presents the methodology used to the validation and inter-comparison, and displays the results obtained in this step.

Chapter 5 explores the deficiencies of the model and discusses potential improvements on the CPTEC method.

Chapter 6 gives a summary of the main results and conclusions, and recommendations for future work.

2 LITERATURE REVIEW

This chapter aims to describe some physical concepts of SDU and the main literature on observation and estimation methods. Concerning estimation methods it focuses mainly on those based on satellite. At the end of the chapter,it is presented an overview of the SDU's network, the efforts to provide reliable datasets, and to characterize this variable over Brazil.

2.1 Sunshine duration

2.1.1 Definition

SDU is defined as the total of hours in a day that the solar disc is not blocked so that direct normal radiation reaches the Earth's surface (WU et al., 2016). SDU has a maximum possible value that corresponds to the day length, which depends on solar declination angle and the latitude of the location, and can be computed as (DUFFIE; BECKMAN, 1991):

$$N = \frac{2}{15} arccos(-tan(\psi)\, tan(\delta)) \tag{2.1}$$

Where ψ is the site's latitude and δ is the solar declination (angles in degrees), given by

$$\delta = 23,45\, sin(\frac{360}{365}\,(284 + n)) \tag{2.2}$$

where n is the julian day.

The amount of sunshine hours depends on astronomical factors (i.e. sun's elevation and azimuth; the sun-Earth distance) and atmospheric constituents (i.e. the chemical abundance in the earth's atmosphere of gases and aerosols are responsible for solar radiation absorption and scattering). Clouds, consisting of liquid water droplets or ice particles, are the primarily responsible for preventing incoming solar rays from reaching the Earth's surface (KANDIRMAZ; KABA, 2014).

2.1.2 Records

There are different methods for measuring SDU, one of the most commonly used is the CS recorder. The instrument that was introduced in 1880, detects sunshine if the

solar radiation beam concentrated by a glass sphere burns a trace on a graduated paper card (HINSSEN; KNAP, 2007). This instrument is widespread all over the world, with time-series longer than 100 years. Despite that, it has some weaknesses as pointed out by Painter (1981) as followed: on days of strong and intermittent sunshine, the burn spreads on paper card, leading to an SDU overestimate. Different locations may require different amounts of energy to burn the paper, e.g. when conditions are cold and wet, more energy is required to burn a trace in the card than when it is warm and dry. Around the dawn period, dew may cover the sphere, reducing the intensity of sunlight passing through. These aspects have been known since it was earliest deployed, so that observers were instructed to take them into account when analyzing the card (KERR; TABONY, 2004). Even so, the flaws are not easily surpassed and have led to a non-negligible degree of subjectivity, affecting sunshine measurements.

Automatic sensors appear to be a more reliable approach to measure SDU. The World Meteorological Organization (WMO) replaced the CS recorders for the pyrheliometer, and redefined SDU as the number of hours for which the direct solar irradiance is reaches or exceeds 120 W/m^2, since investigations have shown that this is the mean threshold irradiance for burning the cards (WORLD METEOROLOGICAL ORGANIZATION - WMO, 2008). The pyrheliometer measures the direct beam solar irradiance, as the sunlight enters the instrument through a aperture and is directed onto a thermopile, converting heat to an electrical signal that can be recorded. The instrument is mounted on a solar track so it can follow the apparent sun position along its path through the sky (BLANC et al., 2014). Although this method is more accurate for measuring SDU, it is also more expensive and requires continuous maintenance, therefore there is a relatively low number of stations with this instrument. In Brazil, for instance, there are less than ten stations equipped with pyrheliometers measuring continuously.

Another automatic method is the pyranometric. It consists in the use of two pyranometers, one measuring global solar irradiance and another with a shadow-ring, to block the direct irradiance, measuring diffuse solar irradiance. From these measurements, the direct solar radiation is calculated. Thereafter, SDU is given by the number of hours that it is above the WMO threshold. There are yet other methods, such as the contrast and the scanning methods, however they are less employed. The SDU measurement networks, mainly those used in the present study are presented and analysed in Chapter 3.

2.1.3 Satellite-derived SDU estimates

SDU's networks are quite sparse and insufficient, due mostly, to geographic and financial reasons, especially on developing countries. Even that the best means of gathering any meteorological data is to record the data directly at meteorological stations, these are point measurements (ZHU et al., 2020), which imply that the recorded data have high temporal resolution but low spatial resolution (KANDIRMAZ; KABA, 2014). To obtain a regional SDU, interpolation methods are required. Due to the number of stations and its heterogeneous distribution, the resulting maps present relatively high uncertainty, especially in locations where there are no observations (GOOD et al., 2010), and are often deficient in representing the complex climate characteristics of geographic sunshine hours at a regional scale (WU et al., 2016). Alternatively, satellite measurements with higher spatial resolution and frequency appears as an useful tool to estimate SDU.

Using geostationary images from the visible band of the Meteosat, Kandirmaz (2006) proposed a simple model for estimation of daily global SDU based on the statistical relation between cloud coverage and SDU, that he found to be linear. To derive an equation that describes this relation, the author starts with the function between the cloud cover index and the total atmospheric transmission factor (i.e. the ratio of global radiation on the ground on a horizontal surface to the horizontal irradiance outside the atmosphere) that can be written as:

$$T_t = T_c + n_t(T_0 - T_c) \tag{2.3}$$

where TT_t, T_c and T_o corresponds to the total, clear and overcast sky transmission factors, respectively. It is convenient to propose the total atmospheric transmission factor as a linear function of cloud cover index, n_t, i.e.:

$$T_t = a - n_t b \tag{2.4}$$

where a and b give a measure of clear sky and cloudiness atmospheric transmissivity, respectively (DIABATÉ et al., 1989), that can be derived from regression equations. Daily transmission factor T_d can also be written as a linear function of daily mean cloud cover index, n_m, as (KANDIRMAZ et al., 2004):

$$T_d = a - n_m b \tag{2.5}$$

However, the relation between daily transmittance factor and daily fractional bright sunshine hours, n/N, can be written by using the well-known Angström–Prescott relation (ANGSTROM, 1924; PRESCOTT, 1940):

$$T_d = d - \frac{n}{N}\, c \tag{2.6}$$

wherein c and d are regression parameters which related to location, season, and state of the atmosphere. Combining the two latter equations, the following relation for s/S can be obtained:

$$\frac{n}{N} = e - n_m\, f \tag{2.7}$$

This implies that the daily mean cloud cover index, n_m, must be directly proportional to n/N. The author used the aforementioned equation to estimate SDU. Primarily, cloud cover index was calculated for each image. To do it, it was used the well-known relation proposed by Cano et al. (1986):

$$n_t = \frac{\rho_t - \rho_{ga}}{\rho_{cm} - \rho_{ga}} \tag{2.8}$$

where ρ_t, ρ_{ga} and ρ_{cm} are the albedo of the pixel in time t, albedo of free-cloud pixels and albedo of compact cloud cover on the same area, respectively. Than taking average of all cloud indices for a day, mean daily cloud cover index was obtained, which ranges from 0 to 1. If daily mean cloud cover index is equal to 0 this means that the pixel had no cloud cover all day long, and if it is equal to 1, than is totally covered by clouds all day long. Thereafter Equation 2.7 with coefficients fitted by linear regression was employed to calculate SDU over Turkey. To validation, mean biased error (MBE) and root mean square error (RMSE) values of monthly mean SDU for eight randomly chosen meteorological stations for each month were calculated, RMSE and MBE values vary between 0.54 h/d (for Diyarbakir station for August 1997) and 2.79 h/d (for Rize station for August 1997) and between -1.82 h/d (for Malatya station for October 1997) and 1.93 h/d (for Rize station for August 1997), respectively. The results suggest that the model performs better on

the stations that exhibits less cloudiness. In this study, the author had only 4 images per day with a resolution of 7 km x 7 km, so he suggests that with a finer resolution the errors would considerably decrease.

Good et al. (2010) developed a method to estimate SDU based on cloud classification data from the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) on board the Meteosat Second Generation (MSG) satellite, which is geostationary above 0° longitude, 0° latitude, provided by Satellite Applications Facility on Nowcasting. The cloud-type data consist of 21 classifications made using a multi-spectral threshold, and provided at full pixel resolution for each 15-minute observation (slot). A pixel can be assigned as bright sunshine or no sunshine according to the cloud type observed. When no cloud is detected, bright sunshine is assumed. Similarly, where opaque cloud is observed, the slot is referred as 'no sunshine'. In cases where semi-transparent or fractional (sub-pixel) cloud types are detected, it is possible that some bright sunshine may be recorded by the in-situ sensors. Therefore, to ensure that the satellite estimates are consistent with the station observations, there is a different approach to these situations: for pixels classified as fractional, a contribution of 0.5 for that slot is assumed, as opposed to a value of 1.0 for other bright sunshine assignments; for pixels classified as semi-transparent clouds, the sun elevation angle (SEA) is taken into account, because as the SEA decreases, the optical path of the sunshine through the cloud increases, which has the effect of attenuating more solar radiation than if the sun were directly overhead. Using the Fu-Liou radiation code, SEA thresholds were determined as 12.0^{o}, 13.8^{o} and 15.3^{o} for very thin, thin and thick cirrus, respectively. And where the SEA is above this threshold, bright sunshine is assumed. For SEA below this threshold, the observation is assigned 'no sunshine'. The estimation of SDU is calculated as the fraction of daytime observations slots that were assigned 'bright sunshine' per total number of available daytime observation slots, multiplied by the number of hours of daylight for that pixel. This calculation is made only when there is a least 90% of all available 15-minute daytime observation slots for a pixel. The validation of the estimates was assessed by comparing the daily pixel-based values with collocated daily station totals obtained over four months in 2008 over the United Kingdom. Since there are differences in observations from CS and Kipp & Zonen instruments, the evaluation was performed separately for each one of the instruments. The results showed that accounting for some bright sunshine for cirrus and fractional cloud types enhanced the accuracy of the estimates. Also, the results are notably better for the Kipp & Zonen compared with the traditional CS instruments. Nevertheless, overall agreement between the satellite and in-situ data is good, the high correlation of the satellite estimate with

the station observations suggests that the spatial pattern of SDU is captured by the satellite data.

Shamim et al. (2012) presented an improved model for global SDU estimation based on the one presented by Kandirmaz (2006). The study was focused on the Brue Catchment, in the southwest of England. They used hourly Meteosat geostationary images in visible band, with a spatial resolution of 2.5 km x 2.5 km, and station data from the Hidrology Radar Experiment (HYREX) project in year of 1994 for training and the year of 1995 for validation. The method suggested by this study distinguishes from the one developed by Kandirmaz (2006), mainly because its approach uses external snow cover information and accounts for the sun-satellite angle effects individually for each pixel in the calculation of the cloud index, also, it uses a finer resolution for the estimates. The validation of the proposed method was performed, and compared with the performance of Kandirmaz (2006) model. Results show an improvement over its predecessor, with the coefficient of determination (r^2) going up from 0.68 to 0.83, and RMSE and MBE coming down to 1.19 h and 0.081 h from 2.37 h and 0.21 h, respectively. The authors highlighted the fact that the method must be studied in other regions, because of the regional dependency of regression parameters.

Kothe et al. (2013) investigated two methods to derive daily and monthly SDU. They extended the Good et al. (2010) method to the wider region of Europe, and considered another one in that the SDU is computed using solar incoming direct radiation (SID) and the WMO threshold for sunshine of 120 W/m^2. Both, cloud-type (used in Good et al. (2010)) and SID datasets are provided by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Climate Monitoring Satellite Application Facility (CMSAF) and their retrievals are based on data from SEVIRI instrument on board Meteosat-8 and Meteosat-9. In the application of the cloud-type method, the authors used hourly data and made a modification in the approach of the semi-transparent clouds so it could cover a wider variety of cirrus clouds. So, in this study, the SEA thresholds to assign the pixel as bright sunshine were derived using hourly data and linear regression between the solar elevation angle and the normalized SID. The method based on SID product is very straightforward. The satellite-based SID product assumes a horizontal plane, and therefore to ensure consistency with in situ observations it was normalized by the solar zenith angle. SDU is then computed whenever the normalized SID exceeds 120 W/m^2. The two methods were compared to each other and to ground measurements focusing on the year 2008. The products intercomparison showed that they

present similar spatial patterns, but that the SDU-cloud-type product typically exhibit lower values than SDU-SID, especially over ocean areas. The comparison with ground measurements showed that for daily estimates, the cloud-type method exhibits an underestimation tendency while the SID method tends to overestimate. The results also indicate that the agreement between the SDU-SID and station data is more stable than for the one based on cloud-type, which is indicated by lower standard deviations and higher Pearson correlation coefficient. On the other hand, the SID method presents a higher bias (+0.8h) compared to the SDU-cloud-type data (-0.1h). The authors investigated the effects of zenith angle on the satellite-based SDU. They found a relationship between the satellite-based SDU and the zenith angle, as both products present negative bias with increasing zenith angle, with increased variance during the summer months. For the monthly SDU estimates, as for the daily evaluation, SDU-SID product performs better than SDU-cloud-type in terms of standard deviation and correlation, but has a bias of larger magnitude. As SDU-SID, performed better than SDU-cloud-type, it was selected to become an operational product within the Regional Climate Centre on Climate Monitoring.

Bertrand et al. (2013) attempted to estimate the SDU over Belgium by exploiting both in situ SDU measurements and data from MSG satellites. In their study, they proposed two approaches to combine these data: The first one consists in fitting regression models (in this case, a linear model) based on the satellite-derived parameters to bright sunshine hours recorded at the station level. They considered five regression models, in four of them, only one parameter from the Meteosat was used, among daily clearness index, daily beam ratio, daily effective albedo and SDU. The fifth one uses the daily clearness index and daily beam ratio combined. Thereafter the site-dependent regression coefficients are spatially interpolated to allow the empirical regression models to be applied to each pixel of the satellite data. The second approach was based on kriging with external drift (KED), that is capable of handle densely sampled ancillary variables highly correlated with the parameter of interest. In this study, KED was used with one of the four MSG-based parameters as a drift, and also with two drifts (i.e. clearness index and beam ratio). The validation was performed using 1 year (i.e. 2011) of daily in situ observation from the Belgian station network in a cross-validation framework. In general, the KED approach with one drift presented the best results, irrespective of the input parameters. The best performance is reached when the satellite-based daily global surface solar radiation estimations were corrected using in situ measurements before the computation of the MSG-based clearness index, and the aftermath is used as input for the KED. Furthermore, the authors showed that the use of this method enhanced the ability

of reproducing the local and regional SDU patterns over Belgium.

Kandirmaz and Kaba (2014) used images from Moderate Resolution Imaging Spectroradiometer (MODIS) on board the polar orbiting satellites Aqua and Terra to estimate daily global SDU in Turkey. The method used was the same described in his previous work (KANDIRMAZ; KABA, 2014), but with a better spatial resolution (250m). Besides the linear regression that was previously employed, the authors proposed a quadratic regression, once that in the mentioned study was concluded that the linear formulation was not adequate for the days of winter in which generally the sky was fully overcast. Data from nine stations were used to find a unique correlation for Turkey, and data from another sixteen stations were used for testing the success of the derived correlations. The analysis of the statistical indicators shows that for months with a high number of overcast days, the quadratic correlation indicates a better statistical relation between the satellite-derived data and the measured one than linear correlation. Whereas if this is not the case, both quadratic and linear models exhibit similar behavior. The overall RMSE and MBE values are consistent with those of previous satellite-based studies in the literature, which shows that although the MODIS data is obtained only twice in a day, the resulting accuracy is comparable with those obtained with geostationary satellite data, that offer many observations in a day. Therefore, daily SDU values can be estimated using images of a polar orbiting satellite for regions where there is no possibility of having any related data.

Wu et al. (2016) tried to obtain SDU by means of FY-2D geostationary meteorological satellite data. The method proposed is based on hourly cloud classification product, and the Shuffled Complex Evolution Algorithm (SCE-UA) was used to calibrate sunshine factors related to different coverage types according to ground measurement data from the Heihe River Basin in 2007. The estimated SDU was validated with ground observation data from 2008 and the spatial distribution was compared with the results of interpolation methods. With the exception of the stations in the mountains, the whole area of study presented a r^2 greater than 0.89, the index of agreement was higher than 0.990, suggesting a good performance. The results of the mean absolute error (MAE) and RMSE implied the presence of outliers. The authors also concluded that the satellite-based product was better able to capture the spatial distribution variation of SDU than the interpolated methods.

Kothe et al. (2017) combined the aforementioned SID product and the cloud index obtained from the Effective Cloud Albedo to estimate SDU. As in Kothe et al.

(2013), direct normalized irradiance (DNI) is derived from SID by normalization with the cosine of the solar zenith angle (SZA). Daily SDU is calculated as the ratio of satellite observations exceeding 120 W/m^2 (sunny slots) to all slots during daylight. The dataset is provided at 30 minutes resolution, and therefore each sunny slots corresponds to 30 minutes of sunlight, however there is a probability that not the whole 30 min are sunny, so the authors proposed weighting the slots using the information of the 24 surrounding grid points. The validation of daily SDU estimates over Europe (although not for all countries in it) showed good agreement, with small uncertainty and high correlation. The evaluation of monthly satellite-derived SDU was performed for all Meteosat disk. It was shown that there is a slight overestimate of SDU, with a bias of 7.5h, mean absolute difference (MAD) of 18.4h, and correlation of 0.91. Besides the high accuracy of SDU, no trend was identified on the normalized bias suggesting that the time series is very stable. Nevertheless, some deficiencies were identified, such as cases of a misinterpretation of snow coverage as clouds, issues with low clouds in West Africa, or a few artifacts in some scenes, that should be further investigated. The success of the results took this method to be implemented operationally in CMSAF, providing a dataset from 1983 to near real time. Bartoszek (2018) used this dataset to extend the validation for Poland, once that it has not been done in the above-mentioned study. The results obtained, endorsed the good performance of the estimates.

2.2 Sunshine duration over Brazil – an overview

The aforementioned worldwide lack of radiation data is also a reality in Brazil. The National Institute of Meteorology (INMET) currently provides both global radiation and SDU data. However, the first started to be measured only recently (for climatological purposes), with its longest time series beginning in 2000, and just for a small number of stations (five sites). Nowadays, INMET provides data of global radiation from 589 stations, nevertheless, continuous recording series are scarce. In order to reduce the absence of information regarding the solar resource, the Impacts, Adaptation and Vulnerabilities Division at Brazilian Institute for Space Research (DIIAV/INPE) began the Brazilian Environmental Data Organization System (SONDA) network, as a research program designed to deliver trustworthy meteorological and solarimetric data to support the Brazilian energy sector (SILVA et al., 2014). The network has 20 measurement stations distributed throughout Brazilian territory. Despite the efforts to provide reliable data, the distribution of these networks is uneven, with regions, such as the Midwest and Northern Brazil, having very few stations with series no longer than 15 years (XAVIER et al., 2016), which make

their use very restrict to agrometeorological studies (BATTISTI et al., 2019). As an alternative, SDU is widely used to estimate this variable, since it is highly correlated with solar radiation, and is registered in many climatological stations (ALMOROX et al., 2020).

Although its most commonly used as proxy for solar radiation (BURIOL et al., 2012; LYRA et al., 2016; SOUZA et al., 2016; RIBEIRO et al., 2018), the SDU has many other applications. It is essential for evapotranspiration estimates (MOURA et al., 2013) and as input parameter for crop simulations (BATTISTI, 2016), key components in agricultural yielding. For this reason, it is a fundamental aspect for a country where agriculture plays such important role as in Brazil (ANDERSON et al., 2016).

SDU was also correlated with several effects on human health, e.g., Benedito-Silva et al. (2007) studied the association between bright sunshine hours and suicide rates and Santos et al. (2020) evaluated the influence of the photoperiod exposure at birth and the chronotype. SDU is also a crucial aspect in the civil construction, as it is related to thermal comfort, hence it is usually used for thermal loads analysis on buildings (BASSO et al., 2015a). Its impact on tourism was also scope of research: Perch-Nielsen (2010) analysed the vulnerability of the beach tourism sector under climate change on a country level, as the attractiveness of a region for touristic activities depends strongly on the local weather and climate.

As previously stated, there is a greater amount of SDU records than other solar radiation data. Notwithstanding, its network presents a very irregular distribution of stations. Due to the continental dimension of Brazil and the difficult access to certain regions, the meteorological network does not have adequate density despite the recognised efforts of the responsible entities (FUNARI; TARIFA, 2017). To overcome this lack of data it is common to adopt techniques to perform spatial interpolations. In this context there are studies focused on both, regional and continental scales.

Funari (1983) aimed to characterize the spatial and temporal SDU distribution in Brazil. To accomplish that, the author used data for the period of 1931-1960 from 204 stations and produced monthly charts using a base map from the Brazilian Institute of Geography and Statistics (IBGE), with a scale of 1: 5.000.000, in azimuthal conformal projection. The relief, natural vegetation, hydrography, geomorphology, cloudiness and rain were considered to trace the isocontour maps of bright sunshine. The author pointed out some key regional features observed in his results. In the Northern region of Brazil the sunshine distribution was very similar to the rain chart, overall, the SDU values were lower in regions with greater amount of rain,

and occurred in the first semester of the year, when precipitation is higher. Over the Northeast region, the observed values were high, mainly due to the low cloudiness. In this region the relief plays a important role in the SDU spatial variability. A strong influence of the orography was observed in the South and Southeast regions; along with the frequent occurrence of frontal cloudiness, with the smallest SDU values observed in the region of Rio Doce Valley and the Serra do Mar.

Tiba et al. (2000) presented a great effort to gather and summarize the information regarding solar resources. Concerning the SDU, the data used are from CS recorders measurements collated from different sources. Since there was no previously knowledge on data quality, it was assumed that the errors did not surpassed 10%. In some cases, the SDU values were extrapolated till 200 km. The authors saliented the macro scale of the project: it was intended to express the relevant aspects of the global distribution of the resource, putting aside particularities typical of the small and regional scales (TIBA et al., 1999). The resultant Brazilian Solar Radiation Atlas consists of a volume on the existing solar radiation data in the country, in a standardized format, with 26 charts of daily solar radiation and sunshine hours, at monthly and annual basis. Besides that, the Atlas also assembled the relevant literature in the theme, with over 50 publications.

New et al. (2002) proposed a global high-resolution data set of monthly averages of relevant climate variables. The author gathered data from several sources. Regarding the measurements of sunshine, 4792 stations were used worldwide. The data were interpolated from networks of station observations using thin plate smoothing splines, with latitude, longitude and elevation as independent predictors. The resulting dataset has a spatial resolution of 10' lat/lon (approximately 18 km $\times$ 18 km at the equator). The authors noticed that large interpolations errors were found over the tropics due to the sparse station networks. They also pointed out that in regions with data scarcity, the stations used to create the datasets were unable to capture local features, and that the interpolated fields represented only of large-scales regimes.

Medeiros et al. (2018) used a Krigging method to evaluate the average SDU in Piaui state. The authors utilized data from 6 stations, comprising the period from 1962 to 2017. They highlight the absence of a fairly dense network to reasonable cover the state area. From their results, they were able to characterize the main features of SDU in Piaui. The annual SDU average observed was 2803.9 hours, with a minimum registered in February (156 hours) and a maximum in August (304.5 hours), that

were related to the cloud cover conditions.

Kozmhinsky et al. (2018) also employed the Krigging technique to characterize the climatic conditions of SDU in the State of Pernambuco. The period analyzed ranged from 1962 to 2016, and data from eight stations were used. Their results indicated lowest SDU values for the coastal and forest areas and high incidence of bright sunshine in the hinterlands and high sertão, mainly conditioned by low cloud cover.

Even though interpolations methods are usually employed to obtain the regional SDU, the accuracy of this techniques rely strongly on the number and spatial distribution of the available meteorological stations. Further, the unequal distribution of the stations, with locations with a large number of stations,e.g., near cities, and areas without any station, makes the resultant interpolation inadequate for representing complex characteristics at regional scale (WU et al., 2016).

With the development and advance of orbital remote sensing technology, this deficiency can be overcame since the satellites can provide observations per hour, or even shorter time intervals, over a wide spatial range with high resolution. Exploiting this advantage, Ceballos and Rodrigues (2008) proposed a simplified, but physically consistent, model to estimate the daily SDU (hereafter named as CPTEC method) based on the linear relationship between cloud cover and pixel reflectance in GOES visible channel. To validate their algorithm, the satellite-based estimates were compared to in situ measurements taken at the Institute of Astronomy, Geophysics and Atmospheric Sciences (IAG) in São Paulo and at Univesity of Ceara in Fortaleza, for 2007. The results indicate that the estimates were accurate within 10% of systematic deviation.

Extending the validation for the Northeastern Brazil, Porfirio (2012) evaluated the CPTEC method considering 53 stations fairly distributed over Piaui, Ceará, Rio Grande do Norte, Paraíba, Pernambuco and Alagoas states, equipped with CS recorders, for 2008. The scatterplots results showed that the satellite-derived estimates underestimate the ground records for clear sky conditions (over 8 hours) and tend to overestimate for overcast sky (less than 2 hours). The spatial distribution of the r^2 and the mean bias error showed an overall good agreement between the estimated and the observed SDU, with values lying within 0.64 and 0.91 for the r^2, and -1.92 and 1.03 for the bias. From the total analysed stations, only 3 of them presented r^2 smaller than 0.75. Furthermore, the author evaluated the cloudiness estimate by changing the surface albedo from 0.09 (as stated in the method) to 0.05 (found using the mean annual value of the visible channel obtained by the MODIS

sensor) for the station located in Natal-RN, which provided a significant improvement to the SDU estimate, showing the importance of a good representativeness of the clear sky reflectance for accurate estimates.

The model of Kothe et al. (2017) was developed for Europe and Africa, nonetheless, the product provides SDU estimates for the entire Meteosat disk. Over the Southern America, only the monthly sums were validated, using less than 40 stations, for the period from 1983 to 2015. The found bias lies between -18 and 24 hours, with a site dependent performance.

The Brazilian territory has an area of about $8,4 \times 10^6\, km^2$, with extensive regions with low-density population and difficult access (CEBALLOS et al., 2004). Despite the efforts of the responsible entities, the current network is insufficient to cover the entire territory given its size, along with the fact that most of the stations are concentrated in urban regions, it is not enough to comprise the whole country with accuracy. The interpolation methods, although relevant to provide information in areas with few records, proved to be inadequate to supply information on regional scales. The employment of satellite-derived data seems to be the way to overpass this issue. The literature about SDU estimates by satellite means over Brazil is still small. The mentioned models that provide SDU estimates over the country were under validated, with extensive information on the performance of the daily SDU estimates only for the Northeast Brazil, for the CPTEC model, and just for a few stations, regarding the monthly sums for the CMSAF method. Therefore, there is still a need to extended this validation for the Brazilian territory.

3 GROUND STATION DATA

A crucial component in the validation process is the consistency check (i.e. comparison) of the remote sensing product with reference measurements which are presumed to be representative of the truth, within their own uncertainties. The choice of the reference data is often based on practicality or on what is considered to be most suitable for the validation procedure (LOEW et al., 2017). Nevertheless, this selection usually regards the following inquiries:

a) Do the database provide scientifically meaningful estimates of the investigated variable?

b) Do these data adequately cover the parameter field?

c) Are the data accurate enough to be able to draw insightful conclusions from the validation process?

d) Are the data publicly available and accessible?

This chapter intends to address these questions aiming to select the appropriate dataset to progress to validation procedure.

3.1 Campbell-Stokes records

The traditional source of SDU data is the CS sunshine recorder. It consists of a spherical glass lens that focuses the Sun's rays onto a treated card placed in a metal holder at the base of the recorder, that burns when direct radiation's intensity overcomes the burning threshold of the recording surface (WOOD; HARRISON, 2011). The WMO specified this threshold as 120 W/m^2 (WORLD METEOROLOGICAL ORGANIZATION - WMO, 2008). When the WMO recommended procedures are adhered to, the SDU measurements resolution and uncertainty is 0.1 hour (STANHILL, 2003).

This instrument is widespread all over the world, with time-series longer than 100 years, and it is the most viable and feasible instruments for measuring sunshine hours in developing countries (ALMOROX et al., 2020). Despite that, it has some known flaws:

- On days of strong and intermittent sunshine, the burn spreads on the paper card, leading to an SDU overestimation (LEGG, 2014);

- Measurements have a relatively high subjectivity degree, since they rely upon the observer's reading (KERR; TABONY, 2004);

- The energy necessary to produce an observable burn in the paper band may vary from the threshold defined by WMO (i.e. 120 W/m^2). Literature presents values ranging from 70 to 285 W/m^2 (BAUMGARTNER, 1979; PAINTER, 1981), because the paper strip responds in different ways according to the atmospheric conditions such as humidity and temperature;

- Eventually, a value of zero is returned when it is almost certain that there was sunshine, as it has been determined that missing data were often reported as zero. A missing data can arise for several reasons, e.g., glass sphere has been removed, inappropriate card to the time of year or instrument not set to the correct latitude so that the sun is not focused on the card, among others (ABBOT, 1986).

In Brazil, the INMET maintains a network with approximately 330 stations measuring SDU with CS recorders (INSTITUTO NACIONAL DE METEOROLOGIA, INMET,). However during the present work's period of interest (September 2013 - December 2017), only 293 stations have data records available. To keep the long-term consistency, only stations with less than 20% of missing observations over the study period were selected. Because of gaps in the heliographic time-series, just 194 stations will proceed for further data quality analysis. These stations are shown in Figure 3.1, and their information is listed in Table A.1 in the appendix. INMET's data are publicly available at `https://bdmep.inmet.gov.br/`.

3.2 INMET's stations - quality control

Unfortunately, the INMET's network does not adopt any measurement data quality control, to ensure the reliability of the records, some quality checks are required. Currently, there is no well established method for SDU quality control, but several attempts are reported in literature. The following two steps are commonly used to SDU data (FENG et al., 2004; REEK et al., 1992; SANCHEZ-LORENZO et al., 2007), and are adapted here:

(i) Extreme possible values: the low extreme was set as zero and the high extreme is obtained using the day length equation, which corresponds to the maximum possible duration of a day (Equation 2.1).

Figure 3.1 - INMET (dots) and SONDA (crosses) stations distribution.

(ii) 'Flatline' check: identifies data intervals with the same value for several consecutive days.

The steps were applied in the presented order. Those records that exceeds the extreme possible values or were identified by the 'flatline' check, were tagged as suspicious and excluded from the following stages. The suitability of the procedure was accessed through visual inspection of the time-series for each individual station data plotted along with the satellite-derived data.

3.2.1 Quality control's conformity analysis

It was not intended in this work to develop a quality control methodology for the heliographic data. Nevertheless, the lack of any quality indicator forced an effort to objective analyse the records with at least minimum requirements to exclude from the following procedures gross erroneous data and assure a fairly reliable reference data.

The first criteria (i) regarded unphysical values, i.e. data with negative values and

data higher than the maximum possible duration of the day, given by Equation 2.1, and very few data were discarded in this step.

The second criteria (ii) was the 'flatline' check: as suggested by Feng et al. (2004), the occurrence of the same variable value for several days is very unlikely to happen, therefore it can be considered suspicious.

This approach, however, present different time intervals for the 'flatline'. In this study 5, 7, and 10 days were tested. The time-series of the data with the correspondent flags were plotted along with the satellite-derived SDU data to verify the conformity of the criteria (not shown).

To clarify the selection of the day interval selected, Figures 3.2 and 3.3 presented. For the southern stations, the use of 5 days for the 'flatline' check misclassified some valid data (that show conformity with the satellite data, i.e. with small differences between the satellite-derived and the observed measures) as can be seen in Figure 3.2a, for June, September and October. This probably occurs due the high frequency of frontal systems in that region, being usually responsible for overcast sky and persist for a few days. The use of 7 days criteria minimize these glitches. Extending the period for 10 days did not show improvements, and even presented cases where the 7 days criteria classified as suspicious correctly, and the 10 days did not, as can be observed in April 2016 (Figure 3.2c).

In stations far northern, the 5 days 'flatline' criteria exhibits some false alarms for events of clear sky, as suggested by satellite data (Figure 3.3a). It is observed in regions where the day length does not present great changes over the year and seasons that clear sky is predominant over day for several days (e.g. semiarid), considering the instrument resolution (6 minutes) it can result in the same SDU value. The use of 7 days criteria eliminated this misclassifications (Figure 3.3b).

The analysis indicated that the 7 days interval presented the best overall performance between the intervals tested. Since the quality control steps proposed aimed to achieve an objective procedure, the 7 days interval displayed a better capability to be employed in different climate conditions, therefore it was selected. Further analysis will use only data that were not tagged as suspicious by the (i) extreme possible value and the (ii) 7 days 'flatline' check.

Although, the application of these 2 criteria enhanced the confiability of the resulting time series, it was noticed that some observations still present some question-

Figure 3.2 - Uruguaiana's station (RS, 83927) time-series: example of the periods flagged with the 'flatline' check method. (a) Top, (b) middle and (c) bottom plot presents, respectively, 5, 7 and 10 days intervals. Heliographic data were plotted as dots, CMSAF's SDU data as circles, and CPTEC's SDU data as crosses. The colors represents the correspondent flag: blue is for "approved" data, green for "suspicious" data. The grey line represent the maximum possible SDU value.

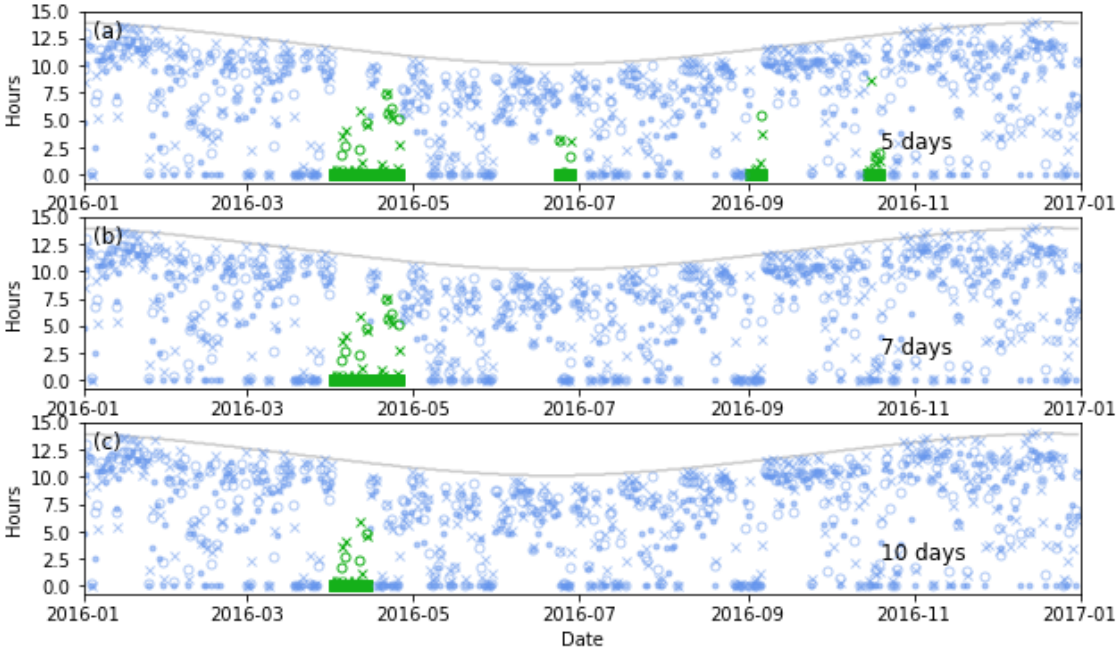
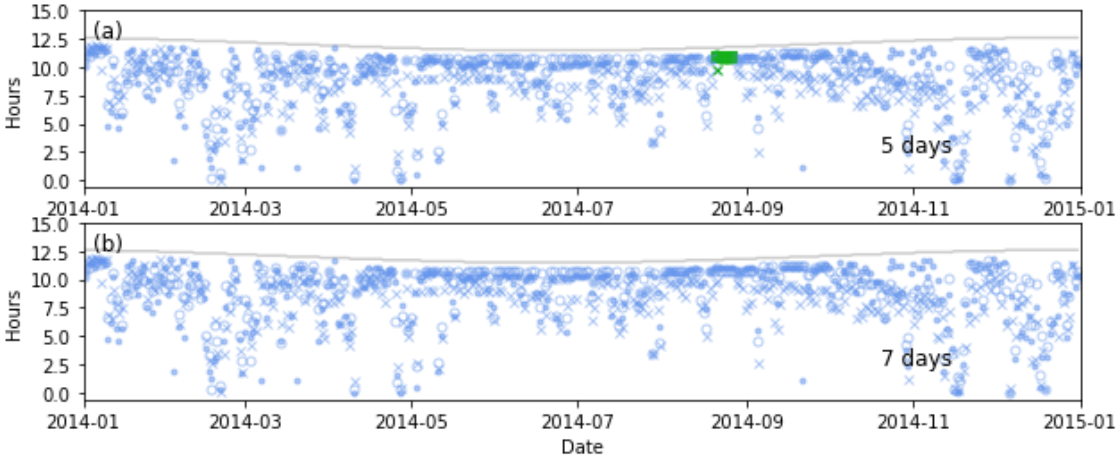


Figure 3.3 - Remanso's station (BA, 82979) time-series: example of the periods flagged with the 'flatline' check method. (a) Top and (b) bottom plot presents, respectively, 5 and 7 days intervals. Heliographic data were plotted as dots, CMSAF's SDU data as circles, and CPTEC's SDU data as crosses. The colors represents the correspondent flag: blue is for "approved" data, green for "suspicious" data. The grey line represent the maximum possible SDU value.



able behavior (e.g. zero values recorded in station observation, while the satellite products report over 9 hours of sunshine). This indicates the need for even deeper investigations to develop a robust quality control methodology, mainly for heliographic records. Some aspects of the ground observations were investigated and are presented in Section 3.4.

3.3 Pyrheliometer and pyranometer data

The SDU can also be registered by means of pyrheliometers. These instruments measure the DNI as they track the apparent sun position along its path through the sky. Using the threshold established by WMO the SDU can be calculated as the sum of observations that DNI $\geq 120 \mathrm{W/m^2}$.

Alternatively, pyranometer can also be employed to obtain DNI, coupling two instruments: one measuring global solar irradiance and another with a shadow-ring to block the direct irradiance, then measuring diffuse solar irradiance. From these measurements, the DNI is estimated using the formula:

$$DNI = \frac{G - D}{cos(SZA)} \tag{3.1}$$

Where G is the global solar irradiance, D is the diffuse solar irradiance and SZA is the solar zenith angle.

There are insufficient stations that use automatic instruments in Brazil to cover its large area, mostly due its costs. Unfortunately, not all stations that record this data make it freely available. The few stations that report data in the period of interest are shown in Figure 3.1, as crosses. Their information is listed in Table A.2 in appendix, along with the available instrument (Pyranomenter or Pyrheliometer). These stations are part of the SONDA project, that includes stations of their own and those that belong to partners. More information about this network can be found at `http://sonda.ccst.inpe.br/index.html`.

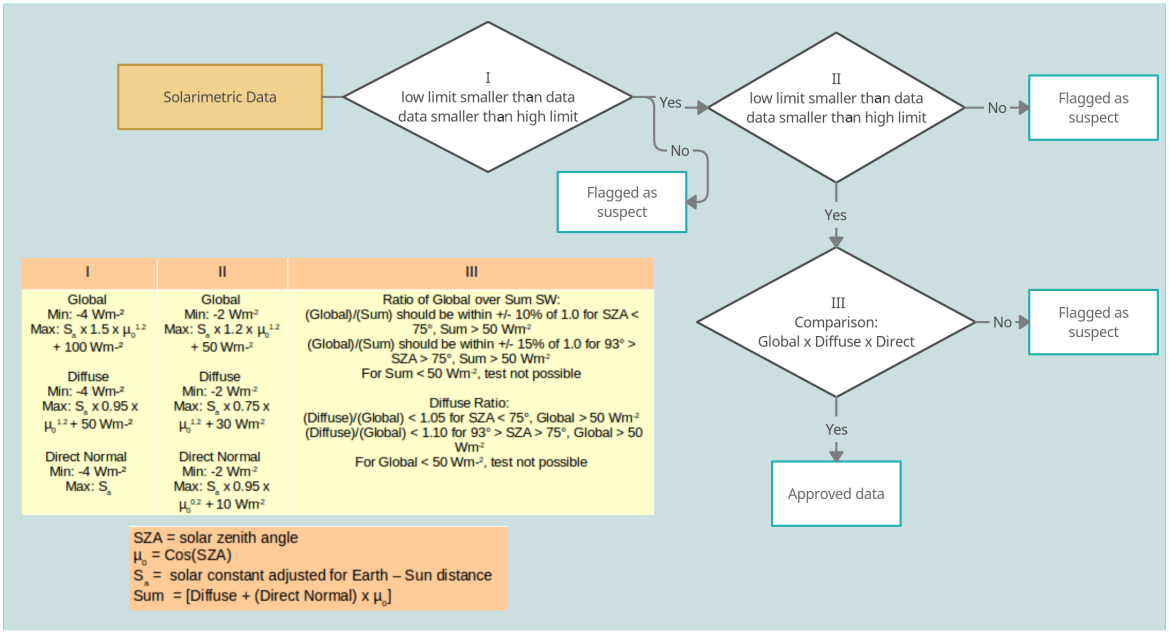
SONDA project adopts a validation process through their stations, which is based on the data quality control strategy employed by Baseline Surface Radiation Network (BSRN) and is provided as ancillary data. This consists in a three steps algorithm that verifies:

- Physically possible limits;

- Extremely rare limits;
- Instrument comparisons.

The criteria adopted are shown in Figure 3.4.

Figure 3.4 - SONDA's data quality control process.



Similarly to the stations from INMET, the selection of SONDA data was based on availability: the present study used only those stations that presented more than 80% of available data not flagged as suspect (during daytime) in the selected period.

3.3.1 Sunshine duration calculation

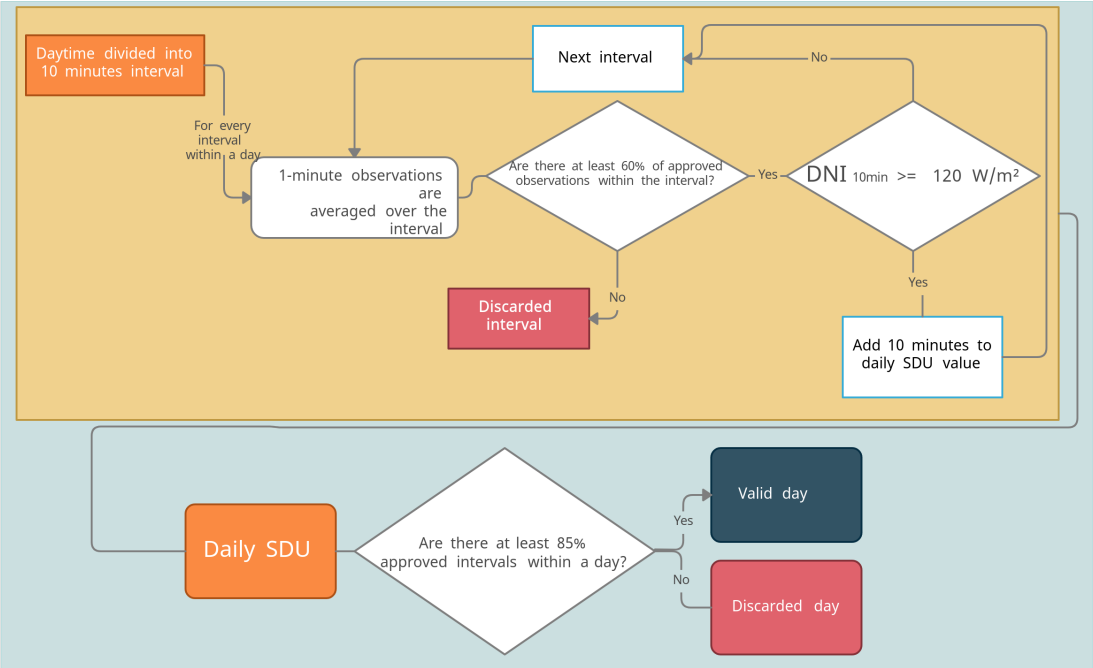
Daily SDU values are calculated by summing up the time intervals that DNI observations exceed the 120 W/m^2 within a day. Nonetheless, to overcome small gaps in the daily data, the 1-minute observation was averaged out over 10-minutes intervals. The intervals mean, however, was calculated only if at least 60% of the interval's observations have been approved by the quality control process, and the correspondent interval was set as approved. Afterwards, for every interval in a day that $\overline{DNI}_{10min} \geq 120\text{W/m}^2$, 10 minutes were summed up to the daily total. Daily SDU values were

considered valid, if 85% of the 10 minutes intervals within the day were approved. This approach was carried out to guarantee consistency within the intervals, and to exclude days with gaps longer than 2 hours.

For the stations that had pyranometric data, instead of pyrheliometric, the global and diffuse data were used to calculate DNI by means of the Equation 3.1. Subsequently, the procedure followed the method explained above.

Figure 3.5 presents the above process's flowchart.

Figure 3.5 - Flowchart of SDU calculation from DNI's observation.



3.4 Ground measurements comparison

One of the major concerns related to the worldwide crescent employment of automatic instruments as replacement for the CS recorder regards the sensitivity threshold. Diverse sets of thresholds for the heliograph can be found in literature, ranging from 70 to 285 W/m^2 (BAUMGARTNER, 1979; PAINTER, 1981), the later much larger than the established 120 W/m^2. This leads to an increasing attention to comparison of sunshine records obtained with different observing system devices.

To investigate the consonance of the WMO 120 W/m^2 criteria and the CS sensitivity threshold in Brazil, it was selected stations that are nearby. However, as highlighted by Baumgartner et al. (2018), parallel records of automated and traditional SDU recording systems are rare. Simultaneous CS and automatic measurements of SDU were available for only 6 sites. The locations and distances between stations are displayed in Table 3.1. The ideal would be that measurements were made at same location; but unfortunately such data was not available.

Table 3.1 - Near stations information. The SONDA's stations that use the pyrheliometer to acquire the data is marked with asterisks.

Site	Distance	INMET's stations			SONDA's stations		
		No	Latitude	Longitude	ID	Latitude	Longitude
Brasília	30.22 km	83377	-15.78	-47.93	10*	-15.6008	-47.7131
Florianópolis	3.82 km	83897	-27.58	-48.56	01*	-27.6017	-48.5178
Natal	7.81 km	82598	-5.91	-35.20	17	-5.8397	-35.2064
Palmas	6.92 km	83033	-10.19	-48.30	19	-10.1778	-48.3619
Petrolina	38.36 km	82983	-9.38	-40.48	11*	-9.0689	-40.3157
São Luis	7.00 km	82280	-2.53	-44.21	16	-2.5933	-44.2122

For these stations, scatterplots were made along with the calculation of the mean difference (MD), the root mean squared difference (RMSD), the correlation coefficient (r) by means of the following equations:

$$MD = \frac{1}{n}\sum_{i=1}^{n}(SDU_{A_i} - SDU_{CS_i}) \tag{3.2}$$

$$RMSD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(SDU_{A_i} - SDU_{CS_i})^2} \tag{3.3}$$

Where SDU_A corresponds to the automatic measurements, i.e. pyrheliometric or pyranometric data, and SDU_{CS} represents the heliographic data.

$$r = \frac{\sum_{i=1}^{n}(SDU_{A_i} - \overline{SDU_A})(SDU_{CS_i} - \overline{SDU_{CS}})}{\sqrt{\sum_{i=1}^{n}(SDU_{A_i} - \overline{SDU_A})^2}\sqrt{(SDU_{CS_i} - \overline{SDU_{CS}})^2}} \tag{3.4}$$

Figure 3.6 presents the scatterplots of daily sunshine amount of CS recorder against

automatic sensors and Table 3.2 displays the statistics results.

Figure 3.6 - Scatterplot of daily sunshine totals obtained from Campbell–Stokes (CS) against automatic (A) recorders.

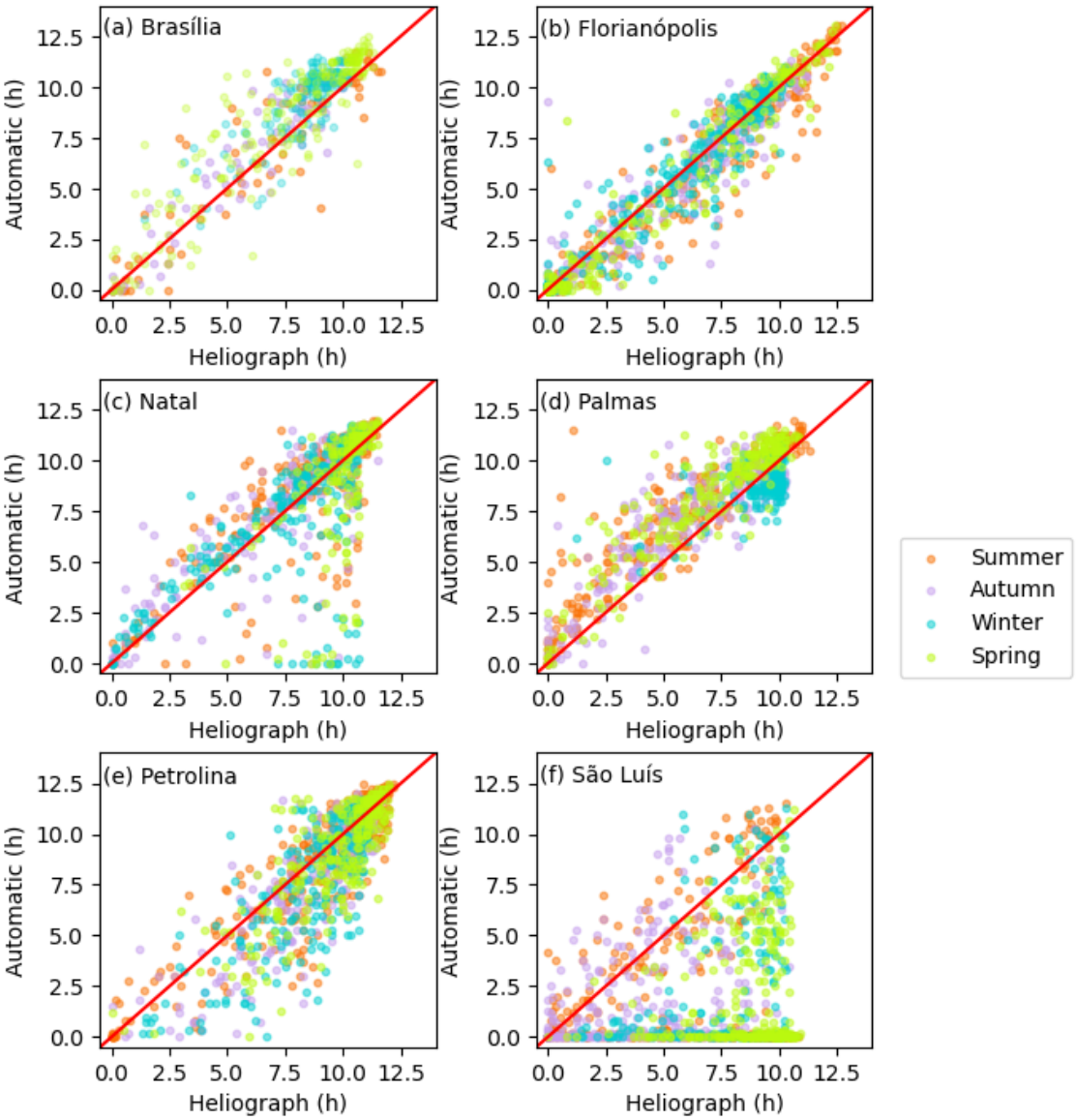


Table 3.2 - Statistics of the comparison of daily SDU ground measurements. Locations where the SONDA's stations use the pyrheliometer to acquire the data is marked with asterisks.

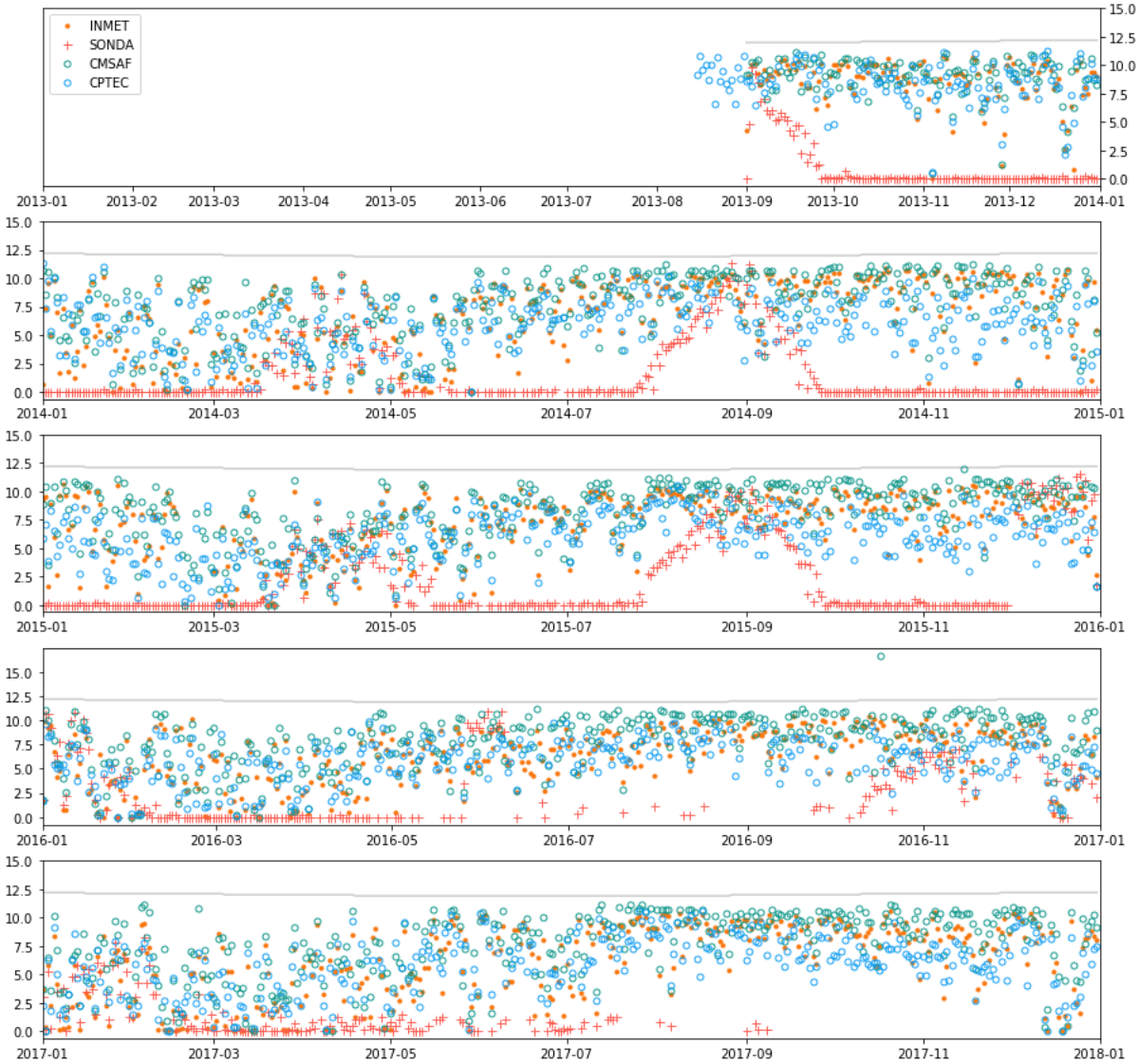
Site	MD (h)	RMSD (h)	r
Brasília - DF *	0.81	1.55	0.92
Florianópolis - SC *	-0.14	1.11	0.96
Natal - RN	-0.12	2.23	0.76
Palmas - TO	0.77	1.67	0.88
Petrolina - PE *	-0.55	1.71	0.86
São Luís - MA	-4.35	5.81	0.23

The Natal and São Luis (Figure 3.6c,d) plots display a very distinctive behavior, with CS records presenting over 10 hours of sunshine, while the automatic system indicates zero. This leads to a deeper investigation of the data from these stations. The time-series of São Luis's station records are shown in Figure 3.7. As can be seen, the automatic data presents systematic zeros, with few exceptions as those in September 2013, 2014 and 2015. The time-series plot of Natal station (not shown) exhibits less erratic behavior, but also presented systematic zeros for extended periods. The SONDA's network do not supply metadata of the stations, so the error can not be assuredly traced back to its origin. Notwithstanding, it can be noticed that, for these sites, the SDU was derived from the DNI estimated by means of pyranometers. It is possible that the inadequate positioning of the shadow ring, at these stations, provide erroneous diffuse irradiance data, that would not be tagged as suspicious, since it would not trigger any of the criteria used for quality control. The DNI estimated from this data would be inaccurate, consequently, the SDU would also be faulty. Given the above, these stations were excluded from further analysis.

Petrolina's station also exhibits suspicious behavior. During winter, for instance, this location presents predominantly clear sky conditions, considering that it corresponds to the dry season at this site. However, the automatic data frequently reported low values of SDU, e.g. less than 2 hours, while the CS measurements displayed up to 8 hours, in some cases (Figure 3.6e). Given the expected SDU behavior for this season (high values due to the lack of cloudiness), it was concluded that the automatic data of this station is questionable. High differences between the SDU measurements from the automatic and the CS recorder are also observed for other seasons. Unfortunately, as exposed above, SONDA project do not provide metadata on the of the station, therefore the error can not be traced back to its source. Further analysis may consider it.

For the other stations, the records are overall, in agreement. Depending on the location some SDU over or underestimate can be observed. As previously mentioned, the threshold sensitivity is a function mainly of temperature and humidity (PAINTER, 1981): when conditions are cold and damp, more energy is required to burn a trace in the card, than when it is warm and dry.

Figure 3.7 - Time series of daily sunshine totals obtained from CS (orange dots), automatic recorders (red crosses) and satellite-derived estimates (circles) for São Luis - MA. The grey line represent the maximum possible SDU value.

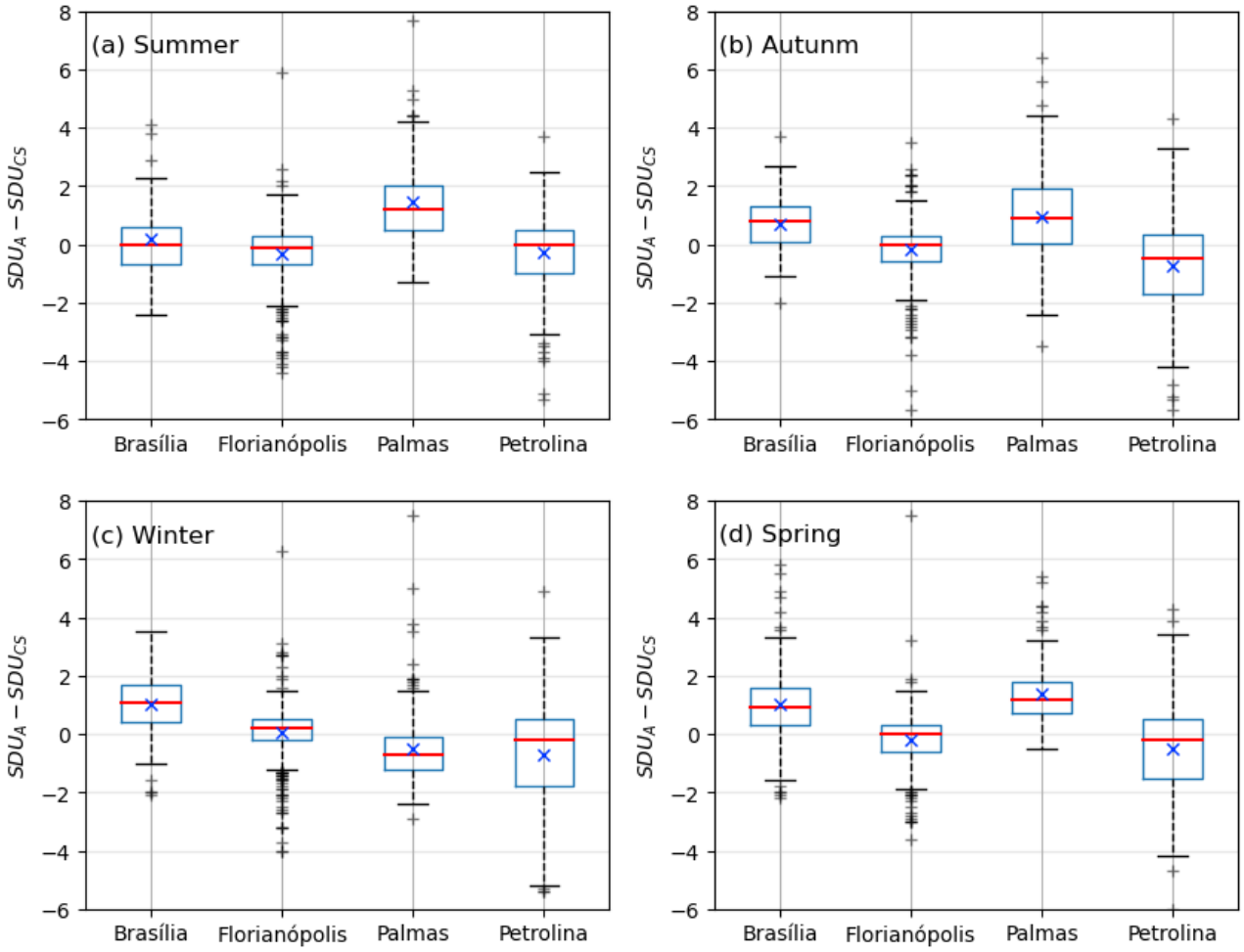


In general, Florianópolis showed the best agreement among the heliographic records

and the automatic data (Figure 3.6b). The station presented the lower MD (-0.14h) e RMSD (1.11h), and the highest coefficient of correlation (0.96).

Brasília and Palmas exhibits similar results with mean differences of 0.81 and 0.77h, respectively (Figure 3.6a,d). The slightly higher values of RMSD (1.55 and 1.67h, respectively) indicates a larger spread compared to Florianópolis. Figure 3.8 presents the box-and-whiskers plot of the differences for these stations separated by season.

Figure 3.8 - Box-and-whiskers plot of daily ground measurements differences. The "whiskers" (lines extending parallel from the boxes) indicate variability outside the upper and lower quartiles. Outliers are plotted as individual crosses. The red line indicates the median and the blue 'x' the mean.



In Brasília, during summer, the scatterplot shows great dispersion, with occurrence of both positive and negative deviations (Figure 3.8a). The other seasons present

similar behavior, with overall overestimate.

For Palmas, a small overestimate is observed through the whole year, with exception of the winter, when underestimate prevail. For this season, it can be seen that the observations concentrates mostly under the 1:1 line, which indicates that the sensitivity threshold is lower than the WMO's threshold during this season. It was expected that during winter and spring, when this location presents driest and hottest periods, the sensitivity threshold would be lower compared to other seasons, but this was observed only for winter.

Petrolina shows higher spread than the other stations considered (RMSD of 1.71h), and in opposite to them, it presented SDU underestimations with MD of -0.55h. The box-and-whiskers plot showed that the underestimate is predominant over the year, but is smaller during summer. Which is consistent with the results obtained in the scatterplot (Figure 3.6e) and above described.

3.4.1 Conclusions

These results show that on average the threshold of 120 W/m^2 is appropriate, but, as expected, the sensitivity threshold can largely change with location, and even season (LEGG, 2014). This is explained for several reasons. The sensitivity depends mainly on the atmospheric conditions, such as air temperature and humidity (KERR; TABONY, 2004). Nonetheless, other factors are also very important, and may influence the necessary energy amount to burn the card, e.g. the type of recording cards used (paper type, colour, quality of printed scale), the properties of the glass used to make the sphere (transparency, colour, scratches), and others (MATUSZKO, 2012). Unfortunately, comparisons between CS and automatic records are limited to a few locations that have overlapping measurements. For a deep investigation of the fitting of the 120 W/m^2 threshold in Brazil, and consequent conformity between the SDU data obtained from the automatic sensor and the CS recorder, it would be necessary detailed information on the maintenance and operation of the stations, but this information is not provided by the analyzed networks.

Regarding the questions stated in the beginning of this chapter, both analyzed datasets brought meaningful information on SDU, being the most common source of SDU records (i.e. data achieved through CS recorders or based on DNI acquire by automatic sensors). The results presented in the Section 3.4 indicated that although the automatic acquired data pass through quality control procedures, the resultant sunshine amount for the analyzed period is questionable. Some studies investigated,

the BSRN criteria for quality check. Urraca et al. (2017) showed that this criteria is very permissive, and their analysis found several error occurrences that were previously approved by the BSRN methodology. For the proposed research, the employment of SONDA data containing faulty records could compromise the results. The employment of erroneous data as "ground truth" in the validation procedure leads to unreliable results, as glitches in the reference dataset can be mistaken as defects of the product (LOEW et al., 2017). Therefore, to the following validation of the satellite-based SDU datasets, the measures from the CS recorder were employed as reference. Besides it has shown more reliability than SONDA records, it also presented a higher number of stations, providing a greater amount of data, allowing a larger Brazilian territory coverage.

4 SATELLITE-BASED SUNSHINE DURATION ESTIMATES VALIDATION

Despite the great importance of SDU for many applications (e.g. yield planning in agriculture (RAO et al., 1998; XUE et al., 2011; HUANG et al., 2012; WANG et al., 2015), analysis of the thermal loads on buildings (SHAO, 1990; BASSO et al., 2015b), input parameter in soil water balance models (WARNANT et al., 1994), proxy for solar irradiation (AKINOGLU, 2008), among others), worldwide the network that records SDU is often sparse and insufficient to capture complex climate characteristics (WU et al., 2016).

Due to its wide spatial and temporal coverage, geostationary satellite data presents a unique opportunity for fill the lack of information in locations that has few or none data. Many authors suggested methods for estimating SDU from satellite data (KANDIRMAZ, 2006; CEBALLOS; RODRIGUES, 2008; GOOD et al., 2010; SHAMIM et al., 2012; WU et al., 2016; KOTHE et al., 2017; ZHU et al., 2020). Among the mentioned studies, only Ceballos and Rodrigues (2008) and Kothe et al. (2017) provide SDU estimates to Brazil. Regarding Brazilian territory, however, little effort was made to furnish information on the performance of satellite-based estimates.

The reliable utilization of data obtained from satellite remote sensing depends on careful validation of the products to ensure its quality (LOEW et al., 2017). Therefore, this chapter aims to validate the CPTEC product for daily SDU estimates by means of comparison with reference ground-measurements over Brazil and inter-compare its results with the CMSAF SDU data for the period 2013-2017. In the first section of this chapter both algorithms for daily SDU estimates (i.e. CPTEC and CMSAF) are described. The Section 4.2 presents the methods used for the validation process, Section 4.3 depicts the validation and inter-comparison results and the last one (Section 4.4) displays the conclusions and the main finds of the chapter.

4.1 Satellite-based SDU estimation algorithms

4.1.1 DSA/CPTEC's method

The CPTEC is part of the INPE and has the mission to produce high quality satellite products and thus to offer relevant information for different Brazilian sectors. Currently, the CPTEC is the most advanced center for weather and climate prediction in the Latin America and develop several kinds of products, from numerical weather models to satellite-derived products. The currently Satellite and Meteoro-

logical Sensors Divison (DISSM) is the responsible for the latter. The DISSM was instituted in 2020, and incorporated the former DSA. The DSA was created in 1986, and through the time established itself as a reference in satellite-derived products generation. Among their products are sea surface temperature, severe weather monitoring, precipitation estimation, sunshine duration and several others (COSTA et al., 2018).

To estimate SDU, visible imagery acquired with GOES processed by the DSA are used. From time to time the GOES platforms are replaced, during the analysed period in this study (2013-2017), GOES-13 was operational and carried the IMAGER sensor on board. The IMAGER visible channel is centered at 0,65 μm, with a bandwidth of 0,2 μm.

The sensor measures the spectral radiance L_λ ($W.m^2.sr^{-1}.\mu m^{-1}$), that represents the mean value at the pixel area. The spectral irradiance at the top of atmosphere is $E_o = \mu S_\lambda$, where S_λ is the solar irradiance at normal incidence in this same spectral interval, and $\mu = cos(SZA)$ is the cosine of the solar zenith angle. Assuming that the reflected radiance is isotropic, the emergent spectral irradiance at the top of atmosphere is $E \uparrow = \pi L_\lambda$ and the reflectance is $R = \frac{E\uparrow}{E_o}$. Operationally, from the satellite visible imagery the reflectance factor (F) and the planetary reflectance (R) are defined as showed in Equation 4.1 (CEBALLOS; RODRIGUES, 2008):

$$F = \pi \frac{L_\lambda}{S_\lambda}; R = f\frac{F}{\mu} \tag{4.1}$$

where the factor f is a function correcting the effects of anisotropic reflection (LUBIN; WEBER, 1995), for the following purposes, f is considered 1 (CEBALLOS et al., 2004). The R is provided by DSA as a by-product of the operational processing of the GL1.2 shortwave model (CEBALLOS et al., 2004).

It is usual to regard the reflectance as a mean value between the cloud reflectance (R_{max}) and the clear sky reflectance (R_{min}) weighted by the fraction of the pixel covered by clouds (C) as showed in Equation 4.2 (CEBALLOS et al., 2004).

$$R = C.R_{max} + (1 - C).R_{min} \tag{4.2}$$

which leads to an estimate of cloudiness (C) as:

$$C = \frac{R - R_{min}}{R_{max} - R_{min}} \tag{4.3}$$

Ceballos et al. (2004) defined the value of R_{max} as 0.465, which corresponds to the transition between a cumuliform and a stratiform cloud field and the R_{min} as 0.09, a reasonable value for continental surface. In the case of $R < R_{min}$, C is set as 0, and if $R > R_{max}$, C=1. In case of $R = 0$ or marked as "invalid" (i.e. $R = -99$), C is also tagged as invalid.

Next, assuming that the average cloud cover assessed by C is also representative of the relative time of cloud passage over a site inside the pixel (PORFIRIO; CEBALLOS, 2017), 1-C corresponds to the relative time of clear sky. The daily SDU is achieved through the Equation 5.1, that is similar to the integration via trapezoidal rule (which consists of a numerical method to approximate the integral value):

$$SDU = (1 - C_1) + \frac{\Delta t}{2}[(1 - C_1) + 2(1 - C_2) + 2(1 - C_3) + \\ + ... + 2(1 - C_{k-1}) + (1 - C_k)] + (1 - C_k) \tag{4.4}$$

where C is the cloudiness parameter (described in Equation 4.3), C_1 corresponds to the first "valid" observation for the pixel, the subscript index corresponds to the number of the image within a day, k is the last valid image of the day, and Δt is the time interval between two consecutive images (for the period analyzed, usually 30 minutes).

On average, for a given pixel, 30 images are available for the daily SDU estimate. However, this value can be smaller and the interval between two consecutive images can be larger than 30 minutes. The daily SDU for a given day is considered invalid, therefore, discarded, if there is a interval greater than three hours: i) between the first image of the day and the sunrise; ii) between consecutive images; and iii) the last image of the day and the sunset and if less then 5 images were available for the estimation.

The spatial resolution of the DSA/CPTEC SDU dataset is 0.04° on a regular latitude-longitude grid of 1800x1800 pixels within latitudes 50°S to 21.96°N and longitudes 100°W to 28.04°W, and cover the time period from February 2007 to near real time.

4.1.2 CMSAF's method

The EUMETSAT's CMSAF was established to contribute to the operational monitoring of the climate and the detection of global climatic changes. With this aim CMSAF's products follow the highest standards and guidelines as lined out by Global Climate Observing System (GCOS) for the satellite data processing.

The SDU is one of the several products of CMSAF based on Surface Radiation Data Set - Heliosat (SARAH) - Edition 2.1 (PFEIFROTH et al., 2019). The data record covers the time period from 1983 to near real time with a spatial resolution of $0.05° \times 0.05°$. In order to derive the SARAH-2 surface parameters, the Heliosat algorithm is used (HAMMER et al., 2003). The Heliosat method presents the opportunity to get a continuous dataset of Effective Cloud Albedo from a combination of Meteosat Visible and InfraRed Imager (MVIRI) and Spinning Enhanced Visible and Infrared Imager (SEVIRI) measurements. To minimizes the impacts of satellite changes and artificial trends due to degradation of satellite instruments, Heliosat includes an integrated self-calibration parameter (MUELLER et al., 2011).

At first, the Effective Cloud Albedo is retrieved by the normalized relation between all sky and clear sky reflection in the visible channel of the Meteosat instruments. This parameter is used to derive cloud index, a measure of the impact of the clouds on the clear sky irradiance. The SPECMAGIC model is used to estimate clear sky irradiance, and then from the combination of cloud index and clear sky irradiance, Surface Incoming Shortwave radiation (SIS) is achieved. Thereafter using the diffuse radiation model of Skartveit et al. (1998) and the cloud index, Surface Incoming Direct radiation (SID) is calculated. By normalizing it with the cosine of the solar zenith angle, the DNI is obtained. The SID and DNI are the basis for the retrieval of SDU (KOTHE et al., 2017).

To the retrieval of the satellite-based SDU, the SARAH-2 30 minutes instantaneous DNI data and the WMO threshold are used (DNI $\geq$ 120 W/m^2). Daily SDU is calculated as the ratio of slots (Meteosat observations) exceeding the DNI threshold, considered as sunny slots, to all slots during daylight (Equation 4.5).

$$SDU = daylength \times \frac{\sum_{i=1}^{daylightslots} W_i}{daylightslots} \tag{4.5}$$

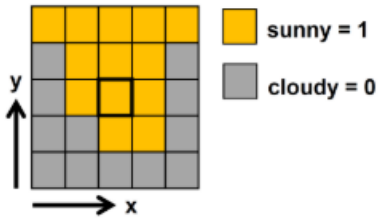
The day length is calculated depending on the date, longitude and latitude and is restricted by a threshold of the solar elevation angle of 2.5° (KOTHE et al., 2013).

W_i is a weight that varies between 0 and 1, and indicates the influence of a single slot depending on the number of surrounding cloudy and sunny grid points.

A grid point at a time slot i is accounted as sunny if DNI is 120 W/m^2 or larger (Equation 4.6). Since SARAH-2 provides instantaneous DNI data every 30 minutes, without weighting, one sunny slot would correspond to a 30 min time window. This is not the case unless it is a bright weather situation. If there are clouds in the vicinity of a grid point, probably not the whole 30 minutes are sunny. The opposite case is also valid. To account this fact, the information of the 24 surrounding grid points (Figure 4.1) and two successive time steps are used.

$$SIn_i = \begin{cases} 1\ if\ DNI(x,y) \geq 120 W/m^2 \\ 0\ if\ DNI(x,y) < 120 W/m^2 \end{cases} \tag{4.6}$$

Figure 4.1 - Demonstration for accounting for surrounding grid points. The target grid point is marked in the center.



SOURCE: Kothe et al. (2017).

For each grid point, the number of sunny points in the 24 grid points in the vicinity plus the center cell grid of interest is summed up (Equation 4.7).

$$\#SIn_i(x,y) = \sum_{m\,=\,y-2}^{m\,=\,y+2} \sum_{n\,=\,x-2}^{n\,=\,x+2} SIn_i(m,n) \tag{4.7}$$

First, for each daytime slot i this is done. Then to also incorporate the temporal shift of clouds, the number of each time step is combined with the number of the previous time step for each pixel (Equation 4.8).

$$\begin{aligned} N_1 &= \#SIn_1 \times 0.04 \\ N_i &= (\#SIn_i + \#SIn_{i-1}) \times 0.02 \end{aligned} \tag{4.8}$$

The factor 0.04 is used for the first time slot of the day, for $i > 1$, 0.02 is used, thus if all 25 grid points are sunny the resulting number N is 1, and 0 in the case that no grid point is sunny.

Thereafter the impact of sunny and cloudy grid points on the temporal length of one time slot is estimated. The fraction of time, which slot i accounts to the daily SDU is derived by Equation 4.9.

$$W_i = \begin{cases} max(N_i, C_1)\ if\ DNI(C_{gp}) \geq 120\ W/m^2 \\ N_i \cdot C_2\ if\ DNI(C_{gp}) < 120\ W/m^2 \end{cases} \tag{4.9}$$

If the center grid point (C_{gp}) DNI$\geq$ 120W/m^2, the grid point is taken as sunny, and the weight (W_i) is equal to N_i, but not smaller than C_1. Otherwise (C_{gp} DNI < 120 W/m^2), W_i is the product of W_i and C_2. These constants were derived empirically through sensitivity tests, by minimizing the bias compared to reference station data in Germany. They are set as $C_1 = 0.4$, that indicates the minimum fraction that a sunny slot can contribute, and $C_2 = 0.05$, the weight for the contribution of a non-sunny slot.

The daily SDU in hours is then derived by Equation 4.5.

4.2 Evaluation methods

To evaluate the satellite products for the period of September 2013 to December 2017, data from INMET's network were used as ground "truth". Detailed information on stations selection, as well as quality control procedures are provided in the previous chapter.

Due to the considerable extension of the Brazilian territory, as well as the great variety of biomes and climates within it, the stations were grouped by climate zones, as suggested in Raichijk (2012). This classification was developed by the IBGE and takes into account the average air temperature and precipitation regimes. The five main regions are illustrated in Figure 4.2. The characteristics of the regions are described below, indicating inside parentheses the number of stations included in each one:

- Equatorial region (27 [1]): has Af/Am climate according to the Köppen-Geiger classification, it presents average annual temperatures between 24 and 27°C, and average annual precipitation of 2300mm. Presents no dry season, or a short dry season, lasting under 3 months.

Figure 4.2 - Spatial distribution of INMET's stations.

- Tropical Equatorial region (43): with Aw/BSh climate according to the Köppen-Geiger classification. Hot, semi-arid with a prolonged dry season (over 8 months). This region comprises the northeast Brazilian Sertão, going south until approximately 10°S (RAICHIJK, 2012).

- Tropical Northeast Oriental region (21): Mainly Aw/BSh climate according to the Köppen-Geiger classification. With average annual temperatures

[1]For the CMSAF evaluation, 22 stations were used, since 5 of the listed stations are out of the METEOSAT disk.

ranging from 24 to 26°C and with maximum precipitation occurring in the late autumn-winter period (PALHARINI; VILA, 2017).

- Tropical Central Brazil region: This main region was sub-divided in a) Warm and b) Mesothermal/Subwarm (hereinafter, Mesothermal), particularly owed to its different precipitation regimes.

  a) Warm (49): Aw climate according to the Köppen-Geiger classification. This region comprises the Brazilian Central Plain. Semi-humid, marked by rainy summer and dry winter (CAVALCANTI, 2009).

  b) Mesothermal (29): Classified as Cw climate, conform to the Köppen-Geiger classification (RAICHIJK, 2012). This region exhibits average annual temperatures between 10 and 18ºC, with dry winter. It includes part of Brazil's Southeast and north of Paraná.

- Humid Temperate region (25): It is defined as Cfa by the Köppen-Geiger classification. This region presents a well distributed precipitation regime (without a dry season), with mild temperatures between 10 and 15°C.

This regionalization of stations allows for a condensed and more robust regional analysis.

In order to compare the satellite-based gridded SDU estimates with the in situ records, the satellite data was extracted at the station sites, by selecting the satellite pixel in which the station is located. Therefore, the monthly MBE, RMSE and the correlation coefficient (r) of the daily SDU were calculated for each station, and then, for each region.

The definitions of the statistical measures are presented below (WILKS, 2011):

$$MBE = \frac{1}{n}\sum_{i=1}^{k}(Z_i - O_i)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(Z_i - O_i)^2}$$

$$r = \frac{\sum_{i=1}^{k}(Z_i - \bar{Z})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{k}(Z_i - \bar{Z})^2}\sqrt{\sum_{i=1}^{k}(O_i - \bar{O})^2}}$$

Thereby, the variable z describes the dataset to be validated (e.g. CPTEC SDU)

and o denotes the reference dataset (i.e. in situ measurements). The individual time step is marked with i and k is the total number of time steps.

4.3 Validation and inter-comparison results

The MBE spatial distribution of daily SDU for the CPTEC product evaluated for each month is displayed in Figures 4.3. On average the difference between the bias evaluated for these regions for the month with its highest and smallest values lies close to 0.6 hours (Tables 4.1-4.3). Most stations south of 15°S, presented small MBE values (overall, ranging from -1 to 1 hour), with no significant variation in magnitude over the year.

Table 4.1 - Mean bias error (MBE), root mean squared error (RMSE), and correlation coefficient (r) for the comparison of daily SDU estimates (derived from satellite data) and INMET's records for the time period 2013–2017, for the Tropical Central Brazil – Warm region. The number of compared daily values is presented in the #obs column. The values in parenthesis correspond to the standard deviation (SD) of the statistical parameters within the region.

	MBE (h)		RMSE (h)		r		#obs
	CPTEC	CMSAF	CPTEC	CMSAF	CPTEC	CMSAF	
JAN	0.34 (1.0)	0.62 (0.61)	1.89 (0.65)	1.71 (0.56)	0.81 (0.09)	0.82 (0.09)	5835
FEB	0.21 (1.03)	0.69 (0.52)	1.84 (0.44)	1.69 (0.41)	0.78 (0.09)	0.79 (0.1)	5502
MAR	-0.07 (1.02)	0.73 (0.54)	1.75 (0.45)	1.67 (0.42)	0.75 (0.09)	0.76 (0.11)	5962
APR	0.02 (0.93)	0.49 (0.62)	1.69 (0.49)	1.52 (0.47)	0.73 (0.09)	0.77 (0.09)	5877
MAY	0.16 (0.68)	0.46 (0.49)	1.5 (0.29)	1.32 (0.37)	0.76 (0.11)	0.82 (0.1)	6119
JUN	0.19 (0.88)	0.54 (0.7)	1.6 (0.36)	1.38 (0.47)	0.66 (0.14)	0.76 (0.12)	5882
JUL	0.07 (0.86)	0.4 (0.64)	1.54 (0.33)	1.24 (0.43)	0.65 (0.17)	0.78 (0.12)	5966
AUG	0.08 (0.95)	0.27 (0.59)	1.54 (0.44)	1.22 (0.46)	0.65 (0.2)	0.71 (0.19)	6006
SEP	0.05 (0.98)	0.15 (0.57)	1.69 (0.47)	1.33 (0.4)	0.72 (0.12)	0.79 (0.1)	7267
OCT	-0.1 (0.98)	0.17 (0.57)	1.74 (0.4)	1.56 (0.36)	0.78 (0.09)	0.79 (0.09)	7501
NOV	0.07 (0.83)	0.36 (0.52)	1.68 (0.3)	1.51 (0.34)	0.83 (0.05)	0.84 (0.07)	7235
DEC	0.55 (0.95)	0.82 (0.64)	1.88 (0.49)	1.85 (0.51)	0.8 (0.08)	0.8 (0.11)	7451

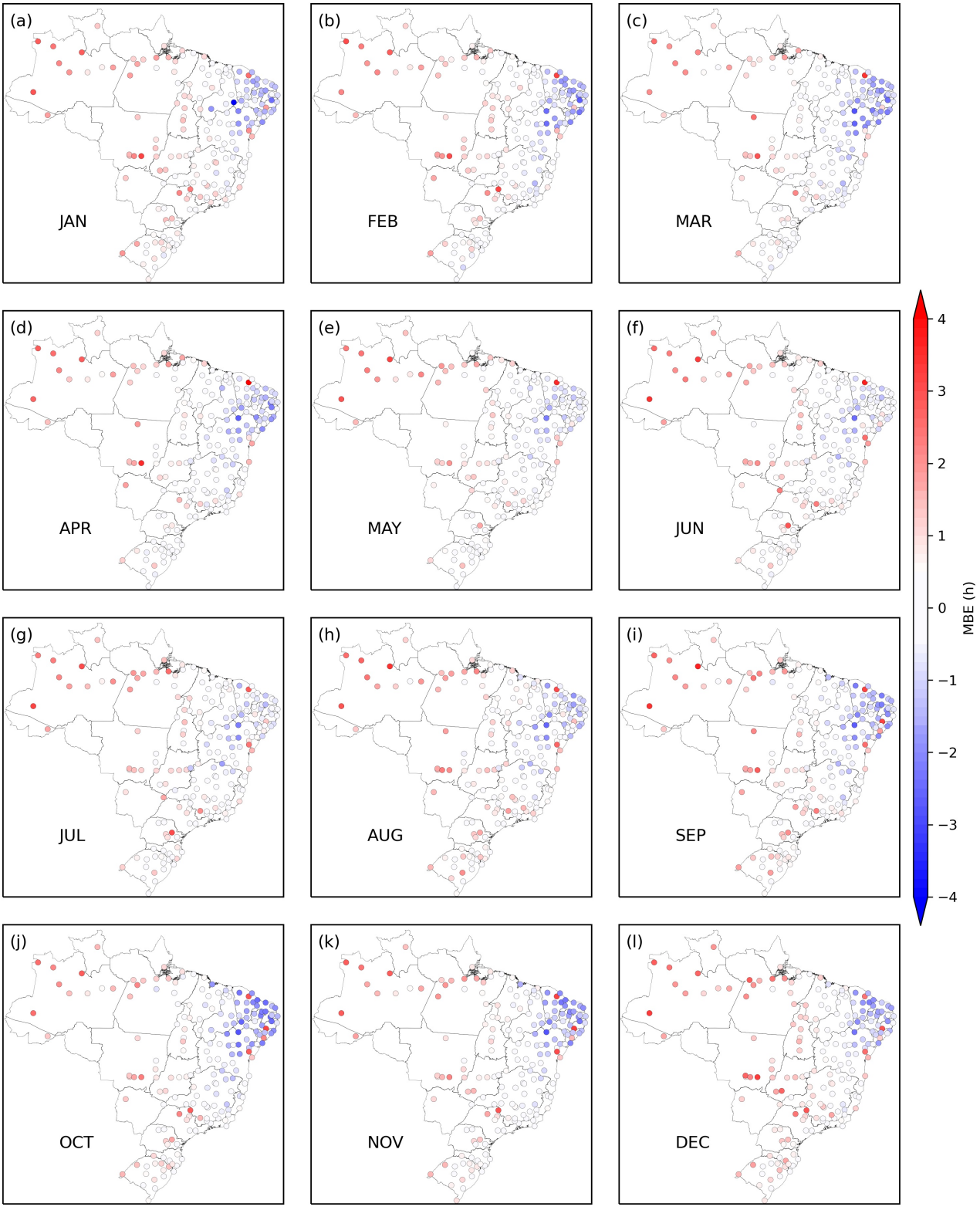
Table 4.2 - Mean bias error (MBE), root mean squared error (RMSE), and correlation coefficient (r) for the comparison of daily SDU estimates (derived from satellite data) and INMET's records for the time period 2013–2017, for the Tropical Central Brazil – Mesothermic region. The number of compared daily values is presented in the #obs column. The values in parenthesis correspond to the standard deviation (SD) of the statistical parameters within the region.

	MBE (h)		RMSE (h)		r		#obs
	CPTEC	CMSAF	CPTEC	CMSAF	CPTEC	CMSAF	
JAN	0.49 (0.74)	0.89 (0.76)	1.77 (0.46)	1.86 (0.6)	0.84 (0.08)	0.83 (0.08)	3504
FEB	0.16 (0.87)	0.76 (0.86)	1.75 (0.49)	1.78 (0.65)	0.8 (0.06)	0.82 (0.07)	3132
MAR	-0.04 (0.72)	0.89 (0.77)	1.65 (0.36)	1.83 (0.56)	0.79 (0.07)	0.8 (0.08)	3458
APR	0.03 (0.57)	0.56 (0.67)	1.5 (0.32)	1.52 (0.43)	0.8 (0.07)	0.84 (0.07)	3340
MAY	0.19 (0.45)	0.5 (0.48)	1.44 (0.3)	1.39 (0.35)	0.81 (0.08)	0.86 (0.08)	3483
JUN	0.3 (0.59)	0.55 (0.6)	1.44 (0.37)	1.42 (0.45)	0.77 (0.1)	0.82 (0.1)	3342
JUL	0.21 (0.56)	0.46 (0.52)	1.36 (0.35)	1.26 (0.37)	0.82 (0.1)	0.87 (0.08)	3375
AUG	0.32 (0.67)	0.45 (0.57)	1.38 (0.4)	1.28 (0.4)	0.82 (0.11)	0.84 (0.13)	3469
SEP	0.36 (0.71)	0.33 (0.67)	1.55 (0.41)	1.36 (0.39)	0.85 (0.06)	0.88 (0.06)	4222
OCT	0.2 (0.84)	0.34 (0.84)	1.76 (0.47)	1.61 (0.54)	0.81 (0.06)	0.86 (0.06)	4331
NOV	0.17 (0.71)	0.47 (0.75)	1.71 (0.47)	1.6 (0.58)	0.83 (0.07)	0.87 (0.07)	4249
DEC	0.57 (0.8)	1.0 (0.81)	1.82 (0.52)	1.98 (0.66)	0.81 (0.09)	0.81 (0.09)	4269

Table 4.3 - Mean bias error (MBE), root mean squared error (RMSE), and correlation coefficient (r) for the comparison of daily SDU estimates (derived from satellite data) and INMET's records for the time period 2013–2017, for the Humid Temperate region. The number of compared daily values is presented in the #obs column. The values in parenthesis correspond to the standard deviation (SD) of the statistical parameters within the region.

	MBE (h)		RMSE (h)		r		#obs
	CPTEC	CMSAF	CTEC	CMSAF	CPTEC	CMSAF	
JAN	0.61 (0.6)	0.55 (0.38)	1.69 (0.44)	1.56 (0.33)	0.86 (0.07)	0.87 (0.06)	3058
FEB	0.37 (0.63)	0.37 (0.48)	1.64 (0.45)	1.54 (0.37)	0.84 (0.07)	0.85 (0.06)	2757
MAR	0.2 (0.54)	0.41 (0.38)	1.47 (0.32)	1.33 (0.29)	0.85 (0.08)	0.88 (0.05)	3037
APR	0.22 (0.47)	0.38 (0.38)	1.46 (0.29)	1.33 (0.33)	0.85 (0.06)	0.89 (0.06)	2914
MAY	0.36 (0.45)	0.3 (0.44)	1.33 (0.37)	1.2 (0.39)	0.86 (0.08)	0.89 (0.08)	2972
JUN	0.55 (0.61)	0.41 (0.59)	1.43 (0.6)	1.26 (0.57)	0.86 (0.1)	0.89 (0.1)	2841
JUL	0.59 (0.63)	0.32 (0.67)	1.43 (0.57)	1.27 (0.6)	0.87 (0.09)	0.89 (0.1)	2925
AUG	0.72 (0.57)	0.17 (0.54)	1.48 (0.46)	1.24 (0.35)	0.88 (0.05)	0.9 (0.05)	2884
SEP	0.78 (0.57)	0.12 (0.43)	1.61 (0.46)	1.25 (0.32)	0.89 (0.05)	0.91 (0.05)	3563
OCT	0.66 (0.53)	0.21 (0.41)	1.72 (0.52)	1.46 (0.4)	0.86 (0.08)	0.88 (0.06)	3605
NOV	0.51 (0.55)	0.2 (0.38)	1.57 (0.34)	1.4 (0.32)	0.9 (0.04)	0.9 (0.05)	3526
DEC	0.7 (0.58)	0.59 (0.49)	1.69 (0.44)	1.64 (0.4)	0.88 (0.07)	0.86 (0.07)	3645

Figure 4.3 - Spatial distribution of monthly MBE (h) between daily SDU estimated through CPTEC's method and INMET's data for the time period 2013-2017. Shades of red correspond to overestimation, while shades of blue correspond to underestimation.



The stations in the north of the country, in general, have higher values of MBE with different tendencies depending on the region. Table 4.4 displays the MBE values for the Equatorial region. It presents positive bias (MBE ranging from 1.16 in October to 1.61 in December), which indicates that on average the CPTEC product overestimates the in situ measurements. The overestimation occurs for most of the stations comprised in this region, and like the southernmost stations, they do not show variations throughout the year (difference between the months with the higher and lowest bias is 0.45h).

Table 4.4 - Mean bias error (MBE), root mean squared error (RMSE), and correlation coefficient (r) for the comparison of daily SDU estimates (derived from satellite data) and INMET's records for the time period 2013–2017, for the Equatorial region. The number of compared daily values is presented in the #obs column. The values in parenthesis correspond to the standard deviation (SD) of the statistical parameters within the region.

	MBE (h)		RMSE (h)		r		# obs
	CPTEC	CMSAF	CPTEC	CMSAF	CPTEC	CMSAF	
JAN	1.42 (0.76)	0.69 (0.43)	2.23(0.59)	1.66 (0.34)	0.71 (0.09)	0.76 (0.09)	3179 (2616)
FEB	1.27 (0.77)	0.72 (0.41)	2.2 (0.67)	1.79 (0.52)	0.68 (0.16)	0.71 (0.17)	2832 (2337)
MAR	1.21 (0.76)	0.74 (0.61)	2.02 (0.62)	1.87 (0.51)	0.7 (0.11)	0.66 (0.13)	3170 (2615)
APR	1.28 (0.86)	0.58 (0.64)	2.2 (0.73)	1.78 (0.55)	0.65 (0.13)	0.67 (0.15)	3089 (2615)
MAY	1.35 (0.83)	0.48 (0.61)	2.22 (0.66)	1.74 (0.49)	0.67 (0.13)	0.7 (0.13)	3146 (2584)
JUN	1.37 (0.93)	0.37 (0.56)	2.2 (0.73)	1.51 (0.53)	0.62 (0.14)	0.69 (0.16)	3005 (2480)
JUL	1.47 (0.84)	0.35 (0.6)	2.17 (0.72)	1.42 (0.61)	0.58 (0.17)	0.65 (0.18)	2986 (2449)
AUG	1.47 (0.87)	0.35 (0.57)	2.17 (0.75)	1.42 (0.54)	0.54 (0.19)	0.55 (0.2)	3040 (2527)
SEP	1.39 (1.03)	0.3 (0.76)	2.26 (0.87)	1.65 (0.66)	0.53 (0.2)	0.52 (0.22)	3705 (3062)
OCT	1.16 (0.89)	0.38 (0.61)	2.12 (0.64)	1.83 (0.46)	0.61 (0.16)	0.53 (0.15)	3826 (3062)
NOV	1.34 (0.85)	0.7 (0.57)	2.23 (0.6)	1.84 (0.4)	0.62 (0.14)	0.61 (0.11)	3791 (3100)
DEC	1.61 (0.82)	0.95 (0.7)	2.32 (0.69)	2.01 (0.58)	0.7 (0.13)	0.65 (0.15)	3872 (3168)

The stations placed in the Northeastern Brazil, contrary to other regions, mainly show negative bias. A distinct pattern found in those regions is that the bias values present a decrease towards the winter months, approaching zero. Tables 4.5 and 4.6 present the statistical parameters for these regions. For the Tropical Equatorial region, the highest MBE value found (in magnitude) was -0.86h in October, and the lowest was -0.08h in May, indicating an overall underestimation tendency. Nonetheless, from Figure 4.3, it can be seen that there is a mixed behavior in this region. Stations located in the Maranhão and Piaui states displayed MBE values close to zero. A contrasting result is found in Guaramiranga-CE: for this stations the bias indicates overestimation, in general over 2h. This is in agreement with the results of Porfirio (2012), that found MBE values of 1.03h, being the exception of the un-

derestimation pattern found by the author for the Northeast Brazil. Regarding the Tropical Northeast Oriental region, the highest bias is registered in March, with a MBE value of -0.94h and the lowest value reported is 0.1h in July.

The seasonality found in these regions may be due to the regional cloudiness regime, since the late autumn-winter period corresponds to the rainy season (PALHARINI; VILA, 2017). Holanda et al. (2017) investigated the SDU and cloudiness patterns in Pernambuco, and they showed that the period from May to July exhibits the highest values of cloud fraction. This interval corresponds to the smallest MBE found for both regions. This suggests that the CPTEC algorithm performance, in this region, is highly related to the cloudiness conditions.

Table 4.5 - Mean bias error (MBE), root mean squared error (RMSE), and correlation coefficient (r) for the comparison of daily SDU estimates (derived from satellite data) and INMET's records for the time period 2013–2017, for the Tropical Equatorial region. The number of compared daily values is presented in the #obs column. The values in parenthesis correspond to the standard deviation (SD) of the statistical parameters within the region.

	MBE (h)		RMSE (h)		r		#obs
	CPTEC	CMSAF	CPTEC	CMSAF	CPTEC	CMSAF	
JAN	-0.39 (1.16)	0.54 (0.76)	1.97 (0.73)	1.58 (0.59)	0.76 (0.08)	0.83 (0.06)	4853
FEB	-0.37 (1.07)	0.75 (0.41)	1.87 (0.44)	1.58 (0.32)	0.76(0.08)	0.81 (0.06)	4498
MAR	-0.57 (1.0)	0.68 (0.55)	1.78 (0.52)	1.57 (0.45)	0.72 (0.08)	0.74 (0.09)	4987
APR	-0.28 (1.05)	0.46 (0.56)	1.72 (0.59)	1.44 (0.41)	0.71 (0.09)	0.76 (0.09)	4929
MAY	-0.08 (0.93)	0.51 (0.5)	1.64 (0.47)	1.38 (0.35)	0.7 (0.08)	0.78 (0.06)	5167
JUN	-0.12 (0.96)	0.46 (0.58)	1.64 (0.49)	1.28 (0.35)	0.61 (0.12)	0.75 (0.11)	4946
JUL	-0.12 (0.89)	0.38 (0.53)	1.57 (0.45)	1.2 (0.33)	0.63 (0.17)	0.78 (0.12)	4967
AUG	-0.19 (1.04)	0.44 (0.54)	1.51 (0.49)	1.01 (0.43)	0.47 (0.15)	0.63 (0.17)	5083
SEP	-0.62 (1.09)	0.28 (0.46)	1.78 (0.5)	1.05 (0.37)	0.53 (0.15)	0.68 (0.18)	6078
OCT	-0.86 (1.2)	0.3 (0.44)	1.92 (0.57)	1.33 (0.34)	0.59 (0.15)	0.64 (0.15)	6362
NOV	-0.71 (1.18)	0.44 (0.48)	2.0 (0.54)	1.33 (0.37)	0.64 (0.16)	0.78 (0.1)	6066
DEC	-0.5 (1.13)	0.65 (0.43)	1.94 (0.44)	1.48 (0.36)	0.75 (0.08)	0.82 (0.06)	6053

Table 4.6 - Mean bias error (MBE), root mean squared error (RMSE), and correlation coefficient (r) for the comparison of daily SDU estimates (derived from satellite data) and INMET's records for the time period 2013–2017, for the Tropical Northeast Oriental region. The number of compared daily values is presented in the #obs column. The values in parenthesis correspond to the standard deviation (SD) of the statistical parameters within the region.

	MBE (h)		RMSE (h)		r		# obs
	CPTEC	CMSAF	CPTEC	CMSAF	CPTEC	CMSAF	
JAN	-0.62 (1.17)	1.48 (1.17)	2.26 (0.48)	2.23 (0.99)	0.61 (0.13)	0.73 (0.11)	2506
FEB	-0.89 (1.22)	1.37 (1.14)	2.32 (0.44)	2.17 (1.0)	0.61 (0.14)	0.7 (0.15)	2193
MAR	-0.94 (1.19)	1.46 (1.1)	2.25 (0.37)	2.21 (0.92)	0.62 (0.1)	0.69 (0.1)	2498
APR	-0.73 (1.09)	1.04 (0.8)	2.07 (0.4)	1.87 (0.63)	0.67 (0.1)	0.74 (0.11)	2429
MAY	-0.16 (0.77)	1.02 (0.64)	1.72 (0.3)	1.74 (0.45)	0.77 (0.08)	0.83 (0.05)	2551
JUN	0.12 (0.94)	1.08 (0.78)	1.85 (0.45)	1.87 (0.65)	0.71 (0.12)	0.79 (0.12)	2460
JUL	0.1 (0.91)	0.98 (0.81)	1.84 (0.4)	1.81 (0.62)	0.72 (0.11)	0.79 (0.12)	2520
AUG	-0.31 (1.15)	0.97 (0.98)	2.01 (0.48)	1.87 (0.81)	0.65 (0.14)	0.74 (0.14)	2500
SEP	-0.42 (1.55)	1.12 (1.36)	2.33 (0.57)	1.98 (1.17)	0.58 (0.14)	0.73 (0.15)	3077
OCT	-0.73 (1.68)	1.25 (1.61)	2.55 (0.6)	2.24 (1.34)	0.52 (0.17)	0.64 (0.18)	3161
NOV	-0.53 (1.56)	1.38 (1.51)	2.4 (0.59)	2.21 (1.33)	0.5 (0.18)	0.64 (0.17)	3033
DEC	-0.32 (1.4)	1.75 (1.25)	2.31 (0.58)	2.49 (1.17)	0.63 (0.13)	0.7 (0.14)	3062

The Figure 4.4 shows the spatial distribution of the MBE of daily SDU for the CM-SAF product. With the exception of the Brazilian Northeast region, the CMSAF product has a similar performance to that of CPTEC. Although for the Equatorial region the CMSAF product also exhibits positive bias values, this overestimation tendency is lower when compared to that of CPTEC. Overall, the bias of the CM-SAF presents values close to 1 hour smaller than those for the CPTEC. For the Tropical Northeast Oriental region, the CMSAF product shows positive bias with considerably high values, generally over 1 hour. The highest MBE values were found during summer (1.75h in December), while the smallest were observed during winter (0.97h in August).

The CMSAF bias values for most of the regions in Brazil are similar for those reported by Kothe et al. (2017) for continental Europe, with the exception of the Tropical Northeast Oriental Region, where the bias is on average at least 1h higher than for the other regions. A similar behavior was found by the author for the Canary Islands (Northwestern Africa), in the daily evaluation (MBE values close to 2h) and for the West coast of Africa, in the monthly sums analysis (bias values up to 50h were found). The author attributed these uncertainties to two causes: the frequently low cloud fields, predominant cloud types in these regions, causes a systematically underestimation of the Effective Cloud Albedo by the Heliosat

algorithm because of the self-calibrating method (HANNAK et al., 2017), leading to an overestimation of SDU; and to the constants used during the SDU estimates. They were derived empirically using data from Germany to take into account the contribution of different sky conditions to the SDU. Consequently, the presence of low warm clouds, that are most frequent over the ocean and subtropical subsidence regions (HUANG et al., 2015) may not be well represented by these constants. This might be the case of the Northeast Brazil, since low clouds with relatively warm tops are the prevailing cloud type due to subsidence in the area associated with the Walker cell (MACHADO et al., 2014).

Figure 4.5 shows the regional MBE values. This figure condensates some of the abovementioned features. Tropical Central Brazil (Warm and Mesothermal) and Humid Temperate regions express low bias values, with CPTEC (CMSAF) displaying lower values for the first (latter). It can also be noted that for the CMSAF product, with exception of the Tropical Northeast Oriental Region, the bias throughout the year lies around 0.5h, whereas the CPTEC MBE exhibits greater variations. The better performance of CMSAF regarding the CPTEC product in the Equatorial region is well marked by, at least, 1h of difference in the bias values for these products. To Northeastern Brazil, the opposite bias sign is evident, and the seasonality in the CPTEC bias is pronounced.

Tables 4.1-4.6 summarize the regional statistical parameters. Regarding the RMSE, both satellites datasets show similar results. For the Equatorial and Tropical Equatorial regions, the RMSE of the CPTEC estimates were higher than those from CMSAF. For the other regions analyzed, the RMSE results present similar magnitudes ranging from 1.2 to 2.55h.

Concerning the correlation coefficient, the southern regions presented the highest values of r, just like the bias, which in these locations had the best results. Nevertheless, on average, all regions showed good agreement between satellite-based datasets and ground measurements. In general, the r exhibits high values, indicating a strong positive correlation.

These results are in agreement with those previously reported in the literature. Using data from Meteosat, Kandirmaz (2006) obtained RMSE varying from 0.54 to 2.79h, and coefficient of determination (r^2) of 0.78. Good et al. (2010) applied cloud classification data from SEVIRI to estimate SDU for the United Kingdom and achieved a r of 0.82. The estimate method explored by Shamim et al. (2012), obtained RMSE values from 0.66 to 2.31h, the authors found an average r^2 of 0.83.

Figure 4.4 - Spatial distribution of monthly MBE (h) between daily SDU estimated through CMSAF's method and INMET's data for the time period 2013-2017. Shades of red correspond to overestimation, while shades of blue correspond to underestimation.

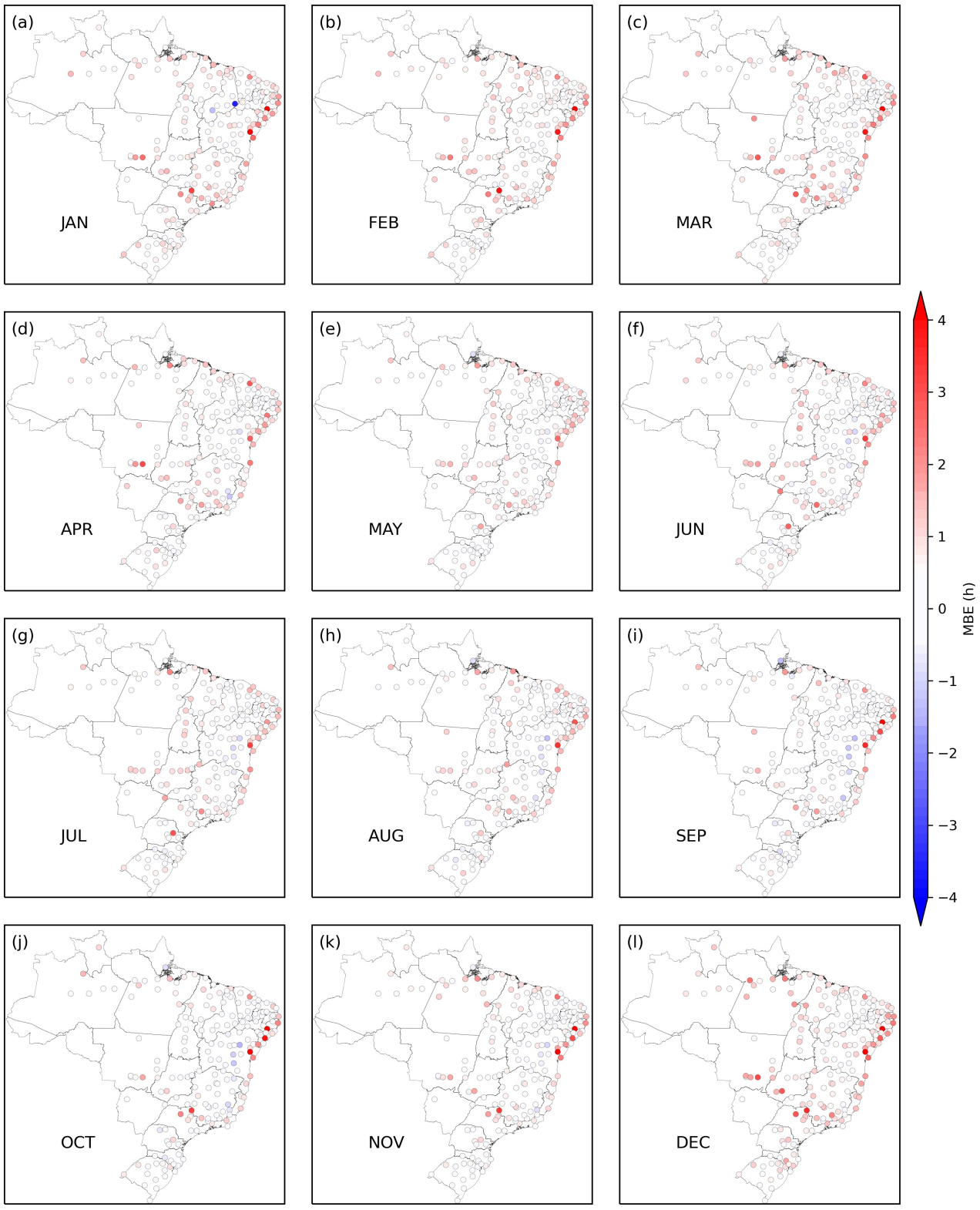
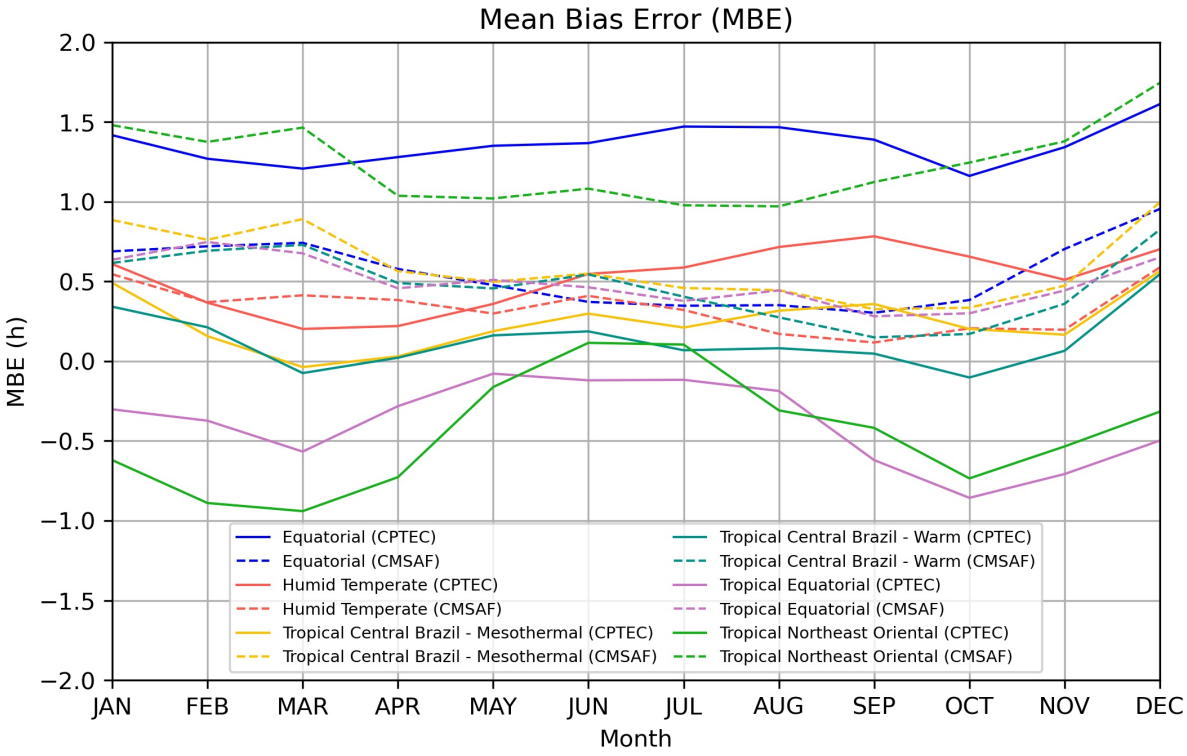


Figure 4.5 - Monthly MBE (h) of daily SDU per region, evaluated as the average of all stations within it. Continuous (dotted) lines presents CPTEC's (CMSAF's) MBE.



The Equatorial and Tropical Northeast Oriental regions presented the highest CPTEC MBE values. For the first, the bias indicated overestimation, to the latter, underestimation. Considering the main features of these regions and the results previously obtained, the deficiency in represent the SDU in these locations seems to be related to the definition of the R_{min} (the clear sky reflectance). To better analyze this situation, bivariate kernel density estimation (KDE) plots were generated.

Figure 4.6 exhibits the KDE plot for the Equatorial region. The precipitation regime, highly correlated to SDU (FUNARI, 1983), is evident: from June to October, the highest densities are located above 8 hours of SDU, corresponding to predominantly clear sky conditions. During the wet season, when partially cover and overcast sky are frequent, there is a more balanced distribution due to a greater range of observed SDU. Regardless the regime, under all sky conditions, it is observed an overestimation tendency. It is noteworthy that, both for CPTEC and CMSAF estimates, there is a non-negligible frequency of high values estimates when the ground measurement

indicates zero SDU. This might not be a deficiency of the satellite-derived models, instead it can be due to false zeros that were not flagged in the quality control procedure. Emphasizing the importance of rigorous criteria for assure the records reliability.

Figure 4.7 shows the KDE plot for the Tropical Northeast Oriental region. It can be seen a very distinct pattern between the CPTEC and the CMSAF plots. In the CPTEC plot, for all months, for values 0 and 6 hours of SDU, the higher densities observed (greenish colors) are positioned very close to the 1:1 line (red line). Indicating that, in the case of a high cloudiness amount, the algorithm performs fairly well. Also, as previously stated, the winter months present an almost perfect distribution along the 1:1 line. This happens because these months have greater cloud cover values, therefore, exhibits a higher frequency of day with partly covered or overcast sky. This result suggests that higher cloud cover amounts mitigate the effect of inadequate representation of R_{min}.

Porfirio et al. (2020) investigated the seasonal R_{min} values for the year 2016, based on images from the 15:00 UTC. They showed that over the Brazilian territory, there is a vast range of R_{min} values (from 0.04 in some parts of Amazonia to 0.10/0.11 in parts of Northeastern Brazil) and also that these values change temporally.

The results presented in this study are coherent with those reported by Porfirio et al. (2020). The overestimation (underestimation) in the Equatorial (Tropical Northeast Oriental) region can be explained by the misrepresentativeness of the R_{min}. Since that the prevailing values of R_{min} in this region is lower (higher) than the defined 0.09, the algorithm frequently considers clear sky (cloudiness) in regions that may not present this condition.

Figure 4.6 - Bivariate kernel density estimates (KDE) for INMET's measurements and (a) CPTEC's estimates and (b) CMSAF's estimates plotted by month for the Equatorial region. Yellowish values corresponds to greater probability of occurrence, and bluish values to smaller probability of occurrence.

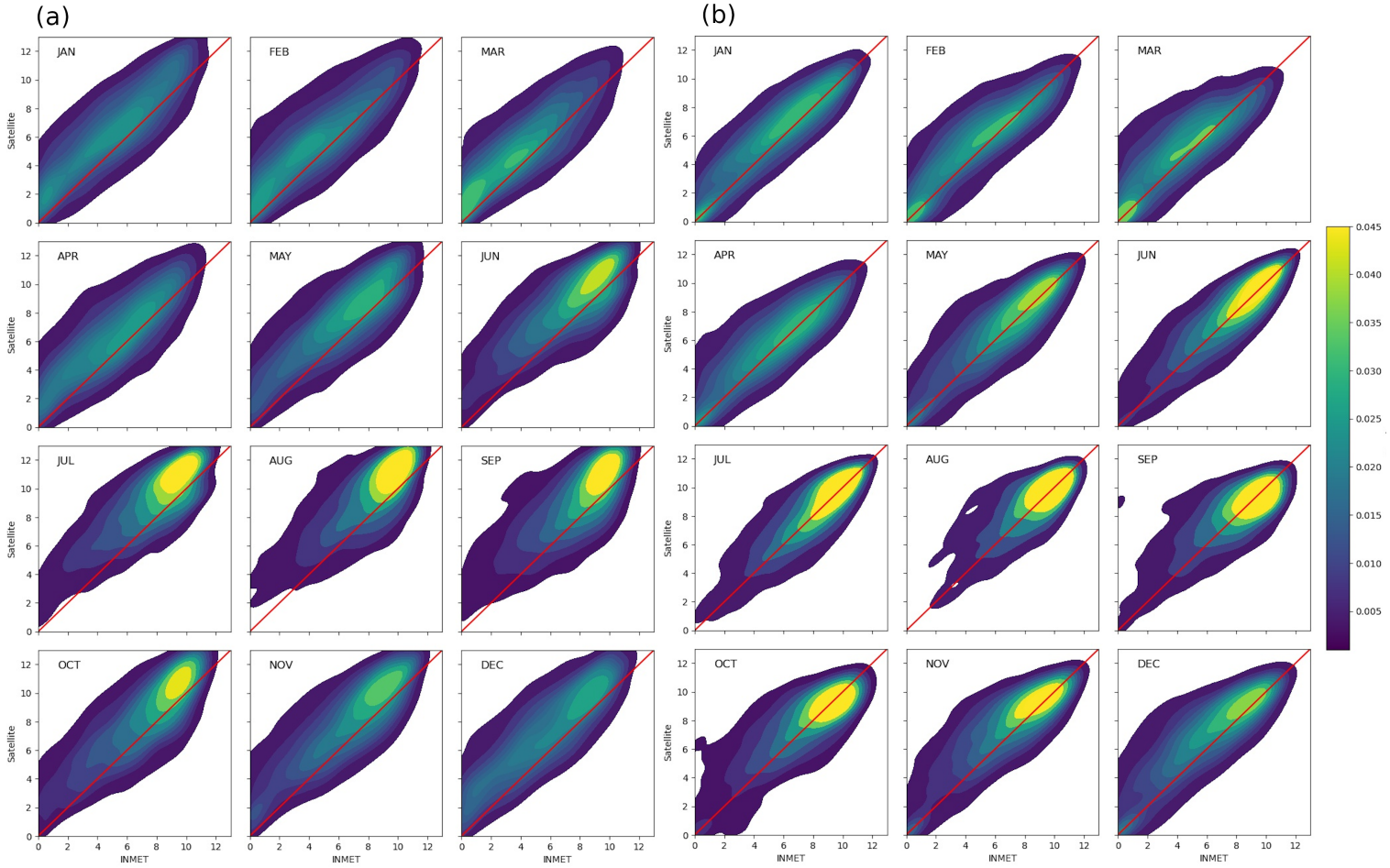
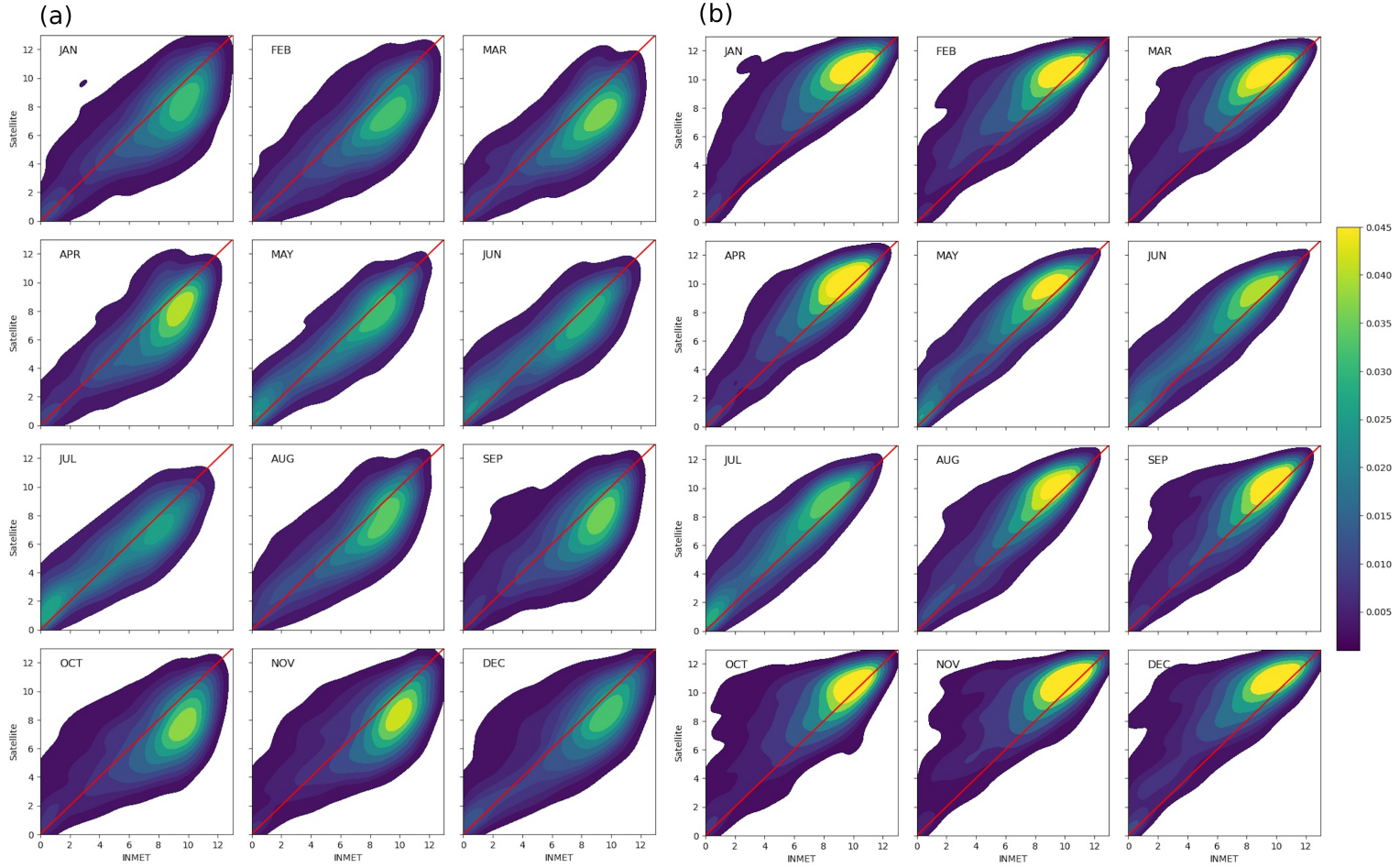


Figure 4.7 - Bivariate kernel density estimates (KDE) for INMET's measurements and (a) CPTEC's estimates and (b) CMSAF's estimates plotted by month for the Tropical Northeast Oriental region. Yellowish values corresponds to greater probability of occurrence, and bluish values to smaller probability of occurrence.



4.4 Conclusions

Within this chapter, the daily satellite-based SDU developed by DSA/CPTEC was evaluated against ground measurements. Its results were compared with the ones obtained from CMSAF SDU estimates. This product was previously validated for Europe regarding the daily values, but for South America, only the monthly sums were considered. Therefore, this study also brings insightful information on the performance of CMSAF product over Brasil.

The MBE results indicated that both products had a good performance for most of the regions. The best performance for both datasets was on the southern locations: CPTEC dataset obtained lower MBE than CMSAF for the Tropical Central Brazil (Warm and Mesothermal). The opposite was found for the Equatorial and Humid Temperate regions. Northeastern Brazil displayed the worse results with higher MBE values for both products, although with contrary signs. For the Tropical Northeast Oriental region, CPTEC product exhibits MBE values ranging from -0.94h for March to 0.1h for July, while the CMSAF presented bias within 0.97h in August and 1.75h in December. For the Tropical Equatorial region, the bias ranged from -0.08 to -0.86h for the CPTEC dataset, and from 0.28 to 0.75h for the CMSAF. The CPTEC showed an underestimation tendency with seasonal variation, where approaching winter months the bias is reduced, probably related to the cloudiness parameterization. Whereas the CMSAF presented an overestimation behavior, which might be due to the combined effects of misrepresentation by the constants used in the estimation and the systematically underestimation of the Effective Cloud Albedo by the Heliosat method in regions with frequent low warm clouds.

The RMSE and r results varied within the regions. The RMSE results of both satellite-derived datasets were very similar, except for the Equatorial and Tropical Equatorial regions, in which RMSE of CPTEC estimates were higher than those from CMSAF. On average, the rs indicated a strong positive correlation between the SDU estimates and the ground measurements. The results obtained were in agreement with the previous reports for satellite-based SDU estimates in the literature.

The bivariate KDE plots for the Tropical Northeast Oriental region suggested that for Northeastern Brazil the CPTEC estimates were close to the observations for days with low SDU values, which indicates that under partly cover and overcast sky, the CPTEC method performs better in this region. For the Equatorial region, the method showed a tendency to overestimate the ground measurements under all sky conditions. These results indicate that the value of R_{min} of 0.09 is inadequate for

these regions. Which suggests that by better representing the clear sky reflectance, the SDU estimate method performance could be improved. This is also in agreement with Porfirio et al. (2020), that reported that the R_{min} over Brazil varies both spatial and temporally. The R_{min} issue on SDU estimative will be explored in the next chapter.

5 POTENTIAL IMPROVEMENTS ON THE SDU CPTEC MODEL

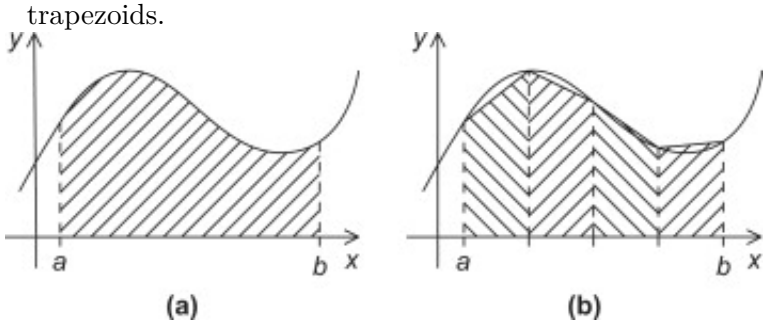
In the previous chapter, the CPTEC model for estimating SDU was evaluated against in situ measurements. The results suggested an overall good performance over Brazil, with MBE and RMSE similar to those reported in the literature. Nonetheless, some regions presented larger bias, with both under- and overestimation tendencies. This behavior can be partially attributed to the inadequate representation of the R_{min} value in these locations. This misrepresentation is in agreement with the results founded by Porfirio et al. (2020). The authors evaluated the Global Solar Irradiance estimates from the GLobal radiation (GL) 1.2 satellite-based model over Brazil. This model was also developed at DSA, and currently ran operationally at CPTEC/INPE. The GL 1.2 adopts the same parameterization for assessing cloud cover as the CPTEC SDU estimation model. Their results indicated that for Southern regions in Brazil, the model exhibits its best performance. It also presents an east-west gradient in the MBE values, with positive bias in the Western part (mostly in North and Midwest) and negative bias in the Eastern. They discussed the error sources, and pointed out that constant value of the R_{min} equal 0,09 is one of the main reason to leading downward solar radiation at surface uncertainties. Since a constant value misleads the strong spatial-temporal variation of the clear sky reflectance over the Brazilian territory. Motivated by their results, this study also investigate the effect of R_{min} on the SDU estimate.

This chapter aims to explore some of the main deficiencies and limitations of the current satellite-based SDU model, with suggestions for further improvement. In the first section, the treatment regarding the beginning and end of the day is investigated. Section 5.2, provides an analysis of the influence of the number of images used in the SDU estimate on the model's quality. In Section 5.3, an attempt to better represent the clear sky reflectance is carried out, and the impact on the model performance is evaluated. The last section summarizes the main results of this chapter.

5.1 Sunrise-sunset constraints

To estimate SDU, the CPTEC model uses a cloudiness parameter (C) to assess the cloud cover, achieved based on the visible imagery for a given time. It is assumed that the average cloud cover assessed by this index represents the relative time of cloud passage over a location inside the pixel area, therefore, 1-C gives the relative time of clear sky. The SDU is then calculated as the sum of clear sky periods. It is accomplished by means of the trapezoidal rule, that is a numerical method for approximating the definite integral. The method is illustrated in Figure 5.1.

Figure 5.1 - The trapezoidal rule: (a) area to be estimated and (b) approximate area using trapezoids.



SOURCE: Pacheco (2011).

Given the formulation of C, it can only be inferred for the instants that have observations. Data obtained in the visible channel corresponds to the sun irradiance reflected by the Earth system, which implies that it is available only in the diurnal period.

As a consequence, with the exception of observations that coincide with the sunrise/sunset, the first image in a day is available after the sunrise, and the last image of the day would be before the sunset. Therefore, the interval between the sunrise and the first image, as well as the last image and the sunset, is not accounted in the integration via the trapezoid rule.

To work around this issue, the current method adds the relative time of clear sky for the first and last image to the total SDU. Nonetheless, this may lead to errors, because the interval that has no data is not properly accounted in the estimation. For instance, if the interval between the image and the sunrise/sunset corresponds to less than 1 hour, this approach shall overestimate the SDU, mainly in sites with overall clear sky conditions within the day. In fact, this can even estimate SDU values greater than the maximum possible day length. Figure 5.2a exemplify this occurrence, it can be seen that some locations in the state of Maranhão present estimations up to approximately 48 minutes beyond the maximum possible day length. Figure 5.2b shows the average C to the correspondent day, the areas with more pronounced overestimation are in general locations that exhibits small values of C, indicating low cloud occurrence.

Figure 5.2 - SDU estimates greater than the maximum possible day length: (a) Difference between the CPTEC SDU estimate and the maximum possible day length for 15/07/2017; (b) Average cloudiness fraction to the correspondent day.

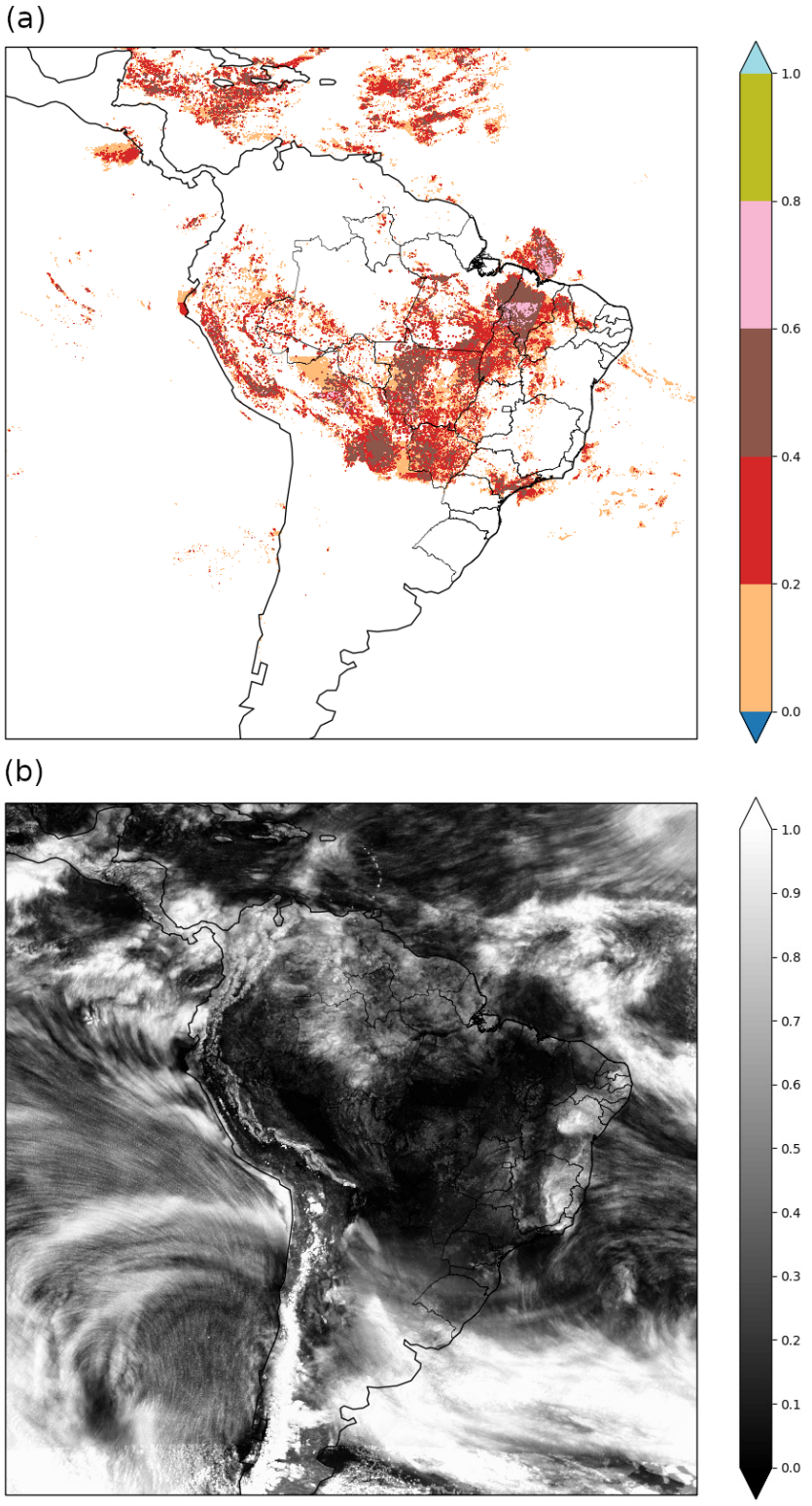
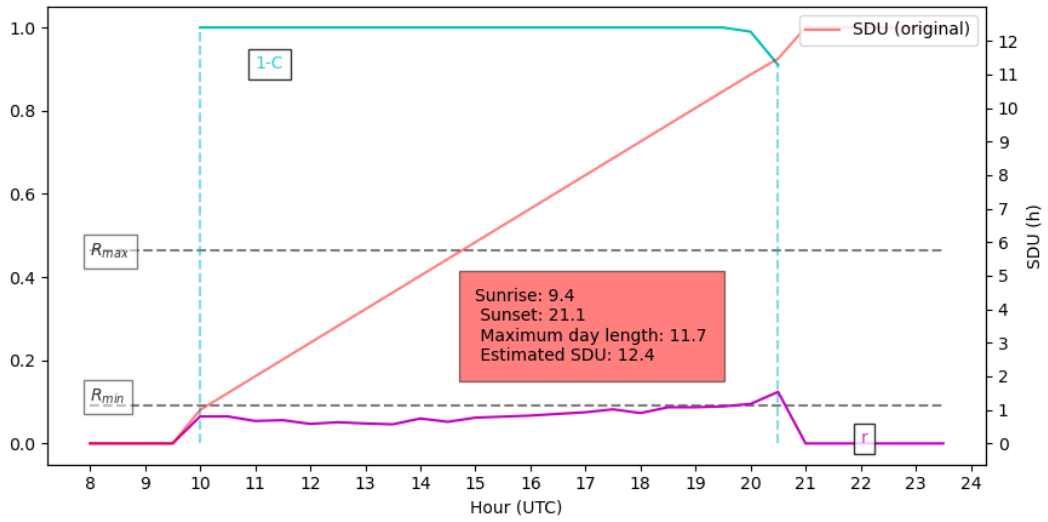


Figure 5.3 illustrates the SDU estimation for 15/07/2017 at Imperatriz-MA station. It was selected because it was located in a region with high SDU overestimation. As shown in the Figure 5.3, the integration of 1-C in the interval between 10 and 20.5 UTC is consistent with the trapezoid rule. The treatment regarding the beginning and the end of the day, however, adds unrealistic contributions to the total SDU. Since the maximum SDU between 9.4 and 10 UTC is 0.6h, instead of the 1 hour added value. The same occurs in the end of the day: the last valid image in the diurnal period is taken at 20.5 UTC, and the sun goes down at 21.1 UTC, therefore, this interval should be the upper limit to the contribution for the total daily SDU.

Figure 5.3 - Example of the SDU estimate for 15/07/2017 at station of Imperatriz - MA (5.53 °S, 47.48 °W).



To overcome this issue, some considerations regarding the sky condition at the sunrise and sunset can be made, to proper account the contribution of the intervals between the first (last) image and the sunrise (sunset) to the total SDU. The two simplest possibilities to it are:

a)

$$SDU = (1 - C_1)\Delta t + \frac{\Delta t}{2}[(1 - C_1) + 2(1 - C_2) + 2(1 - C_3) + \\ + ... + 2(1 - C_{k-1}) + (1 - C_k)] + (1 - C_k)\Delta t \quad (5.1)$$

b)

$$SDU = \sum_{i=1}^{k}(1 - C_i)\Delta t \quad (5.2)$$

The first one corresponds to consider that the sky condition at the sunrise (sunset) is similar to the moment of the first (last) image. Considering the usual interval between images (30 minutes) this is a reasonable assumption. The second one would assume that the clear sky fraction is zero at the sunrise/sunset. Both considerations can yet lead to errors, since the sky condition in the moments that do not have images, may not be correctly represented by these assumptions, and the quantitative advantages of these approaches must be further investigated. Nevertheless, adding time constraints will prevent estimates greater than the maximum day length, and account for the intervals without data being consistent with the trapezoid rule.

5.2 Amount of images in the SDU estimation

Due to the formulation of the algorithm for SDU estimation, the interval between consecutive images has great impact in the model performance. It assumes that the cloud cover in a given moment assessed through the C (cloudiness parameter) is also representative of the relative time of cloud passage over a site inside the pixel. This is a reasonable assumption over the usual 30 minutes interval between images because the lifetime of a fair-weather cumulus cloud is less than a half-hour and its life cycle would be completed within a small displacement (less than 12 km) (PORFIRIO; CEBALLOS, 2017). However, as this interval increases, this premise becomes less plausible.

There are several reasons for the interval between images to get larger than the usual 30 minutes. It can occur due to maintenance, failure of the satellite instrument, data storage issues, or during GOES rapid scan operation, when scanning of the Southern Hemisphere is curtailed sharply, and an image is available only every 3 hours (BOEING, 2006; COSTA et al., 2018).

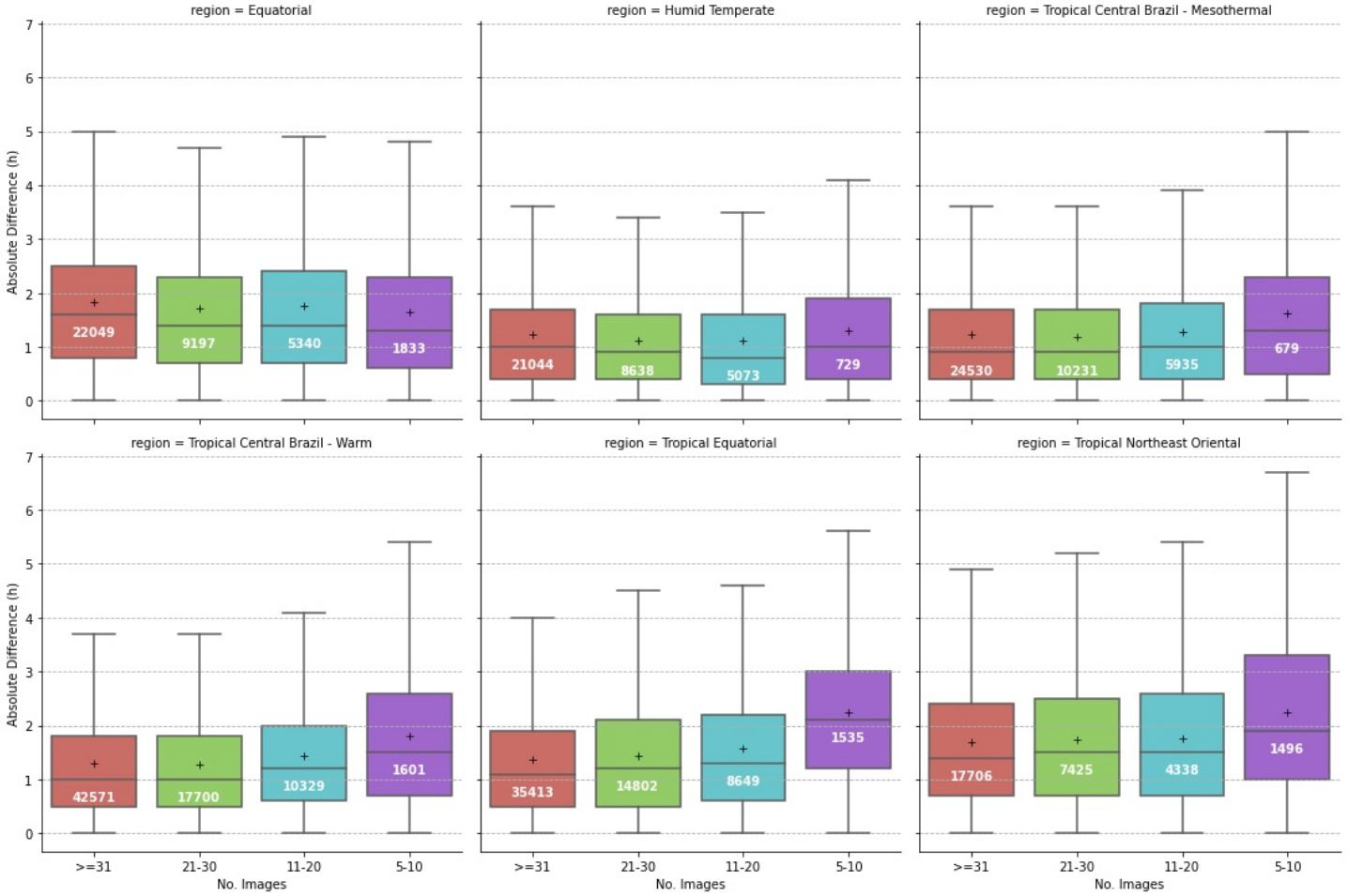
To assess the impact of the number of available images in the final estimation, the data was grouped into 4 categories: between (i) 5-10, (ii) 11-20, (iii) 21-30, and (iv)

over 30 images used. Afterwards, it was calculated the absolute deviation between the CPTEC dataset and the ground measurements for each region, considering the aforementioned categories. For this step, it was used the same period considered in previous analysis, i.e. from September 2013 to December 2017. The results are shown in the box-and-whisker plot in Figure 5.4.

It can be seen that, with exception of the Equatorial region, when the number of available images for the estimation is reduced, the absolute difference becomes higher and with greater dispersion, as expected. For the Equatorial region the MAD does not presents significant variation, neither the spread (indicated by the whiskers) shows fluctuations related to the number of images in the estimation. The Tropical Northeast Oriental region exhibits in general the highest MAD for all categories, as well as higher spread. The difference between MAD for the category with over 30 images and the one with the smallest number of observations is approximately 0.5h. For the remaining regions, the MAD ranges from approximately 1.2h for the "over 30 images" category to close to 2h for the category with the smallest number of observations. Figure 5.4 also shows that estimations made with 5-10 images are less frequent that other amounts and days with over 30 images available are the more frequent.

These results indicate that the uncertainty of SDU product increases with decreasing number of available images. It is suggested, then, to provide quality flags with the satellite-based SDU data indicating the number of observations available for the calculation. This is a commonly implemented measure, providing useful information to the data users and enabling judgment on the fitness-for-purpose of the data for their specific applications (NIGHTINGALE et al., 2018).

Figure 5.4 - Boxplot of the absolute difference between CPTEC SDU estimate and in situ measurements. The data was divided within regions and the number of images used for the daily estimation. The values inside the boxes corresponds to the number of observations included in each bin. The crosses indicate the mean absolute deviation.



5.3 Clear sky reflectance assessment

In Chapter 2 a spatial variation in the MBE values was found, with an east-west gradient. This feature is coherent with the results obtained by Porfirio et al. (2020), during the evaluation of the global solar irradiance estimated from GL1.2. It was expected since the global irradiance and SDU are highly correlated (STANHILL, 2003). Besides that, the reflectance used in the SDU estimate is a GL1.2 model subproduct, and the cloud cover is assessed in the same fashion in both estimation models.

Porfirio et al. (2020) analyzed the main errors sources in estimating the global irradiance, among them, they highlighted the importance of a proper assessment of the clear sky reflectance to the estimation of the cloud cover. The authors noticed that an underestimation of R_{min} leads to an overestimation of cloud cover, implying an underestimation of ground solar irradiance by the GL model, and vice versa. The effects to the SDU estimation are similar.

Previously, Porfirio (2012) and Porfirio and Ceballos (2017) investigated, respectively, the CPTEC SDU and DNI estimates over Northeast Brazil. Porfirio (2012) stated that the constant value for R_{min} can especially affect the SDU estimates results, since it varies geographically. Porfirio and Ceballos (2017) suggested that to proper account for cloud cover, it is convenient to adapt R_{min} to the environment, since it exhibits changes spatial and temporally. The authors obtained seasonal R_{min} fields, and used it to the DNI estimates, achieving overall good performance.

Several attempts to properly assess the clear sky reflectance are reported in literature, e.g. Perez et al. (2002) defines the lower reflectance as the average of the 10 lowest pixels in a sliding time window of 18 days in summer and 5 days during winter. Rigollier et al. (2004) uses an iterative process to filter out the minimum values that can be result of cloud shadows, which is the approach used by the CMSAF.

To investigate the R_{min} fluctuations and its effects in the model results, the central months of each season were selected (i.e. January, April, July, October), and monthly R_{min} fields were generated for the analyzed period, i.e. from October 2013 to October 2017. The methodology performed was based in the one described by Porfirio and Ceballos (2017): R_{min} fields were achieved by taking the minimum (not null) reflectance values for each pixel considering the available images in the interval between 14 and 16 UTC, within the target month. Thereafter, the fields were smoothed by performing 3x3 pixel means. The time window of one month was chosen following the consideration of Rigollier et al. (2004). They stated that this

window should be long enough to filter out cloud contamination but short enough to consider seasonal changes in the albedo surface. Seventeen monthly R_{min} fields were obtained, i.e. one R_{min} field to October 2013, one to January 2014, and so on.

To illustrate the resulting fields, Figure 5.5 displays the average R_{min} for each considered month in the analyzed period, e.g. Figure 5.5a presents the mean R_{min} field for January within the period of 2014-2017. The above mentioned characteristics can be observed: the fields present great variability within the country. Over Brazil, higher reflectance values are observed in Northeast Brazil and lower values in the Amazon region for all seasons. Most of the regions display seasonal variations in reflectance.

These monthly R_{min} fields were used as input to the CPTEC SDU model, and new estimates were performed following the same steps of the original model (Section 4.1.1). For each pixel, the R_{min} value was defined by the correspondent pixel in the R_{min} monthly field, e.g., for the estimates for January 2014, the monthly R_{min} field for the respective year and month was used.

From the previous analysis, it could be seen that the worst performance obtained with the original product regarded the Equatorial and Northeastern regions with higher MBE values, although with opposite tendencies. To verify if the modification brought improvements to the model's performance, the SDU estimates from the modified product were evaluated for these regions by calculating the MBE, RMSE and r for the new dataset for each considered month for the above mentioned period (following the definitions from Section 4.2). Then, the results were compared to the ones from the original product.

Figure 5.6 presents the comparison of the spatial distribution of the MBE obtained for the original and the modified product. It is shown that the bias for the most of the stations in the Equatorial region was significantly reduced for all considered months, with some stations presenting bias even 1 hour smaller. For the Tropical Equatorial and Tropical Northeast Oriental regions, it can be seen that the modification in the R_{min} does not have much effect on the bias. Stations in Maranhão and Piauí states, presented low bias both in the original and the modified product. The seasonality in the bias, with smaller underestimation tendencies towards the winter are still present after the modifications.

The obtained results showed that the modifications in the R_{min} field presented a more significant impact in the estimates for the stations in the Equatorial Region

than in the ones from Northeastern locations. It may be due to the greater magnitude of the difference between the "true" clear sky reflectance and the previous constant value (i.e., 0.09) for the Amazon region in comparison for Northeast Brazil.

Figure 5.5 - Average R_{min} obtained for each analysed month for the period from Oct/2013 to Oct/2017. Blueish (Redish) colors corresponds to low (high) reflectance values.

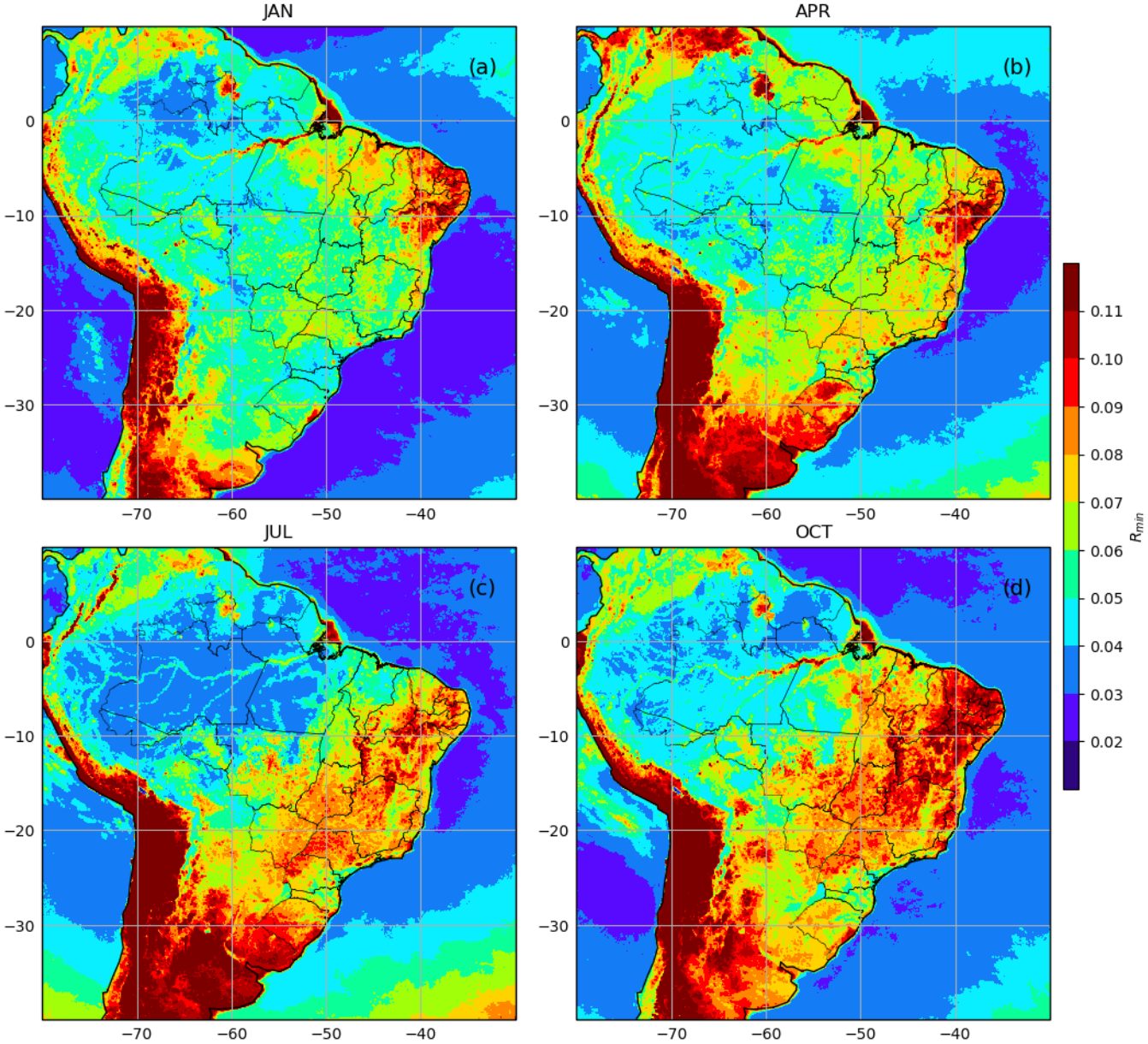
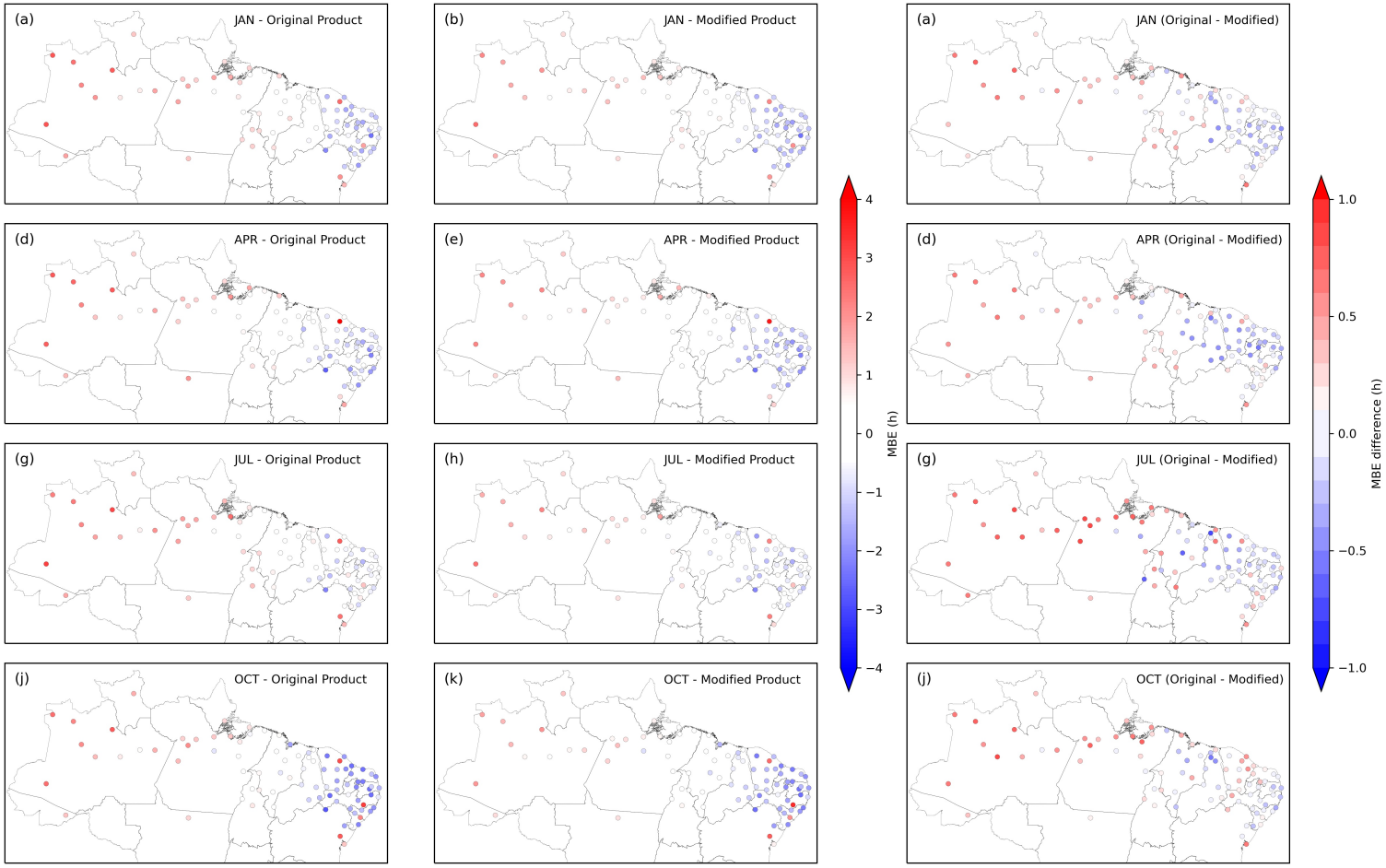


Figure 5.6 - Spatial distribution of the MBE. The first column presents the MBE for the original product; in the second column are displayed the MBE for the modified R_{min}, and the last column shows the difference of the absolute values of MBE of the original and the modified product. Each plot corresponds to a month. For the first two columns the red corresponds to positive MBE values, indicating a tendency to overestimation, while blue corresponds to negative MBE values, and a tendency for underestimation. In the last column the positive (negative) values also represented by red (blue) colors indicate station where the modified product obtained MBE values lower (higher) than the original product.



Figures 5.7-5.9 shows the KDE for INMET's measurements and the evaluated products for Equatorial, Tropical Equatorial and Tropical Northeast Oriental regions, respectively. For both regions it can be seen that the modification in the R_{min} field induced a slight reduction in the spread.

For the Equatorial region, Figure 5.7 shows that the higher concentration of observations (yellowish colors) are closer to the 1:1 line for the modified product than for the original one, as expected due to the reduction observed on the MBE.

The plot for the Tropical Equatorial region shows that for January, under clear sky conditions the underestimation observed in the original product seems to be reduced, although for other skies conditions it does not show a significant impact. The other months considered have not shown meaningful changes.

The Tropical Northeast Oriental region has not presented great modifications in the estimates for the overcast and partially covered sky (between 0 and 6 hours, approximately). For clear sky conditions, however, it seems that the change in the R_{min} field introduced a slightly stronger underestimation tendency.

Table 5.1 presents the statistical parameters obtained per region for the considered months. Although not all stations within a region exhibit identical patterns, this grouping was considered useful for summarizes the analysis throughout the Brazilian territory.

The previously pointed reduction in the spread is illustrated by smaller RMSE values, although it is more significant for the Equatorial region than for the other evaluated ones. The MBE values for the this region reinforce the overall smaller overestimation tendency of the modified product. The other regions showed a regional tendency to underestimation a little higher for the modified product than for the original one, which was also previously noticed in the KDE plots. The correlation coefficient results did not show significant improvements.

Figure 5.7 - Bivariate kernel density estimates (KDE) for INMET's measurements and the original product (left panel), and the modified product (right panel) plotted for each considered month for the Equatorial Region. Yellowish values corresponds to greater probability of occurrence, and bluish values to smaller probability of occurrence.

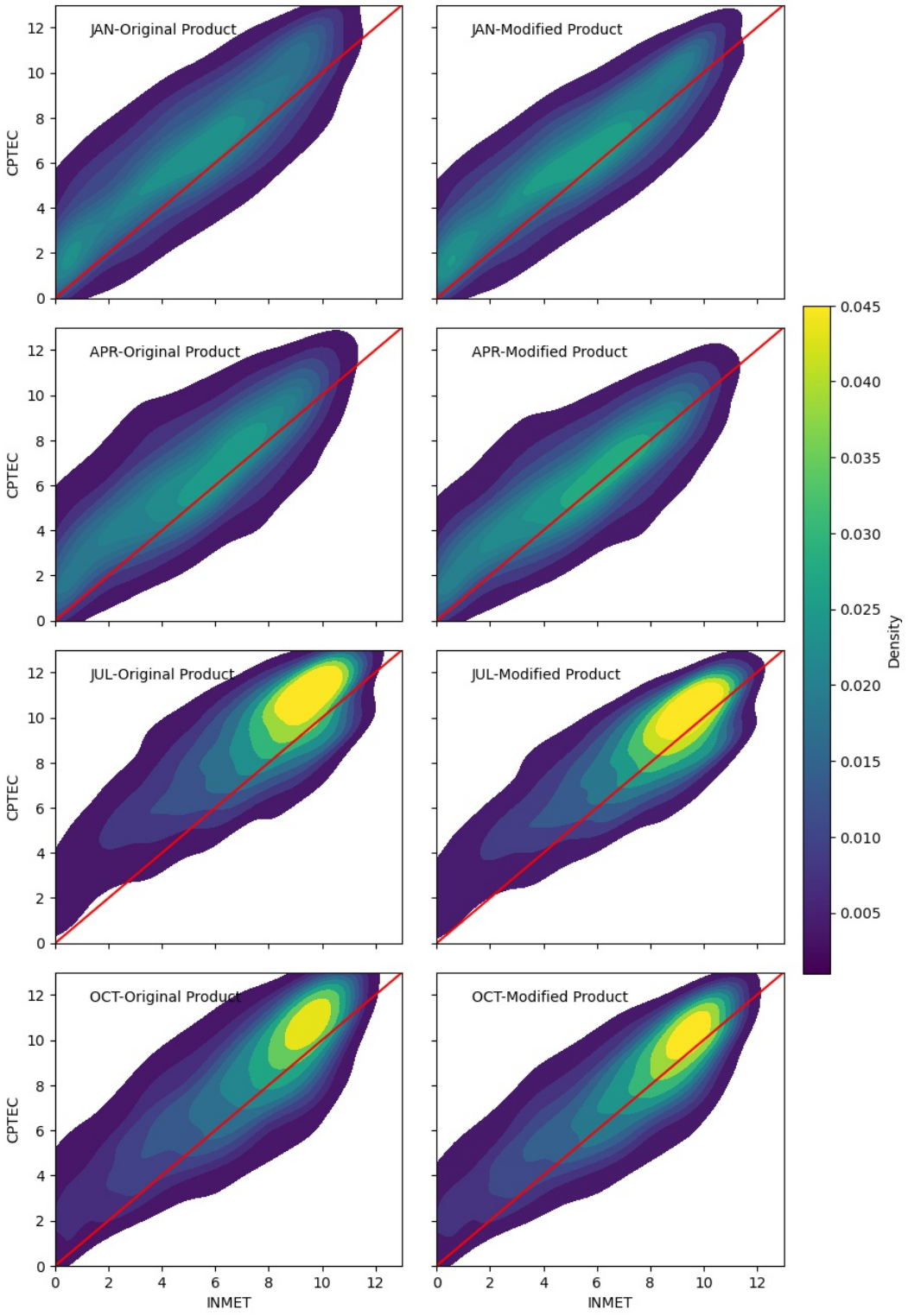


Figure 5.8 - Bivariate kernel density estimates (KDE) for INMET's measurements and the original product (left panel) and the modified product (right panel) plotted for each considered month for the Tropical Equatorial Region. Yellowish values corresponds to greater probability of occurrence, and bluish values to smaller probability of occurrence.

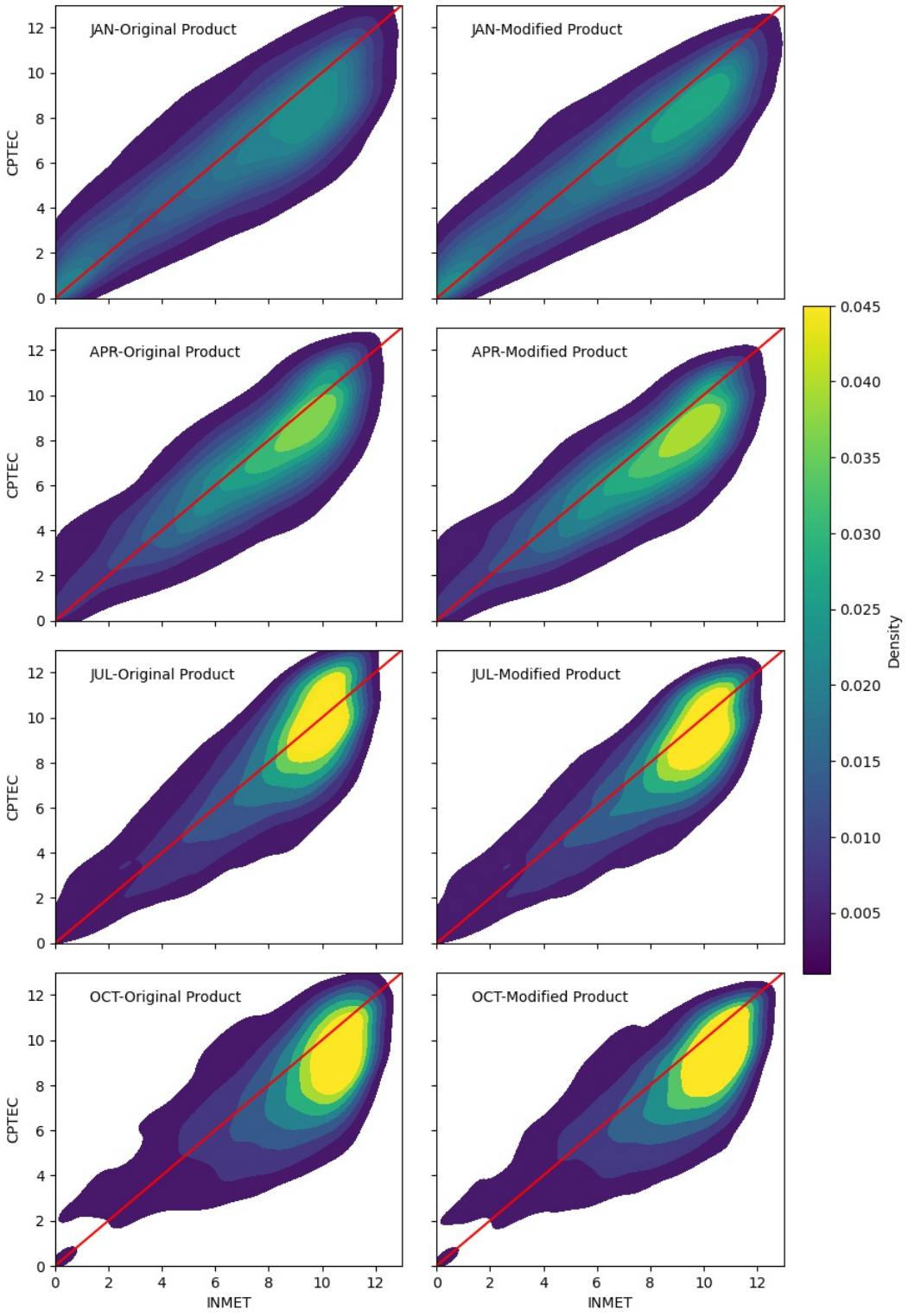


Figure 5.9 - Bivariate kernel density estimates (KDE) for INMET's measurements and the original product (left panel) and the modified product (right panel) plotted for each considered month for the Tropical Northeast Oriental Region. Yellowish values corresponds to greater probability of occurrence, and bluish values to smaller probability of occurrence.

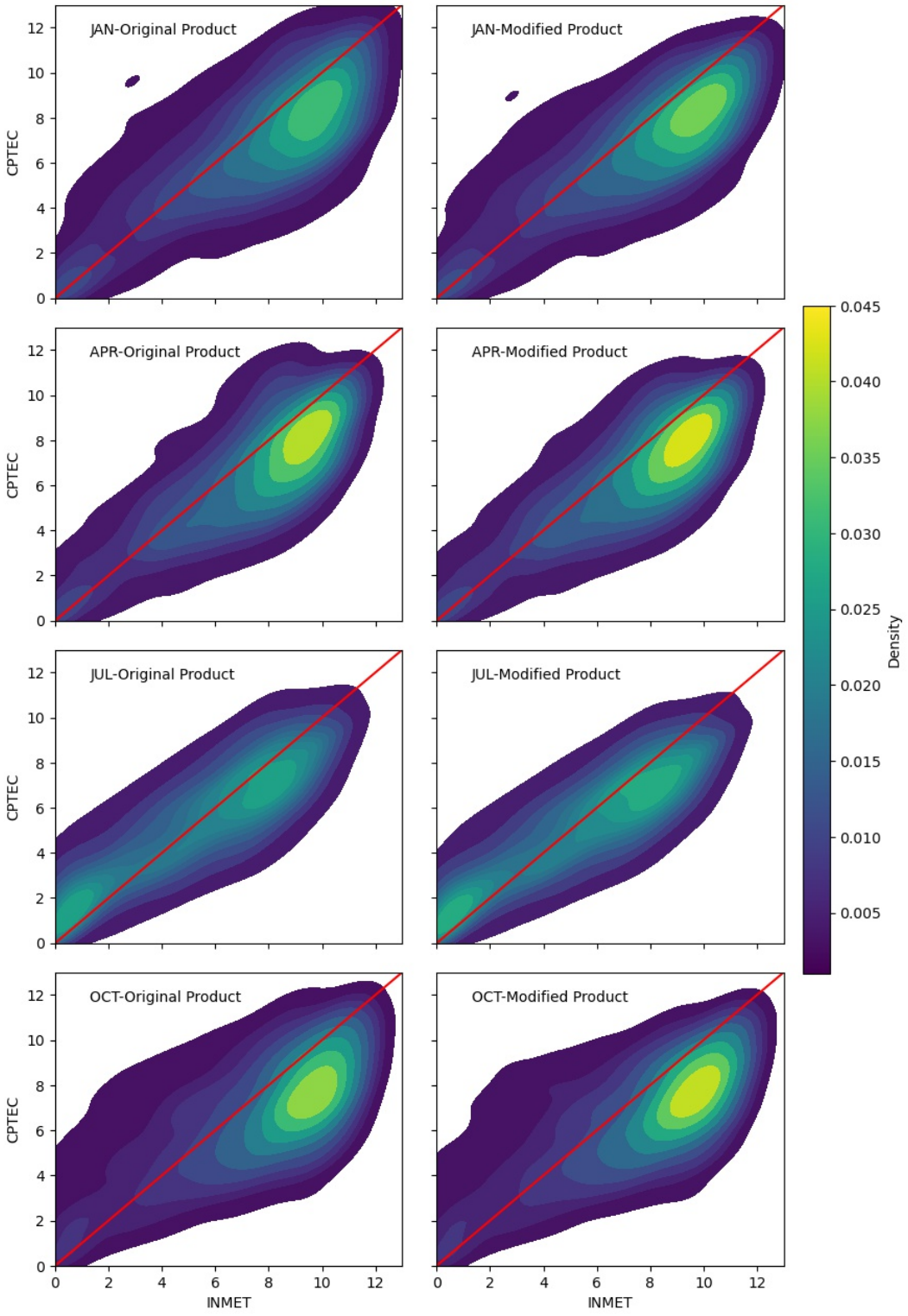


Table 5.1 - Mean bias error (MBE), root mean squared error (RMSE) and correlation coeficient (r) for the comparison of daily SDU estimates (derived from satellite data) and INMET's records for the time period 2013–2017. "CPTEC" corresponds to the original product and "CPTEC-Modified" stands for the dataset with R_{min} modified. The statistic parameters were evaluated per region by averaging all stations within the region. The number of compared daily values are is the same that for Tables 4.1-4.6.

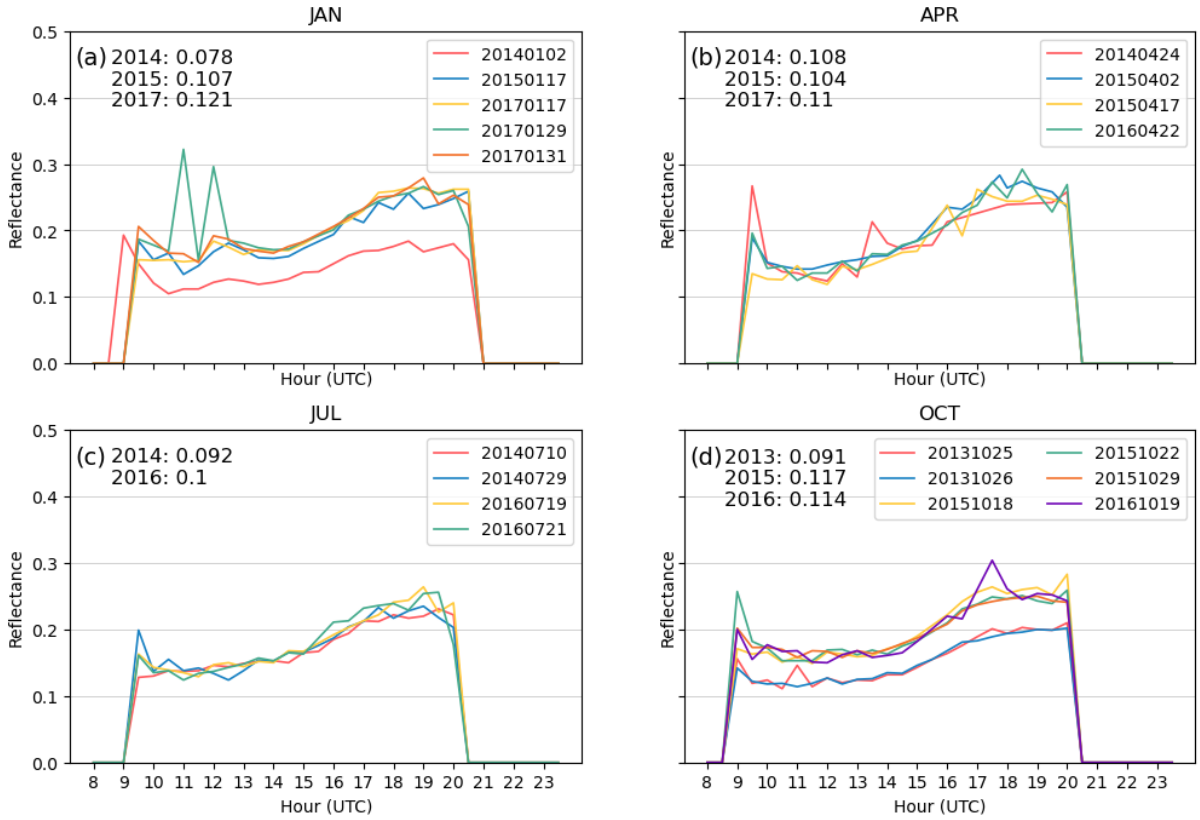
	MBE (h)		RMSE (h)		r	
Equatorial Region						
	CPTEC	CPTEC-Modified	CPTEC	CPTEC-Modified	CPTEC	CPTEC-Modified
JAN	1.42 (0.76)	1.07 (0.62)	2.23(0.59)	1.94 (0.44)	0.71 (0.09)	0.73 (0.09)
APR	1.28 (0.86)	0.92 (0.76)	2.2 (0.73)	1.98 (0.58)	0.65 (0.13)	0.65 (0.13)
JUL	1.47 (0.84)	0.81 (0.79)	2.17 (0.72)	1.75 (0.56)	0.58 (0.17)	0.61 (0.17)
OCT	1.16 (0.89)	0.75 (0.7)	2.12 (0.64)	1.84 (0.5)	0.61 (0.16)	0.62 (0.16)
Tropical Equatorial Region						
	CPTEC	CPTEC-Modified	CPTEC	CPTEC-Modified	CPTEC	CPTEC-Modified
JAN	-0.39 (1.16)	-0.6 (1.09)	1.97 (0.73)	1.9 (0.75)	0.76 (0.08)	0.78 (0.07)
APR	-0.28 (1.05)	-0.61 (0.98)	1.72 (0.59)	1.75 (0.53)	0.71 (0.09)	0.71 (0.09)
JUL	-0.12 (0.89)	-0.48 (0.76)	1.57 (0.45)	1.53 (0.43)	0.63 (0.17)	0.64 (0.17)
OCT	-0.86 (1.2)	-0.9 (0.99)	1.92 (0.57)	1.8 (0.48)	0.59 (0.15)	0.61 (0.15)
Tropical Northeast Oriental Region						
	CPTEC	CPTEC-Modified	CPTEC	CPTEC-Modified	CPTEC	CPTEC-Modified
JAN	-0.62 (1.17)	-0.79 (1.13)	2.26 (0.48)	2.21 (0.45)	0.61 (0.13)	0.65 (0.13)
APR	-0.73 (1.09)	-0.83 (1.0)	2.07 (0.4)	2.05 (0.4)	0.67 (0.1)	0.67 (0.1)
JUL	0.1 (0.91)	-0.13 (0.88)	1.84 (0.4)	1.8 (0.38)	0.72 (0.11)	0.73 (0.11)
OCT	-0.73 (1.68)	-0.85 (1.67)	2.55 (0.6)	2.51 (0.61)	0.52 (0.17)	0.54 (0.17)

To provide a closer understanding on the results of the modified product obtained for the Tropical Equatorial and Tropical Northeast Oriental regions, analysis of the diurnal cycle of the reflectance under clear sky conditions was carried out. Porfirio and Ceballos (2017), when evaluating the DNI estimates from the CPTEC model over Northeastern Brazil, noticed that consider a constant R_{min} value through the day, without accounting for diurnal variations, could lead to non-negligible errors.

To find clear sky days, the relative SDU was used. It is defined as the actual sunshine duration, and the maximum possible sunshine duration (ANIS et al., 2019). Values close to the unit represent days without or with a low frequency of clouds. So, for the stations localized within the mentioned regions, the time series of the relative SDU was generated, and days with relative SDU greater than 0.95 were selected. Thereafter, the diurnal cycle of the reflectance for each station was plotted.

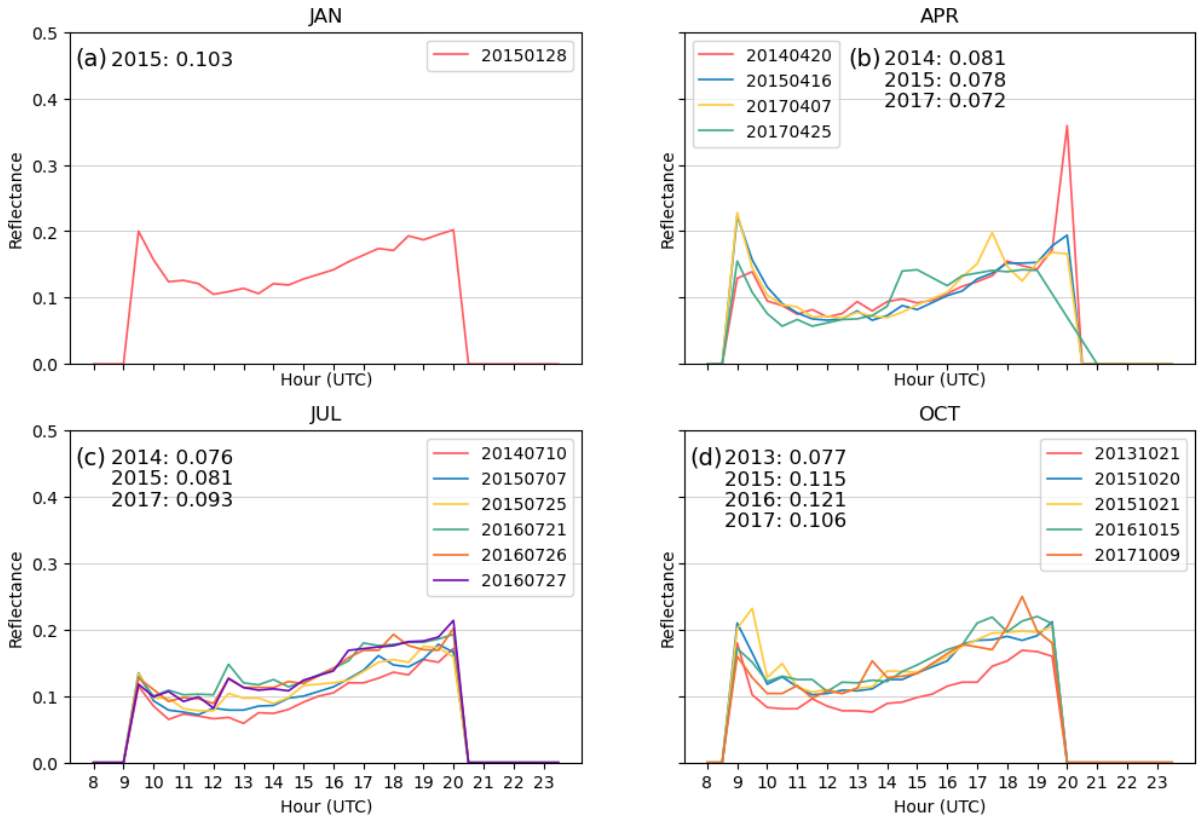
As examples, Figures 5.10 and 5.11 exhibit the diurnal cycle of reflectance for two stations (Petrolina-PE and Apodi-RN).

Figure 5.10 - Diurnal cycle of reflectance for Petrolina-PE (9.38°S, 40.48°W) station for (a) January, (b) April, (c) July and (d) October. Each line correponds to a day indicated in the legend. The R_{min} for each month and year for the pixel correspondent to the station site is indicated in each plot.



Firstly, is important to elucidate that not all plotted days consist of perfectly cloudless days. Despite of that, the results for both stations show lower reflectance values in the time interval between approximately 10 and 15 UTC, compared to other periods of the day. In fact, in some cases, the reflectance values close to the end of the day are almost twice the observed for the first half of the day, i.e., the plot for 26/10/2013 at Petrolina-PE, around 12 UTC the observed reflectance is about 0.11, however at 19 UTC it is close to 0.2. The amplitude of this difference, however, seems to be site and seasonal dependent.

Figure 5.11 - Diurnal cycle of reflectance for Apodi-RN (5.61°S, 37.81°W) station for (a) January, (b) April, (c) July and (d) October. Each line correponds to a day indicated in the legend. The R_{min} for each month and year for the pixel correspondent to the station site is indicated in each plot.



These findings may explain the results obtained for the modified product. The proper assessment of the R_{min} proved to be essential to provide good SDU estimates. Nonetheless, the diurnal variation of the clear sky reflectance was not considered in the SDU estimates. Previous studies demonstrated that the reflectance is influenced not only by surface properties and environmental factors, but also by the solar illumination angle in relation to the scene Kollenkark et al. (1982), Ranson et al. (1985), Deering and Eck (1987), and in some cases, it can present variations over 140% within a day.

The simplified approach used, i.e. R_{min} constant in a given day, may induce SDU underestimation, principally under clear sky conditions, as the minimum reflectance for some periods, even if it corresponds to clear sky, will not be as low as the values found around the first half of the day.

These results indicate that diurnal cycles of R_{min} on a seasonal basis could provide better SDU estimates. This proved to be useful by Porfirio and Ceballos (2017), that obtained improved DNI estimates by considering variable R_{min} values during the diurnal period.

Besides, although the reflectance diurnal cycle have been only examined in the Northeastern regions, similar results are also expected for other locations. Therefore, the contribution of using variable R_{min} values during the day, adjusted for the location and the season, to the improvement of the model, should be further investigated.

5.4 Summary

During the validation process, it is important to acknowledge the main source of errors and the limitations of the proposed method. After providing quantitative measures of the CPTEC model performance, this chapter intended to carefully address these features, and then give alternatives to overcome these issues.

Firstly, the lack of time constraints was investigated. It showed that the current method for accounting for the intervals that do not have observations leads to errors. For instance, in cases that the interval between the sunrise and the first valid image or/and the last image and the sunset is lower than one hour induces overestimates. This is more pronounced in locations with lower frequency of clouds along the day, which can even generate estimates higher than the maximum duration of the day. Regarding this topic, two possible solutions were suggested. Nonetheless, the potential of providing better estimates of each of the suggested equations must be investigated.

Section 5.2 examined the impact of the number of observations used in the estimations in the model performance. It was presented the absolute deviation per region for different amounts of images available for employment in the estimation. It could be observed that the estimates generated with fewer observations within a day (e.g. 5-10 images) exhibited higher absolute differences and larger spread, with exception of the Equatorial region, where the absolute difference does not show fluctuations related to the number of images. Therefore, it was suggested that this information be furnished with the estimates. This is a common practice in the remote sensing (NIGHTINGALE et al., 2018), as in most of cases the uncertainty of the estimation increases with decreasing number of available observations.

Last, the importance of the proper assessment of the R_{min} was analyzed. A new dataset with monthly R_{min} fields were generated for the central month of each season. The statistical parameters were obtained for the regions that the original product has shown its worse performance. Then, the results of the modified and the original product were compared. It could be seen that the modification of the R_{min} brought improvements for the Equatorial region, although has not exhibited a significant impact in the Tropical Equatorial and Tropical Northeast Oriental regions. The diurnal cycle of the reflectance indicated that to correctly consider the clear sky reflectance in the cloudiness parameterization, a variable R_{min} along the day must be considered beyond the location and season. Notwithstanding, the analysis of the diurnal cycle was performed only for the Northeastern Brazil, similar results are expected for other locations. Therefore the consideration of a variable R_{min} along the diurnal period, derived for a specific location and season should be extended to other regions, and the potential of this in improving the model's performance must be further examined.

6 MAIN CONCLUSIONS AND FUTURE WORK

This work has investigated the performance of the CPTEC model for SDU estimates over Brazil. To accomplish that, an analysis on the readily available SDU datasets was performed. After the proper selection of the reference dataset, the validation and inter-comparison of the CPTEC and CMSAF daily SDU estimates considering the period from September 2013 to December 2017 was carried out. Given the results, the main deficiencies of the CPTEC method were explored, and potential improvements were suggested. Section 6.1 gives a summary of the main results and conclusions of this work, regarding the specific objectives outlined in Chapter 1. In addition, Section 6.2 proposes future studies based on concerns that have been raised during this study.

6.1 Summary and conclusions

SDU is one of the most important parameters in climate monitoring (KOTHE et al., 2013). Given its relevance, long SDU time-series are accumulated at networks all over the world. Concerning the Brazilian territory, there are two networks that provide freely available relevant data for SDU studies: the INMET's network, which encompasses approximately 330 stations that assemble SDU records from CS recorders; and SONDA's network, that provides solar radiation data measured with automatic sensors at 20 stations. These data are further employed to derive SDU using the threshold defined by WMO, i.e. DNI $\geq 120\ W/m^2$ (WORLD METEOROLOGICAL ORGANIZATION - WMO, 2008). The commitment to maintain these networks is fundamental to deliver trustworthy information on SDU for climate studies. Notwithstanding, the distribution of these stations is uneven with regions, such as the Midwest and Northern Brazil, having very few stations. Besides, ground measurements are point-based observations, then to gather information in the vicinity, some authors proposed different methods of interpolation to provide information to locations with no records. Despite their efforts, unfortunately, the resultant data has high interpolations errors in regions with few stations available, being inadequate to characterize the SDU's high spatial variability.

Alternatively, methods based on remote sensing data have been proposed for SDU estimate. Ceballos and Rodrigues (2008) proposed a simplified method to estimate SDU leveraging the high spatial and temporal resolutions of GOES. The perfomance of the model was previously assessed by Porfirio (2012) for the Brazilian Northeast. The results indicated that the model has an overall good performance. Kothe et al. (2017) proposed a method based on the DNI achieved through Meteosat for

SDU estimation. The CMSAF product also delivers SDU estimates over Brazil, nonetheless, the accuracy of this model regarding the country was only assessed for the monthly sums. Therefore, the present study intended to extend the validation of CPTEC daily SDU estimates for the entire Brazilian territory, inter-compare its results with the ones obtained from the CMSAF method, and explore the model's deficiencies aiming to provide potential improvements.

Considering the primary requirements to select the reference dataset for the validation process presented in Chapter 3, the readily available data on SDU was gathered. For the analyzed period, applying the criteria proposed, INMET's network provided records for 194 stations, and SONDA's network presented solar radiation time-series at 10 stations. With the former dataset, quality flags are delivered as ancillary data, thus no additional quality control was employed. Regarding the INMET's records, some quality control procedures were utilized to discard gross errors. Afterwards, the measurements for locations with stations from both networks were compared. The results showed that the SDU records obtained from the SONDA's network for São Luis-MA and Natal-RN contained faulty data. Besides that, the results for Petrolina-PE were suspicious for the winter period. This indicated that although the automatic acquired data pass through quality checks, the resultant SDU dataset for the analyzed period may contain erroneous data. Hence, for the purpose of this work, the SONDA dataset was considered inappropriate for the present study, because the employment of faulty records could compromise the validation procedure. Therefore, to the following procedures, the measurements from INMET's network was used as reference. This dataset displayed more reliability than SONDA records, and also provided a greater amount of data, allowing a larger coverage of the country.

For the validation, the period from September 2013 to December 2017 was selected. It was chosen because it corresponds to the operation of the same platform (i.e. GOES-13) and provides a reasonable time interval for statistical analysis. The available stations were grouped into climate zones, this regionalization allowed a condensate and more robust assessment. The daily satellite-based SDU datasets, i.e. the CPTEC and CMSAF estimates, were evaluated against the reference measurements regarding the individual stations and the proposed regions, for each month. The results were analyzed by means of the MBE, the RMSE and the r for the climate zones, and the spatial distribution of the MBE.

The assessment showed that both products exhibit an overall good performance. The best results were obtained for the stations south of 15°S, encompassed by the

Tropical Central Brazil (Warm and Mesothermic) and Humid Temperate regions, where the CPTEC dataset exhibited lower MBE for the first, and the CMSAF for the latter. These stations presented MBE ranging from -1 to 1h for the individual assessment, and from -0.1 to 1h for the regional one. The Equatorial region also presented, in general, an overestimation tendency. For the CPTEC method, the regional MBE stayed between 1.16 and 1.61h, and for the CMSAF model, between 0.3 to 0.95h, indicating that for this region CMSAF method performs better.

The CPTEC product presented a tendency to underestimate the SDU records for the regions located in Northeast Brazil. The Tropical Equatorial region presented a mixed behavior, the stations located in Maranhão and Piaui states exhibited bias close to zero, the other stations presented negative bias for most of the year. The regional MBE ranged between -0.08 to -0.86h. The MBE for the Tropical Northeast Oriental region presented the highest bias, in magnitude, in March (-0.94h) and the lowest in July (0.1h). These two last regions presented a distinct pattern from the others: the bias exhibited a seasonal variation, approaching zero, towards the winter. This feature was attributed mainly to the cloudiness regime of these locations. Regarding these regions, the CMSAF estimates presented higher positive bias than for the aforementioned ones. Similar behavior was found by Kothe et al. (2017) for areas with frequent low warm clouds, such as the Brazilian Northeast. The authors credited these errors to the combined effect of misrepresentation by the constants used in the estimation and the systematically underestimation of the Effective Cloud Albedo by the Heliosat method in locations with persistent low cloud fields.

Concerning the RMSE and r results, as for the MBE, these parameters varied within regions. The RMSE for both satellite-derived datasets were similar, with exception for the Equatorial and Tropical Equatorial regions, in which RMSE of CPTEC estimates were higher than those from CMSAF. In general, the r indicated a strong positive correlation between SDU estimated and the reference dataset. The obtained results were in agreement with those previously reported in literature.

To further examine the results for the regions with the worst CPTEC model's performance, bivariate KDE plots were generated. The results suggested that for the Tropical Northeast Oriental region, the CPTEC estimates were close to the observations for days with low SDU values, which indicates that under partly cover and overcast sky, the CPTEC method performs better. Reassuring that for this region, the model's performance is related to the cloudiness regime. For the Equatorial region, the overestimation tendency were present under all sky conditions.

Given the results, some aspects of the CPTEC method were investigated, aiming to provide meaningful insights on potential improvements. Firstly, the treatment regarding the periods without observations, i.e. interval between sunrise-first image and last image-sunset, was analyzed. It could be seen, that the current model approach to account for these intervals may lead to errors. Which are more pronounced in sites with low cloud frequency within the day, in some cases generating estimates higher than the maximum duration of the day. Two simple suggestions are provided, although the considerations might yet lead to errors, adding time constraints will at least prevent estimates greater than the maximum day length. Quantitative advantages of the proposed approaches must be further explored.

The second aspect analyzed was the impact of the number of available images in the SDU estimates. It was shown that decreasing the number of available images, the uncertainty of the SDU product increases. In fact, this is frequent in remote sensing products. To circumvent this issue, a commonly employed measure is to inform data users of the amount of observations available for the estimate as quality flags, enabling judgment on the fitness-for-purpose of the data for their specific applications (NIGHTINGALE et al., 2018).

Lastly, an attempt to better represent the clear sky reflectance was performed. To ascertain its potential on improving the model, SDU estimates were generated for the central month of each season, using monthly R_{min} fields. These fields were achieved by taking the minimum (not null) reflectance values in the interval between 14 and 16 UTC, within the target month. The resultant SDU estimates were compared to the reference dataset for the regions that presented the worst results in the previous assessment of the model. The modification of R_{min} brought improvements for the Equatorial region, but did not presented a significant impact in the Tropical Equatorial and Tropical Northeast Oriental regions. Afterwards the diurnal cycle of R_{min} for days with predominantly clear sky was analyzed. The results suggested that to proper consider the clear sky reflectance in the cloudiness parameterization, a variable R_{min} within daytime must be considered beyond the location and season.

The results obtained in this study showed that the CPTEC method presents a good performance over Brazilian territory, being an reliable alternative to provide information to different applications. Notwithstanding, further investigation on the main deficiencies of the method must be carried out to provide further improvements.

6.2 Future work

Although the results of the CPTEC model's performance has shown an overall good performance, additional analysis must be accomplished:

Regarding the validation, periods correspondent to other GOES platforms should be evaluated and compared to the results obtained in the present work, to assure that the SDU time-series is sufficiently homogeneous and stable, which are fundamental requirements for application on climate research (GLOBAL CLIMATE OBSERVING SYSTEM, GCOS, 2016).

Given the high resolution of the CPTEC SDU estimates, generate annual and seasonal maps to update the previous assessed knowledge, and provide further characterization of its variability on regional and local scales.

Perform experiments implementing the proposed suggestions for the beginning and end of the day treatment to quantitatively assess the potential of these considerations in improving the model's performance.

Extend the analysis on the impact of the available observations in the estimation. For subsequently, on an operational basis, provide the user with quality flags based on these results.

Investigate the diurnal cycle of R_{min} for different regions of Brazil, and evaluate SDU estimates generated with variable R_{min} along the diurnal period, derived for a specific location and season, against ground measurements.

Considering the importance of reliable ground measurements for model's validation, this work showed that the employment of the BSRN criteria for quality check on the SONDA data was insufficient to assure the data quality. Regarding the data obtained through CS recorders, there is no widely accepted methodology for quality control. Therefore, it is suggested to conduct studies to develop rigorous criteria for quality control concerning the relevant data for SDU research.

REFERENCES

ABBOT, P. **Guidelines on the quality control of surface climatology data**. [S.l.]: WMO, 1986. 20

AKINOGLU, B. G. Recent advances in the relations between bright sunshine hours and solar irradiation. In: BADESCU, V. (Ed.). **Modeling solar radiation at the earth's surface**. [S.l.]: Springer, 2008. p. 115–143. 1, 35

ALMOROX, J.; ARNALDO, J.; BAILEK, N.; MARTí, P. Adjustment of the angstrom-prescott equation from campbell-stokes and kipp-zonen sunshine measures at different timescales in Spain. **Renewable Energy**, v. 154, p. 337–350, Jul 2020. Available from: <http://dx.doi.org/10.1016/j.renene.2020.03.023>. 14, 19

ANDERSON, M. C.; ZOLIN, C. A.; SENTELHAS, P. C.; HAIN, C. R.; SEMMENS, K.; YILMAZ, M. T.; GAO, F.; OTKIN, J. A.; TETRAULT, R. The evaporative stress index as an indicator of agricultural drought in Brazil: an assessment based on crop yield impacts. **Remote Sensing of Environment**, v. 174, p. 82–99, 2016. 14

ANGSTROM, A. Solar and terrestrial radiation. report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation. **Quarterly Journal of the Royal Meteorological Society**, v. 50, n. 210, p. 121–126, 1924. 1, 8

ANIS, M. S.; JAMIL, B.; ANSARI, M. A.; BELLOS, E. Generalized models for estimation of global solar radiation based on sunshine duration and detailed comparison with the existing: a case study for India. **Sustainable Energy Technologies and Assessments**, v. 31, p. 179–198, 2019. 72

BARTOSZEK, K. Actual sunshine duration in Poland–comparison of satellite and ground-based measurements. **Przegląd Geofizyczny**, 2018. 13

BASSO, T. M.; NOGUEIRA, C. E. C.; SILVA, D. S. Eficiência energética na construção civil no brasil. **Acta Iguazu**, v. 4, n. 1, p. 48–56, 2015. 14

_____. Eficiência energética na construção civil no Brasil. **Acta Iguazu**, v. 4, n. 1, p. 48–56, 2015. 35

BATTISTI, R. **Calibration, uncertainties and use of soybean crop simulation models for evaluating strategies to mitigate the effects of climate change in Southern Brazil**. Thesis (PhD in Agricultural Systems Engineering) — Universidade de São Paulo, Piracicaba, 2016. 14

BATTISTI, R.; BENDER, F.; SENTELHAS, P. C. Assessment of different gridded weather data for soybean yield simulations in Brazil. **Theoretical and Applied Climatology**, v. 135, n. 1, p. 237–247, 2019. 14

BAUMGARTNER, D.; PÖTZI, W.; FREISLICH, H.; STRUTZMANN, H.; VERONIG, A.; FOELSCHE, U.; RIEDER, H. A comparison of long-term parallel measurements of sunshine duration obtained with a campbell-stokes sunshine recorder and two automated sunshine sensors. **Theoretical and Applied Climatology**, v. 133, n. 1-2, p. 263–275, 2018. 27

BAUMGARTNER, T. **Die Schwellenintensität des Sonnenscheinautographen campbell-stokes an wolkenlosen Tagen**. [S.l.]: Schweizerische Meteorologische Zentralanstalt, 1979. 20, 26

BENEDITO-SILVA, A. A.; PIRES, M. L. N.; CALIL, H. M. Seasonal variation of suicide in Brazil. **Chronobiology International**, v. 24, n. 4, p. 727–737, 2007. 14

BERTRAND, C.; DEMAIN, C.; JOURNÉE, M. Estimating daily sunshine duration over Belgium by combination of station and satellite data. **Remote Sensing Letters**, v. 4, n. 8, p. 735–744, 2013. 1, 11

BLANC, P. et al. Direct normal irradiance related definitions and applications: the circumsolar issue. **Solar Energy**, v. 110, p. 561–577, 2014. 6

BOEING. **GOES N data book**. Boeing Satellite Development Center, 2006. Available from: <http://goes.gsfc.nasa.gov/text/goes.databookn.html>. 61

BURIOL, G. A.; ESTEFANEL, V.; HELDWEIN, A. B.; PRESTES, S. D.; HORN, J. F. C. Estimativa da radiação solar global a partir dos dados de insolação, para Santa Maria-RS. **Ciência Rural**, v. 42, n. 9, p. 1563–1567, 2012. 14

CANO, D.; MONGET, J.-M.; ALBUISSON, M.; GUILLARD, H.; REGAS, N.; WALD, L. A method for the determination of the global solar radiation from meteorological satellite data. **Solar Energy**, v. 37, n. 1, p. 31–39, 1986. 8

CAVALCANTI, I. F. **Tempo e clima no Brasil**. [S.l.]: Oficina de textos, 2009. 42

CEBALLOS, J. C.; BOTTINO, M. J.; SOUZA, J. M. D. A simplified physical model for assessing solar radiation over Brazil using goes 8 visible imagery. **Journal of Geophysical Research: Atmospheres**, v. 109, n. D2, 2004. 17, 36, 37

CEBALLOS, J. C.; RODRIGUES, M. L. Estimativa de insolaçao mediante satélite geoestacionário: resultados preliminares. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 15., 2008. **Proceedings...** [S.l.], 2008. 2, 16, 35, 36, 77

COSTA, S. et al. A successful practical experience with dedicated geostationary operational environmental satellites goes-10 and-12 supporting Brazil. **Bulletin of the American Meteorological Society**, v. 99, n. 1, p. 33–47, 2018. 36, 61

DEERING, D. W.; ECK, T. F. Atmospheric optical depth effects on angular anisotropy of plant canopy reflectance. **International Journal of Remote Sensing**, v. 8, n. 6, p. 893–916, 1987. 74

DIABATÉ, L.; MOUSSU, G.; WALD, L. Description of an operational tool for determining global solar radiation at ground using geostationary satellite images. **Solar Energy**, v. 42, n. 3, p. 201–207, 1989. 7

DUFFIE, J. A.; BECKMAN, W. A. **Solar engineering of thermal processes**. New York: John Wiley & Sons, 1991. 5

FENG, S.; HU, Q.; QIAN, W. Quality control of daily meteorological data in China, 1951–2000: a new dataset. **International Journal of Climatology**, v. 24, n. 7, p. 853–870, May 2004. Available from: <http://dx.doi.org/10.1002/joc.1047>. 20, 22

FUNARI, F. L. **Insolação, radiação solar global e radiação líquida no Brasil**. Dissertação (Mestrado) — Universidade de São Paulo (USP), São Paulo, 1983. 14, 51

FUNARI, F. L.; TARIFA, J. R. Insolação, radiação solar global e radiação líquida no Brasil. **Revista do Instituto Geológico**, v. 38, n. 2, p. 49–83, 2017. 14

GLOBAL CLIMATE OBSERVING SYSTEM, GCOS. **The Global Observing System for Climate: implementation needs**. [S.l.]: WMO, 2016. 81

GOOD, E. et al. Estimating daily sunshine duration over the UK from geostationary satellite data. **Weather**, v. 65, n. 12, p. 324–328, 2010. 2, 7, 9, 10, 35, 49

HAMMER, A.; HEINEMANN, D.; HOYER, C.; KUHLEMANN, R.; LORENZ, E.; MÜLLER, R.; BEYER, H. G. Solar energy assessment using remote sensing technologies. **Remote Sensing of Environment**, v. 86, n. 3, p. 423–432, 2003. 38

HANNAK, L.; KNIPPERTZ, P.; FINK, A. H.; KNIFFKA, A.; PANTE, G. Why do global climate models struggle to represent low-level clouds in the west african summer monsoon? **Journal of Climate**, v. 30, n. 5, p. 1665–1687, 2017. 49

HINSSEN, Y. B. L.; KNAP, W. H. Comparison of pyranometric and pyrheliometric methods for the determination of sunshine duration. **Journal of Atmospheric and Oceanic Technology**, v. 24, n. 5, p. 835–846, 2007. 6

HOLANDA, R. M. de; MEDEIROS, R. M. de; GONÇALVES, E. M. Flutuação da insolação e nebulosidade no município de Caruaru–PE, Brasil. In: WORKSHOP INTERNACIONAL SOBRE ÁGUA NO SEMIARIDO BRASILEIRO, 3., 2017. **Anais...** [S.l.], 2017. 47

HUANG, L.; JIANG, J. H.; WANG, Z.; SU, H.; DENG, M.; MASSIE, S. Climatology of cloud water content associated with different cloud types observed by a-train satellites. **Journal of Geophysical Research: Atmospheres**, v. 120, n. 9, p. 4196–4212, 2015. 49

HUANG, Y. L.; XIU, S. Y.; ZHONG, S. Q.; ZHENG, L.; SUN, H. Division of banana for climatic suitability based on a decision tree. **Journal of Tropical Meteorology,**, v. 28, p. 140–144, 2012. 1, 35

INSTITUTO NACIONAL DE METEOROLOGIA, INMET. **Mapas das estações**. Available from: <https://mapas.inmet.gov.br/>. Access in: 02 May 2021. 20

KANDIRMAZ, H. A model for the estimation of the daily global sunshine duration from meteorological geostationary satellite data. **International Journal of Remote Sensing**, v. 27, n. 22, p. 5061–5071, 2006. 2, 7, 10, 35, 49

KANDIRMAZ, H. M.; KABA, K. Estimation of daily sunshine duration from terra and aqua modis data. **Advances in Meteorology**, v. 2014, 2014. 1, 5, 7, 12

KANDIRMAZ, H. M.; YEGINGIL, L.; PESTEMALCI, V.; EMRAHOGLU, N. Daily global solar radiation mapping of Turkey using Meteosat satellite data. **International Journal of Remote Sensing**, v. 25, n. 11, p. 2159–2168, 2004. 7

KELLER, A.; FREDERIKSEN, P.; HäNDEL, M. N.; JACOBSEN, R.; MCGRATH, J. J.; COHEN, A. S.; HEITMANN, B. L. Environmental and individual predictors of 25-hydroxyvitamin d concentrations in Denmark measured from neonatal dried blood spots: the d-tect study. **British Journal of Nutrition**, v. 121, n. 5, p. 567–575, 2019. 1

KERR, A.; TABONY, R. Comparison of sunshine recorded by campbell–stokes and automatic sensors. **Weather**, v. 59, n. 4, p. 90–95, 2004. 6, 20, 32

KOLLENKARK, J.; VANDERBILT, V.; DAUGHTRY, C.; BAUER, M. Influence of solar illumination angle on soybean canopy reflectance. **Applied Optics**, v. 21, n. 7, p. 1179–1184, 1982. 74

KOTHE, S.; GOOD, E.; OBREGÓN, A.; AHRENS, B.; NITSCHE, H. Satellite-based sunshine duration for Europe. **Remote Sensing**, v. 5, n. 6, p. 2943–2972, 2013. 1, 10, 13, 38, 77

KOTHE, S.; PFEIFROTH, U.; CREMER, R.; TRENTMANN, J.; HOLLMANN, R. A satellite-based sunshine duration climate data record for Europe and Africa. **Remote Sensing**, v. 9, n. 5, p. 429, 2017. 2, 3, 12, 17, 35, 38, 39, 48, 77, 79

KOZMHINSKY, M.; MEDEIROS, R. M. de; HOLANDA, R. M. de; SILVA, V. de P. Average insolation interpolated by the krigagem method for the state of Pernambuco–Brazil. **Journal of Hyperspectral Remote Sensing**, v. 8, n. 2, p. 78–84, 2018. 16

LEGG, T. Comparison of daily sunshine duration recorded by campbell–stokes and kipp and zonen sensors. **Weather**, v. 69, n. 10, p. 264–267, 2014. 19, 32

LOEW, A. et al. Validation practices for satellite-based earth observation data across communities. **Reviews of Geophysics**, v. 55, n. 3, p. 779–817, Sep 2017. Available from: <http://dx.doi.org/10.1002/2017RG000562>. 19, 33, 35

LUBIN, D.; WEBER, P. G. The use of cloud reflectance functions with satellite data for surface radiation budget estimation. **Journal of Applied Meteorology**, v. 34, n. 6, p. 1333–1347, 1995. 36

LYRA, G. B.; ZANETTI, S. S.; SANTOS, A. A. R.; SOUZA, J. L. de; LYRA, G. B.; OLIVEIRA-JÚNIOR, J. F.; LEMES, M. A. M. Estimation of monthly global solar irradiation using the hargreaves–samani model and an artificial neural network for the state of alagoas in northeastern Brazil. **Theoretical and Applied Climatology**, v. 125, n. 3, p. 743–756, 2016. 14

MACHADO, L. A. et al. The chuva project: how does convection vary across Brazil? **Bulletin of the American Meteorological Society**, v. 95, n. 9, p. 1365–1380, 2014. 49

MATUSZKO, D. Influence of the extent and genera of cloud cover on solar radiation intensity. **International Journal of climatology**, v. 32, n. 15, p. 2403–2414, 2012. 32

MCGRATH, J.; SELTEN, J.-P.; CHANT, D. Long-term trends in sunshine duration and its association with schizophrenia birth rates and age at first registration—data from Australia and the Netherlands. **Schizophrenia research**, v. 54, n. 3, p. 199–212, 2002. 1

MEDEIROS, R. M. de; HOLANDA, R. M. de; FRANÇA, M. V. de. Interpolação da insolação média para o estado do Piauí–Brasil. **Revista de Geografia (Recife)**, v. 35, n. 5, 2018. 15

MOURA, A. R. C.; MONTENEGRO, S. M.; ANTONINO, A. C. D.; AZEVEDO, J. R. G. d.; SILVA, B. B. d.; OLIVEIRA, L. M. Evapotranspiração de referência baseada em métodos empíricos em bacia experimental no estado de Pernambuco-Brasil. **Revista Brasileira de Meteorologia**, v. 28, n. 2, p. 181–191, 2013. 14

MUELLER, R.; TRENTMANN, J.; TRÄGER-CHATTERJEE, C.; POSSELT, R.; STÖCKLI, R. The role of the effective cloud albedo for climate monitoring and analysis. **Remote Sensing**, v. 3, n. 11, p. 2305–2320, 2011. 38

NASTOS, P. T.; MATZARAKIS, A. Weather impacts on respiratory infections in Athens, Greece. **International Journal of Biometeorology**, v. 50, n. 6, p. 358–369, 2006. 1

NEW, M.; LISTER, D.; HULME, M.; MAKIN, I. A high-resolution data set of surface climate over global land areas. **Climate Research**, v. 21, n. 1, p. 1–25, 2002. 15

NIGHTINGALE, J. et al. Quality assurance framework development based on six new ecv data products to enhance user confidence for climate applications. **Remote Sensing**, v. 10, n. 8, p. 1254, 2018. 62, 75, 80

PACHECO, P. **An introduction to parallel programming**. [S.l.]: Elsevier, 2011. 58

PAINTER, H. E. The performance of a campbell-stokes sunshine recorder compared with a simultaneous record of the normal incidence irradiance. **The Meteorolgical Magazine**, v. 110, p. 102–109, 1981. 6, 20, 26, 30

PALHARINI, R. S. A.; VILA, D. A. Climatological behavior of precipitating clouds in the northeast region of Brazil. **Advances in Meteorology**, v. 2017, 2017. 42, 47

PERCH-NIELSEN, S. L. The vulnerability of beach tourism to climate change—an index approach. **Climatic Change**, v. 100, n. 3, p. 579–606, 2010. 14

PEREZ, R.; INEICHEN, P.; MOORE, K.; KMIECIK, M.; CHAIN, C.; GEORGE, R.; VIGNOLA, F. A new operational model for satellite-derived irradiances: description and validation. **Solar Energy**, v. 73, n. 5, p. 307–317, 2002. 64

PFEIFROTH, U.; KOTHE, S.; TRENTMANN, J.; HOLLMANN, R.; FUCHS, P.; KAISER, J.; WERSCHECK, M. **Surface Radiation Data Set - Heliosat (SARAH) - Edition 2.1**. [S.l.]: Satellite Application Facility on Climate Monitoring (CM SAF), 2019. 38

PORFIRIO, A.; CEBALLOS, J. A method for estimating direct normal irradiation from GOES geostationary satellite imagery: validation and application over Northeast Brazil. **Solar Energy**, v. 155, p. 178–190, 2017. 37, 61, 64, 72, 75

PORFIRIO, A.; CEBALLOS, J. C.; BRITTO, J.; COSTA, S. Evaluation of global solar irradiance estimates from gl1. 2 satellite-based model over brazil using an extended radiometric network. **Remote Sensing**, v. 12, n. 8, p. 1331, 2020. 52, 56, 57, 64

PORFIRIO, A. C. S. **Estimativa de irradiação solar direta normal mediante satélite: um estudo para o nordeste brasileiro**. Dissertação (Mestrado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais, São José do Campos, 2012. 3, 16, 46, 64, 77

PRESCOTT, J. Evaporation from a water surface in relation to solar radiation. **Transactions of the Royal Society of South Australia**, v. 46, p. 114–118, 1940. 8

RAICHIJK, C. Observed trends in sunshine duration over South America. **International Journal of Climatology**, v. 32, n. 5, p. 669–680, 2012. 40, 41, 42

RANSON, K. J.; BIEHL, L. L.; BAUER, M. E. Variation in spectral response of soybeans with respect to illumination, view and canopy geometry. **International Journal of Remote Sensing**, v. 6, n. 12, p. 1827–1842, 1985. 74

RAO, P. S.; SARASWATHYAMMA, C. K.; SETHURAJ, M. R. Studies on the relationship between yield and meteorological parameters of para rubber tree (hevea brasiliensis). **Agricultural and Forest Meteorology**, v. 90, n. 3, p. 235–245, 1998. 1, 35

REEK, T.; DOTY, S. R.; OWEN, T. W. A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. **Bulletin of the American Meteorological Society**, v. 73, n. 6, p. 753–765, 1992. 20

RIBEIRO, A. d. A.; ANDRADE JÚNIOR, A. S. d.; SILVA, E. M. d.; BASTOS, E. A.; SIMEÃO, M. Estimativa da radiação solar global a partir dos dados de insolação no estado do Piauí, Brasil. **Revista Brasileira de Agricultura Irrigada**, v. 12, n. 1, p. 2348–2356, 2018. 14

RIGOLLIER, C.; LEFÈVRE, M.; WALD, L. The method heliosat-2 for deriving shortwave solar radiation from satellite images. **Solar Energy**, v. 77, n. 2, p. 159–169, 2004. 64

SANCHEZ-LORENZO, A.; BRUNETTI, M.; CALBó, J.; MARTIN-VIDE, J. Recent spatial and temporal variability and trends of sunshine duration over the iberian peninsula from a homogenized data set. **Journal of Geophysical Research**, v. 112, n. D20, Oct 2007. Available from: <http://dx.doi.org/10.1029/2007JD008677>. 20

SANTOS, R. C.; SILVA, A. C. P.; SANTOS, M. J. d.; BARBOSA, M. R.; COIMBRA, D. G.; GITAÍ, D. L. G.; ANDRADE, T. G. de. Environmental temperature as a mediator on the association between photoperiod at birth and chronotype. **Chronobiology International**, v. 37, n. 11, p. 1662–1668, 2020. 14

SHAMIM, M. A.; REMESAN, R.; HAN, D.-w.; EJAZ, N.; ELAHI, A. An improved technique for global daily sunshine duration estimation using satellite imagery. **Journal of Zhejiang University Science A**, v. 13, n. 9, p. 717–722, 2012. 2, 10, 35, 49

SHAO, J. Calculation of sunshine duration and saving of land use in urban building design. **Energy and Buildings**, v. 15, n. 3-4, p. 407–415, 1990. 1, 35

SILVA, P. E. D. da; MARTINS, F. R.; PEREIRA, E. B. Quality control of solar radiation data within sonda network in Brazil: preliminary results. In: EUROSUN INTERNATIONAL CONFERENCE ON SOLAR ENERGY AND BUILDINGS, 2014. **Proceedings...** [S.l.], 2014. 13

SKARTVEIT, A.; OLSETH, J. A.; TUFT, M. E. An hourly diffuse fraction model with correction for variability and surface albedo. **Solar Energy**, v. 63, n. 3, p. 173–183, 1998. 38

SOUZA, J. L. D.; LYRA, G. B.; SANTOS, C. M. D.; FERREIRA JÚNIOR, R. A. F.; TIBA, C.; LYRA, G. B.; LEMES, M. A. M. Empirical models of daily and monthly global solar irradiation using sunshine duration for Alagoas state, northeastern Brazil. **Sustainable Energy Technologies and Assessments**, v. 14, p. 35–45, 2016. 14

STANHILL, G. Through a glass brightly: some new light on the campbell–stokes sunshine recorder. **Weather**, v. 58, n. 1, p. 3–11, 2003. 19, 64

TIBA, C.; FRAIDENRAICH, N.; LYRA, F.; NOGUEIRA, A. **Atlas solarimétrico do Brasil: banco de dados terrestres**. Recife: UFPE: [s.n.], 2000. 32 p. 15

TIBA, C.; GALLEGOS, H. G.; FRAIDENRAICH, N.; LYRA, F. Technical note on the development of spatial temporal solar radiation maps: a brazilian case study. **Renewable Energy**, v. 18, n. 3, p. 393–408, 1999. 15

URRACA, R.; GRACIA-AMILLO, A. M.; HULD, T.; PISON, F. J. Martinez-de; TRENTMANN, J.; LINDFORS, A. V.; RIIHELÄ, A.; SANZ-GARCIA, A. Quality control of global solar radiation data with satellite-based products. **Solar Energy**, v. 158, p. 49–62, 2017. 33

WANG, H.; LIU, D.; LIN, H.; MONTENEGRO, A.; ZHU, X. NDVI and vegetation phenology dynamics under the influence of sunshine duration on the tibetan plateau. **International Journal of Climatology**, v. 35, n. 5, p. 687–698, 2015. 1, 35

WARNANT, P.; FRANÇOIS, L.; STRIVAY, D.; GÉRARD, J.-C. CAB: a global model of terrestrial biological productivity. **Global Biogeochemical Cycles**, v. 8, n. 3, p. 255–270, 1994. 1, 35

WILKS, D. S. **Statistical methods in the atmospheric sciences**. [S.l.]: Academic Press, 2011. 42

WOOD, C. R.; HARRISON, R. G. Scorch marks from the sky. **Weather**, v. 66, n. 2, p. 39–41, 2011. 19

WORLD METEOROLOGICAL ORGANIZATION - WMO. **WMO guide to meteorological instruments and methods of observation.** [S.l.]: WMO, 2008. 1, 6, 19, 77

WU, B.; LIU, S.; ZHU, W.; YU, M.; YAN, N.; XING, Q. A method to estimate sunshine duration using cloud classification data from a geostationary meteorological satellite (FY-2D) over the Heihe River Basin. **Sensors**, v. 16, n. 11, p. 1859, 2016. 2, 7, 12, 16, 35

XAVIER, A. C.; KING, C. W.; SCANLON, B. R. Daily gridded meteorological variables in Brazil (1980–2013). **International Journal of Climatology**, v. 36, n. 6, p. 2644–2659, 2016. 13

XUE, L.; WANG, C.-c.; WANG, C.-l.; SHEN, S.-h. Agricultural climatic regionalization for Longyan cultivation in Guangdong province. **Journal of Tropical Meteorology**, v. 27, p. 403–409, 2011. 1, 35

ZHU, W.; WU, B.; YAN, N.; MA, Z.; WANG, L.; LIU, W.; XING, Q.; XU, J. Estimating sunshine duration using hourly total cloud amount data from a geostationary meteorological satellite. **Atmosphere**, v. 11, n. 1, p. 26, 2020. 2, 7, 35

APPENDIX A

Table A.1 - INMET's weather stations with less than 20% missing data during the analyzed period, i.e., September 2013 to December 2017.

No	Site	Latitude	Longitude	Altitude
	Equatorial			
82212	FONTE BOA - AM	-2.53	-66.16	55.57
82915	RIO BRANCO - AC	-9.96	-67.80	160.00
82336	ITACOATIARA - AM	-3.13	-58.43	40.0
82188	BREVES - PA	-1.68	-50.48	14.74
82145	TRACUATEUA - PA	-1.06	-46.9	36.00
82178	OBIDOS - PA	-1.91	-55.51	37.00
82353	ALTAMIRA - PA	-3.21	-52.21	74.04
83214	MATUPA - MT	-10.25	-54.91	285.00
82317	TEFE - AM	-3.83	-64.70	47.00
82326	CODAJAS - AM	-3.83	-62.08	48.00
82361	TUCURUI - PA	-3.76	-49.66	40.00
82067	IAUARETE - AM	0.61	-69.18	120.00
82562	MARABA - PA	-5.36	-49.13	95.00
82141	SOURE - PA	-0.73	-48.51	10.49
82113	BARCELOS - AM	-0.96	-62.91	40.00
82445	ITAITUBA - PA	-4.28	-55.99	45.00
82263	CAMETA - PA	-2.25	-49.50	23.90
82181	MONTE ALEGRE - PA	-2.00	-54.10	145.85
82246	BELTERRA - PA	-2.63	-54.95	175.74
82184	PORTO DE MOZ - PA	-1.73	-52.23	15.93
82861	CONCEICAO DO ARAGUAIA - PA	-8.26	-49.26	156.85
82610	EIRUNEPE - AM	-6.66	-69.86	104.00
82191	BELEM - PA	-1.43	-48.43	10.00
82331	MANAUS - AM	-3.10	-60.01	61.25
82098	MACAPA - AP	-0.05	-51.11	14.46
82024	BOA VISTA - RR	2.82	-60.66	83.00
82106	S G DA CACHOEIRA UAUPES - AM	-0.11	-67.00	90.00
	Tropical Equatorial			
82882	PAULISTANA - PI	-8.13	-41.13	374.22
82590	APODI - RN	-5.61	-37.81	150.00
82588	MORADA NOVA - CE	-5.11	-38.36	43.62
82480	PIRIPIRI - PI	-4.28	-41.78	161.12
82970	ALTO PARNAIBA - MA	-9.10	-45.93	285.05
82578	TERESINA - PI	-5.08	-42.81	74.36

(Continue)

Table A.1 - Continuation.

82791	PATOS - PB	-7.01	-37.26	249.09
82690	SERIDO CAICO - RN	-6.46	-37.08	169.85
82198	TURIACU - MA	-1.56	-45.36	44.06
82493	JAGUARUANA - CE	-4.78	-37.76	11.71
82298	ESPERANTINA - PI	-3.90	-42.25	87.05
82693	CRUZETA - RN	-6.43	-36.58	226.46
82686	IGUATU - CE	-6.36	-39.29	217.67
82382	CHAPADINHA - MA	-3.73	-43.35	103.50
82784	BARBALHA - CE	-7.31	-39.30	409.03
82487	GUARAMIRANGA - CE	-4.26	-39.00	870.67
82287	PARNAIBA - PI	-3.08	-41.76	79.50
82768	BALSAS - MA	-7.53	-46.03	259.38
82296	LUZILANDIA - PI	-3.41	-42.28	49.00
82586	QUIXERAMOBIM - CE	-5.16	-39.28	79.50
82564	IMPERATRIZ - MA	-5.53	-47.48	123.30
82583	CRATEUS - CE	-5.16	-40.66	296.82
82780	PICOS - PI	-7.03	-41.48	207.93
82376	ZE DOCA - MA	-3.26	-45.65	45.28
82571	BARRA DO CORDA - MA	-5.50	-45.23	153.00
82280	SAO LUIS - MA	-2.53	-44.21	50.86
82879	SAO JOAO DO PIAUI - PI	-8.35	-42.25	235.33
82659	ARAGUAINA - TO	-7.20	-48.20	228.52
82886	CABROBO - PE	-8.51	-39.33	341.46
82689	SAO GONCALO - PB	-6.75	-38.21	233.06
82789	TRIUNFO - PE	-7.81	-38.11	1105.0
82777	CAMPOS SALES - CE	-7.00	-40.38	583.50
82870	VALE DO GURGUEIA - PI	-8.41	-43.71	265.00
82983	PETROLINA - PE	-9.38	-40.48	370.46
82392	SOBRAL - CE	-3.73	-40.33	109.62
82678	FLORIANO - PI	-6.76	-43.01	123.27
82753	OURICURI - PE	-7.90	-40.04	459.28
82460	BACABAL - MA	-4.22	-44.76	25.07
82676	COLINAS - MA	-6.03	-44.25	179.75
82397	FORTALEZA - CE	-3.81	-38.53	26.45
82863	PEDRO AFONSO - TO	-8.96	-48.18	187.00
82765	CAROLINA - MA	-7.33	-47.46	192.83
82594	MACAU - RN	-5.15	-36.57	17.34
	Tropical Northeast Oriental			

(Continue)

Table A.1 - Continuation.

82596	CEARA MIRIM - RN	-5.65	-35.65	61.35
83096	ARACAJU - SE	-10.95	-37.04	4.72
82900	RECIFE CURADO - PE	-8.05	-34.95	10.00
83192	CIPO - BA	-11.08	-38.51	145.31
82795	CAMPINA GRANDE - PB	-7.22	-35.88	547.56
82890	ARCOVERDE - PE	-8.41	-37.08	680.70
82797	SURUBIM - PE	-7.83	-35.71	418.32
83221	FEIRA DE SANTANA - BA	-12.18	-38.96	230.68
82798	JOAO PESSOA - PB	-7.10	-34.86	7.43
83097	PROPRIA - SE	-10.21	-36.84	19.92
82989	AGUA BRANCA - AL	-9.28	-37.90	605.34
83229	SALVADOR ONDINA - BA	-13.01	-38.53	51.41
83195	ITABAIANINHA - SE	-11.11	-37.81	208.00
82990	PAO DE ACUCAR - AL	-9.75	-37.43	19.10
82893	GARANHUNS - PE	-8.88	-36.51	822.76
82696	AREIA - PB	-6.97	-35.68	574.62
82996	PORTO DE PEDRAS - AL	-9.18	-35.43	50.02
82992	PALMEIRA DOS INDIOS - AL	-9.44	-36.70	274.90
83398	CANAVIEIRAS - BA	-15.66	-38.95	3.87
82598	NATAL - RN	-5.91	-35.20	48.60
82994	MACEIO - AL	-9.55	-35.77	84.12
	Tropical Central Brazil – Warm			
83179	BARRA - BA	-11.08	-43.16	401.58
83364	PADRE RICARDO REMETTER - MT	-15.78	-56.06	140.00
83374	GOIAS - GO	-15.91	-50.13	512.22
83244	ITABERABA - BA	-12.51	-40.28	249.89
83377	BRASILIA - DF	-15.78	-47.92	1159.54
83286	CORRENTINA - BA	-13.33	-44.61	549.47
83441	SALINAS - MG	-16.15	-42.28	471.32
83379	FORMOSA - GO	-15.54	-47.33	935.19
83076	STa R DE CASSIA IBIPETUBA - BA	-11.01	-44.51	450.30
83190	SERRINHA - BA	-11.63	-38.96	359.63
83648	VITORIA - ES	-20.31	-40.31	36.20
83483	PIRAPORA - MG	-17.35	-44.91	505.24
83236	BARREIRAS - BA	-12.15	-45.00	439.29
83270	CANARANA - MT	-13.47	-52.27	430.00
83033	PALMAS - TO	-10.19	-48.30	280.00
83064	PORTO NACIONAL - TO	-10.71	-48.41	239.20

(Continue)

Table A.1 - Continuation.

83339	CAETITE - BA	-14.06	-42.48	882.47
83319	NOVA XAVANTINA - MT	-14.70	-52.35	316.00
83338	ESPINOSA - MG	-14.91	-42.80	569.64
83358	POXOREO - MT	-15.83	-54.38	450.00
83446	GUARATINGA - BA	-16.73	-39.54	194.67
83361	CUIABA - MT	-15.61	-56.10	145.00
83479	PARACATU - MG	-17.24	-46.88	712.00
83186	JACOBINA - BA	-11.18	-40.46	484.74
83368	ARAGARCAS - GO	-15.90	-52.23	345.00
83228	PEIXE - TO	-12.01	-48.35	242.49
83550	SAO MATEUS - ES	-18.70	-39.85	25.04
83182	IRECE - BA	-11.30	-41.86	747.16
83513	NHUMIRIM - MS	-18.98	-56.62	89.00
83242	LENCOIS - BA	-12.56	-41.38	438.74
83464	JATAI - GO	-17.91	-51.71	662.86
83437	MONTES CLAROS - MG	-16.68	-43.84	652.00
83452	JURAMENTO - MG	-16.77	-43.66	648.00
83334	FORMOSO - MG	-14.93	-46.25	840.00
83090	MONTE SANTO - BA	-10.43	-39.29	464.60
83393	PEDRA AZUL - MG	-16.00	-41.28	648.91
83565	PARANAIBA - MS	-19.75	-51.18	331.25
83344	VITORIA DA CONQUISTA - BA	-14.88	-40.79	874.81
83386	JANUARIA - MG	-15.45	-44.00	473.71
83376	PIRENOPOLIS - GO	-15.85	-48.96	740.00
82979	REMANSO - BA	-9.63	-42.10	400.51
83288	BOM JESUS DA LAPA - BA	-13.26	-43.41	439.96
83470	RIO VERDE - GO	-17.80	-50.91	774.62
83292	ITUACU - BA	-13.81	-41.30	531.43
83184	MORRO DO CHAPEU - BA	-11.21	-41.21	1003.27
83492	TEOFILO OTONI - MG	-17.86	-41.51	349.11
83363	SAO VICENTE - MG	-15.82	-55.42	786.99
83114	IGUABA GRANDE - RJ	-22.85	-42.19	5.57
82975	BOM JESUS DO PIAUI - PI	-9.07	-44.37	330.60
	Tropical Central Brazil – Mesothermic			
83538	DIAMANTINA - MG	-18.23	-43.64	1296.12
83669	SAO SIMAO - SP	-21.48	-47.55	617.39
83531	PATOS DE MINAS - MG	-18.51	-46.43	940.28
83579	ARAXA - MG	-19.60	-46.94	1023.61

(Continue)

Table A.1 - Continuation.

83851	SOROCABA - SP	-23.48	-47.43	645.00
83577	UBERABA - MG	-19.73	-47.95	737.00
83738	RESENDE - RJ	-22.45	-44.44	439.89
83639	CAPARAO - MG	-20.51	-41.90	843.18
83630	FRANCA - SP	-20.58	-47.36	1026.20
83736	SAO LOURENCO - MG	-22.10	-45.01	953.20
83676	CATANDUVA - SP	-21.11	-48.93	570.00
83587	BELO HORIZONTE - MG	-19.93	-43.93	915.00
83687	LAVRAS - MG	-21.75	-45.00	918.84
83781	SAO PAULO MIR de SANTANA - SP	-23.50	-46.61	792.06
83592	CARATINGA - MG	-19.73	-42.13	609.65
83767	MARINGA - PR	-23.40	-51.91	542.00
83695	ITAPERUNA - RJ	-21.20	-41.90	123.59
83692	JUIZ DE FORA - MG	-21.76	-43.36	939.96
83766	LONDRINA - PR	-23.31	-51.13	566.00
83718	CORDEIRO - RJ	-22.02	-42.36	505.92
83514	CAPINOPOLIS - MG	-18.71	-49.55	620.60
83689	BARBACENA - MG	-21.25	-43.76	1126.00
83683	MACHADO - MG	-21.68	-45.94	873.35
83726	SAO CARLOS - SP	-21.96	-47.86	856.00
83635	DIVINOPOLIS - MG	-20.17	-44.87	788.35
83642	VICOSA - MG	-20.76	-42.86	712.20
83582	BAMBUI - MG	-20.03	-45.00	661.27
83533	BOM DESPACHO - MG	-19.68	-45.36	695.00
83557	BOA ESPERANÇA - MG	-18.54	-40.27	128.95
	Humid Temperate			
83927	URUGUAIANA - RS	-29.75	-57.08	62.31
83811	IVAI - PR	-25.00	-50.86	808.00
83948	TORRES - RS	-29.35	-49.73	4.66
83872	INDAIAL - SC	-26.90	-49.21	86.13
83897	FLORIANOPOLIS - SC	-27.58	-48.56	1.84
83914	PASSO FUNDO - RS	-28.21	-52.40	684.05
83985	PELOTAS - RS	-31.78	-52.41	13.00
83919	BOM JESUS - RS	-28.66	-50.43	1047.50
83836	IRATI - PR	-25.46	-50.63	836.95
83923	URUSSANGA - SC	-28.51	-49.31	48.17
83907	SAO LUIZ GONZAGA - RS	-28.40	-55.01	245.11
83936	SANTA MARIA - RS	-29.70	-53.70	95.00

(Continue)

Table A.1 - Continuation.

83891	LAGES - SC	-27.81	-50.33	936.83
83967	PORTO ALEGRE - RS	-30.05	-51.16	46.97
83980	BAGE - RS	-31.33	-54.10	242.31
83842	CURITIBA - PR	-25.43	-49.26	923.50
83997	SANTA VITORIA DO PALMAR - RS	-33.51	-53.35	24.01
83942	CAXIAS DO SUL - RS	-29.16	-51.20	759.60
83912	CRUZ ALTA - RS	-28.63	-53.60	472.50
83920	SAO JOAQUIM - SC	-28.30	-49.93	1415.00
83813	CASTRO - PR	-24.78	-50.00	1008.80
83964	ENCRUZILHADA DO SUL - RS	-30.53	-52.51	427.75
83916	LAGOA VERMELHA - RS	-28.21	-51.50	840.00
83887	CAMPOS NOVOS - SC	-27.38	-51.20	946.67
83883	CHAPECO - SC	-27.11	-52.61	679.01

(Conclusion.)

Table A.2 - SONDA's stations with less than 20% missing data during the analyzed period, i.e., September 2013 to December 2017.

ID	Site	Latitude	Longitude	Altitude	Instrument
10	Brasília - DF	-15.6008	-47.7131	1023	Pyrheliometer
13	Cachoeira Paulista - SP	-22.6896	-45.0062	574	Pyrheliometer
01	Florianópolis - SC	-27.6017	-48.5178	31	Pyrheliometer
04	Joinville - SC	-26.2525	-48.8578	48	Pyranometers
17	Natal - RN	-5.8397	-35.2064	58	Pyranometers
19	Palmas - TO	-10.1778	-48.3619	216	Pyranometers
11	Petrolina - PE	-9.0689	-40.3197	387	Pyrheliometer
16	São Luiz - MA	-2.5933	-44.2122	40	Pyranometers
08	São Martinho da Serra - RS	-29.4428	-53.8231	489	Pyrheliometer
05	Sombrio - SC	-26.0956	-49.8133	15	Pyranometers