



Fakultät V - Mathematik und Naturwissenschaften
Institut für Physik

Master Thesis

Improved Heliosat4 cloud parallax and
shadow corrections using gridded cloud top
height data

Luis Guillermo Gomez Bobadilla

15. November 2024

Reviewers:
Prof. Dr. Phillip Huke
Dr. Jorge Enrique Lezaca Galeano

Acknowledgement

I would like to express my deepest gratitude to all those who have supported me throughout the journey of completing this master's thesis.

First and foremost, to my family, their love and support have been my foundation. To my daughter Sontje, who is the engine that pushes me forward and shines my whole world, and to my wife Franzi, thank you for your patience, understanding, and constant encouragement. You have always believed in me and provided the emotional strength necessary to persevere through this challenging journey.

I am immensely grateful to my colleague and mentor Dr. Jorge Lezaca. Your counsel, encouragement, and constantly teaching me new stuff every time we sit down and work together, have been crucial in allowing me to become a better scientist and researcher. I truly appreciate the support and the opportunities you provided that have enriched my learning experience.

I would also like to extend my heartfelt thanks to my leader Dr. Marion Schroedter-Homscheidt and Prof. Dr. Philipp Huke. Your guidance, expertise, and unwavering support have been invaluable. Thank you for your insightful feedback, encouragement, and for pushing me to strive for excellence. Your mentorship has been instrumental in shaping this research.

I would also like to thank everyone else in the EnMet team, who have contributed immensely to the completion of this thesis. I am honored to be part of such an awesome team.

The in-situ database used in this work was developed by Raphaël Jolivet and Yves Marie Saint-Drenan, from the research center O.I.E of Mines-Paristech. Special thanks and recognition for the amazing work in gathering and building the database from which the reference data was obtained.

Thank you all.

Statement

I declare that I have written this thesis independently and that I have not used any sources other than the ones sources and aids.

Luis Guillermo Gomez Bobadilla
Matrikelnummer 6412527
Oldenburg, den 15. November 2024

Abstract

The CAMS radiation model utilizes images from geostationary meteorological satellites to assess cloud presence and cloud properties to estimate solar irradiance at the Earth's surface. Previous studies have investigated the effects of satellite viewing parallax and shadow effects of the cloud position, but a correction for such effects has not yet been implemented on the current CAMS model. The parallax and shadow displacements are based on cloud top height, solar and satellite viewing angles. This study focuses on an improved geometrical correction for a future implementation of these effects on the current model to obtain an improved estimation of Global Horizontal Irradiance (GHI) using the Heliosat-4 Method. These adjustments were estimated by: (i) retrieving pixel-wise satellite estimated cloud top temperature data to calculate cloud top height (CTH) using atmospheric temperature profiles. (ii) Using CTH on geolocated cloudy pixels and viewing angles of each pixel to estimate the ground displacement from the parallax effect. (iii) The geolocation of the pixels for each of the displacements are found using haversine formulae and solar angles of those locations are determined to calculate the shadow displacement. (iv) Cloud Optical Thickness (COT), Cloud Coverage, and Cloud Type of the displaced pixels are shifted to new geolocated points and a bilinear interpolation is performed to estimate the values of those parameters on the original gridded map. (v) Global horizontal irradiance (GHI) of pixels where ground meteorological stations are available is estimated using the Heliosat-4 method. The corrected GHI values found were compared against solar surface irradiance (SSI) ground observation data and CAMS uncorrected data. Results showed an overall 0,71% MBE, 1,61% MAE and 2,55% RMSE reduction, when validated against the current model. The parallax/shadow adjusted irradiation estimates enable better solar resource retrieval. This demonstrates the potential of the implementation of the correction method with the Heliosat method to better estimate SSI and advance solar modeling capabilities. Further refinements to the parallax correction approach and model parametrization are discussed as avenues for future research.

Keywords: satellite viewing parallax, cloud shadow projection, cloud top height (CTH), cloud optical thickness (COT), global horizontal irradiance (GHI), solar surface irradiance (SSI), Heliosat-4 method

Contents

1. Introduction	1
2. The Heliosat Method	5
3. Improved Geometrical correction	9
3.1. Description of the improved geometrical correction	9
3.2. Calculation of Cloud Height	11
3.2.1. Using ceilometers	12
3.2.2. Using stereo-imaging	12
3.2.3. Using atmospheric profiles	13
4. Data	16
4.1. APOLLO_NG Data	16
4.2. Satellite Data from CM-SAF	19
4.3. Differences between APOLLO_NG and CM-SAF Datasets	20
4.4. Cloud Height	24
4.4.1. Approach 1: Using Cloud Top Temperature Data from APOLLO_NG (APOCTT)	24
4.4.2. Approach 2: Using Cloud Top Temperature Data from CM-SAF (SAFCTT)	25
4.4.3. Approach 3: Using Cloud Top Height Data directly from APOLLO_NG (SAFCTH)	25
4.5. SSI Ground observations	26
5. Results and Analysis	29
5.1. Evaluation Methodology	29
5.1.1. Absolute metrics	29
5.1.2. Relative metrics	30
5.2. Overall relative error metrics	32
5.3. Relative error metrics over shorter time periods	35
5.4. Correction performance	37
5.4.1. Improved Stations	38
5.4.2. Partially Improved Stations	40
5.4.3. Degraded Stations	41
5.4.4. General performance of the correction methods	44
5.5. Possible reasons for the lack of improvement	45
5.6. Final remarks	49
5.6.1. Key Findings	49
5.6.2. Concluding Thoughts	50
6. References	51
A. Haversine formula	54

B. Implementation	55
B.1. Data collection	55
B.2. Data processing	56
B.3. Application of Heliosat-4	56
B.4. Validation	56
C. Overall absolute error metrics	58
D. Seasonal Absolute Error Metrics	61

List of Figures

1.1. 2D Representation of cloud position	3
2.1. Schematic structure of Heliosat-4 Method [image taken from Qu et al., 2017]	6
3.1. Schematic representation of the relative geometry sun-pixel position-satellite	10
3.2. Ceilometer Lufft CHM 15K [image taken from DLR Eye2Sky network] . .	12
3.3. Cloud camera Mobotix Q25 with its own ventilation and heating system [image taken from DLR Eye2Sky network]	13
3.4. ISA Atmospheric Temperature Profile	14
3.5. AFGL and McClatchey Atmospheric Temperature Profiles	15
4.1. Example COT data [timestamp: '2021-10-21T11:00']	18
4.2. Example CTT data with non-physical values encountered [timestamp: '2021-07-09T15:00']	19
4.3. CTT Datasets comparison between APOLLO_NG and CM-SAF [timestamp: '2021-04-12T13:00']	21
4.4. Probability Distribution of CTT values	22
4.5. Time series comparison between CTT datasets	23
4.6. Overview of the different correction approaches	24
4.7. Selected stations for the implementation	27
4.8. Example QC on BSRN-BUD station [time range: '2021-01-01 - 2022-01-01']	28
5.1. Relative error comparison for 2021 data	34
5.2. Accumulative Relative Error Metrics per Season	36
5.3. Probability distribution of Cloud Type across seasons	37
5.4. Locations of best-performing stations	38
5.5. Percentage improvement for 'Improved' stations	39
5.6. Locations of 'Partially Improved' Stations	40
5.7. Percentage improvement for OUJ, UNZ and PAL stations	41
5.8. Locations of 'degraded' stations	42
5.9. Percentage improvement for degraded stations	43
5.10. Locations of FLO and RUN stations	44
5.11. Schematic representation of limitations of the geometrical correction . . .	46
5.12. Total parallax and shadow displacements for Sonnblick Station [Times-tamp: 2021-10-12T15:00]	46
5.13. Topographic map of BUD station	47
5.14. Topographic map of CNR station	47
5.15. Topographic map of PAY station	48
5.16. Topographic map of SON station	48
B.1. Overview of the implementation algorithm	57

C.1. GHI absolute error comparison for 2021 data	60
D.1. Seasonal Absolute error metrics	61

List of Tables

- 4.1. Data sources and resolutions for various cloud and irradiance measurements 16
- 4.2. Meteorological Stations retrieved from THREDDS Server 26
- 5.1. Relative Metrics for 2021 Data 33
- 5.2. Average performance metrics for all correction methods (full data) 44
- C.1. Absolute Metrics for 2021 Data 59

1. Introduction

Accurate modeling of solar irradiation is critical for applications in weather forecasting, climate research, and solar energy systems. It is essential for resource assessments, contributing for more precise yield calculations and planning. It also helps in the management of electricity grids, energy trade, and the solar power plants operations. It significantly impacts energy systems performance, affecting both photovoltaic systems and solar thermal power plants. Moreover, the integration of solar radiation data into energy systems is very important, as it provides information that is more that ever vital for the optimum operation of the energy grid.

The CAMS Radiation Service (CRS) offers historical values of Solar Surface Irradiance, also known as downwelling shortwave. Development of these applications is necessary to meet the needs of European and national policy developments and requirements of commercial and non-commercial downstream services. These services include planning, monitoring, efficiency enhancements, and integrating solar energy systems into energy grids (Copernicus Atmosphere Monitoring Service, 2024).

Irradiation at the surface is influenced by various atmospheric factors. It is mainly affected by cloud formation, but to a fewer extent, also by aerosol optical properties, as well as concentrations of water vapor and ozone. Other elements like ground albedo and elevation alter irradiation on the ground. CRS uses satellite images to compute Solar Surface Irradiance (SSI) based on data of aerosols, ozone, and water vapor taken from the CAMS global forecasting system. It is currently available for the field-of-view of the Meteosat Second Generation (MSG) satellite. This area encompasses Europe, Africa, the Atlantic Ocean, and the Middle East, spanning from -66° to 66° in both latitude and longitude. In the latest version (v4.6), the cams radiation model has included the field-of-view of the HIMAWARI satellite (West Asia and Australia).

The complex nature of clouds is commonly the most influential factor for estimating SSI (Cebecauer et al., 2011). Clouds have complex shapes even without going into integral length scales of the dynamics acting on them. In fact, there is a strict restriction because of the pixel size of the image, namely the image spatial resolution, even though the latest generation of geostationary satellite can go under-1-kilometer resolutions. The complex geometry of the cloud is determined by the sky conditions, that is: the amount of water vapor and other elements in the atmosphere, the local temperature as well as the weather conditions in the given location and time-of-year, the wind speed that carries the cloud through the sky, and in general the atmospheric fluid dynamics acting on the cloud.

There is already a big enough challenge when considering the cloud as an object with depth/thickness. Nevertheless, to attempt to model the cloud in a 2-dimensional image, CRS has developed cloud detection algorithms for satellite imaging that use multi-spectral data to identify and classify cloud pixels. Due to their unique spectral characteristics, clouds differ from other surface features. Cloud pixels often exhibit high reflectance in visible wavelengths and low thermal emission in thermal wavelengths.

In essence, the reflectivity detected by the satellite sensors is used to classify pixels as cloudy or cloud-free. So a cloud mask is generated to indicate the presence and extent of clouds within the image (Dagestad, 2005). A cloud mask can be as simple as a binary value (0 for clear sky, 1 for cloud), used i.e. in the APOLLO Scheme, but a more optimal approach is used in the APOLLO_NG cloud retrieval scheme, which replaces the previously used APOLLO scheme and introduces a probabilistic cloud mask instead of binary cloud masking. (Schroedter-Homscheidt et al., 2022).

Using radiative transfer models and algorithms that utilize measurements from multiple spectral bands, including visible and near-infrared wavelengths, cloud optical depth (COD) can also be determined. COD is a measure of the capacity of the cloud to absorb and scatter solar radiation. Simply put: the COD indicates how opaque the cloud is, determining the extent in which the cloud prevents sunlight to pass through it.

Satellite-derived data can also provide the cloud top temperature. This will become important later when finding a way of determine the cloud top height. But for now, the satellite itself is not capable of scanning the complete geometry or determine the cloud height, relative to the elevation of the actual position of the cloud.

Without going in too much detail the image is a discreet visualization of a 3D sky in a 2D map. So it registers the distorted outline of the clouds on the ground projection. The shape of that outline depends on the distance from the satellite and the cloud, the viewing angle of the same, the cloud height and surface of the ground in which the cloud is projected. All of this and with the help of Figure 1.1, it is clear that the actual position of the cloud is different from the apparent position of the cloud in the image.

Furthermore, the cloud's actual position relative to the sun at the moment when the image is taken, projects a shadow which is not located in apparent position or actual position. This is only True if the sun is not directly at the zenith of the cloud or directly behind the satellite, in which case the shadow location would be exactly the cloud's actual or apparent position respectively.

This raises an important point in both how the projection of the cloud and its shadow is distorted in the satellite image: the further the cloud is from the satellite, i.e. the sub satellite point, the more it will distort the projection. Moreover, the relative position of the sun with relation to the cloud will distort the shadow more or less. This of course, if the sun is not directly at its zenith. If it were, there would be no distortion. When the cloud is farther away from the sun in the horizon, i.e. at sunrise or sunset, the cloud's shadow would be projected further away.

The presence of a given cloud naturally affects the irradiance over the shaded surface. In particular, the shadow of the given cloud has a negative impact on the direct SSI received. Its presence has also an influence on the diffused irradiance because of scattering / refraction effects. So to accurately calculate the SSI of any given location using the Heliosat method a correction of the cloud's position in real-time is required.

The CRS model primarily relies on two main tools: APOLLO_NG and the Heliosat Method. The APOLLO (AVHRR Processing Over Clouds, Land and Ocean) is a tool used for cloud detection and cloud property retrieval. The APOLLO_NG (APOLLO Next Generation) builds upon the physical principles and incorporates advanced cloud detection algorithms, offering more accurate cloud masking, improved identification of cloud types, and better quantification of cloud properties (such as cloud optical depth (COD), temperature, and phase) (Klüser et al., 2015). The Heliosat method is then

used in order to estimate SSI under different sky conditions, namely clear or cloudy skies by using the cloud information obtained from APOLLO_NG. In essence, it uses the detected clouds on satellite images and their optical properties to estimate shortwave irradiance at ground level. It takes into account different atmospheric conditions and satellite sensor characteristics alongside the physical variables, e.g. solar and viewing (satellite) zenith and azimuth angles.

The current CRS model does not include the effects of parallax and shadow of the cloud position. These parallax and shadow effects have an influence on determining accurately the irradiance (Cebecauer et al., 2011, Perez et al., 2010). Before exploring the methods used to address this issue, it is important to understand the nature of the problem. Satellites capture raw digital counts of incoming radiation in a 2-dimensional grid represented by pixels. The registered clouds within these grid are merely the projections of them as seen by the observer, i.e. the satellite. This projection causes a parallax effect: the cloud appears in a different position in the image from where it truly is in the atmosphere. The position of a cloud is known by knowing the viewing angles relative to the satellite. In addition to this, the position of the same cloud relative to the sun projects a shadow on a different position. This is better illustrated in Figure 1.1.

To visualize this, imagine a cloud captured in a satellite image, projected onto the ground as a 2D image. The projection, shown in Figure 1.1, depends on the satellite's viewing angle and the cloud's actual position in the sky. Without any additional information, such as cloud elevation data, the cloud's apparent position differs from its actual position in 3D space.

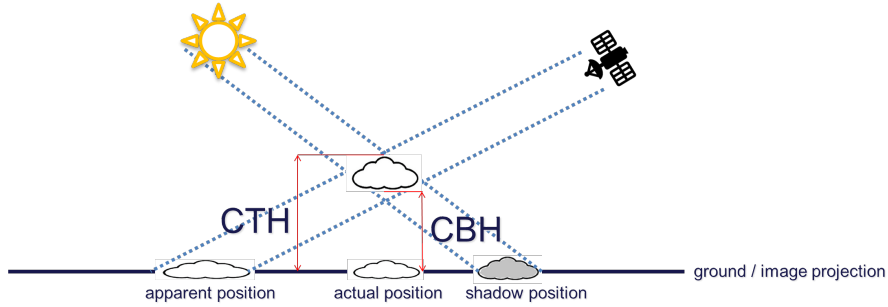


Figure 1.1.: 2D Representation of cloud position

Both the parallax and shadow corrections are crucial for accurately calculating SSI under real-time conditions. The purpose of this investigation is to find a mechanism for improving such effects. Previous work does not show any positive impact due to the sensitivity of the geometrical height, extension and overall structure of the cloud (Qu et al., 2017). Furthermore, it is important to note that work by Qu et al., 2017 was done before the SEVIRI instrument's calibration was re-visited, where an inter-calibration study found a systematic deviation of MSG SEVIRI VIS channels (Meirink et al., 2013) and implemented in CAMS (Schroedter-Homscheidt et al., 2022).

This work follows results obtained by Roy et al., 2023 using cloud top temperatures as derived by EUMETSAT from satellite observations instead of using typical cloud heights as most previous studies done using Heliosat-3 Method. A positive impact was shown for the Meteosat East satellite at two stations in India with a large distance to

the sub-satellite point in the field of view (Roy et al., 2023, in review).

These results are now tested using Heliosat-4 Method for the full field-of-view of the MSG Prime satellite over Europe and Africa and further elaborated. Furthermore, a comparison against another operational cloud retrieval done by the EUMETSAT CM-SAF is conducted to clarify the suitability of the CAMS processing chain results for the purpose.

2. The Heliosat Method

Now that the problem with the cloud parallax and shadow projection was described, it is important to understand how does CRS uses Heliosat method to estimate solar surface irradiance.

The Original Heliosat method is described in (Cano et al., 1986), and later on iterations of the method (Beyer et al., 1996)(Rigollier et al., 2004)(Pagola et al., 2014), and (Hammer et al., 2015). In Hammer et al., 2015 the authors start by calculating a relative albedo (ρ) for each pixel in the visible satellite image using the measured pixel radiance/count (c), offset count (c_0) representing clear sky conditions, and solar geometry:

$$\rho = \frac{c - c_0}{0.7f \cdot \cos(sza) \cdot \cos(sza)^{0.15}} \quad (2.1)$$

where f is the correction for the variability of the Sun-Earth distance.

A cloud index (n) is then derived by normalizing the relative albedo against estimated values for clear sky (ρ_g) and overcast conditions (ρ_c):

$$n = \frac{\rho - \rho_g}{\rho_c - \rho_g} \quad (2.2)$$

The cloud index n theoretically varies between 0 (clear sky) and 1 (overcast) and serves as a quantitative metric for cloud cover fraction for each image pixel. All previous versions of the Heliosat Method are called cloud index methods, since their physical modelling was developed for converting cloud index in irradiance.

Thus the Heliosat method calculates the reflectivity of each pixel in the satellite image, which is a measure of how much radiation is being reflected back into space by the atmosphere. To obtain the reflectivity, the raw digital counts recorded by the satellite sensors are normalized against the incoming solar radiation at the top of the atmosphere. This is represented by the Bidirectional reflectance distribution function (BRDF).

A key step in the Heliosat algorithm involves identifying the typical range of reflectivity values for each pixel over time. Through an analysis of a time series of satellite images, the algorithm determines the "typical lowest and highest reflectivities" observed at each pixel. The lowest reflectivity typically corresponds to clear sky conditions, where there is minimal cloud cover, while the highest reflectivity is associated with overcast conditions, where thick clouds dominate. These two extremes serve as reference points for estimating the solar irradiance at the Earth's surface under different atmospheric conditions (Dagestad, 2005).

To enhance the accuracy of these basic calculations, the Heliosat algorithm incorporates several corrections. One important correction accounts for the scattered radiation

from air molecules, which can influence the reflectivity values. By subtracting this scattered radiation using an empirical formula, the algorithm makes the reflectivity values more consistent across different geometries between the sun, ground, and satellite. This adjustment, known as making the reflectivity more Lambertian, ensures that reflectivity measurements taken under different conditions can be more reliably compared (Dagestad, 2005).

Another correction addresses the fact that even under the thickest cloud cover, it is never completely dark at the Earth's surface. Empirical observations have shown that when the reflectivity is close to the highest identified value, a small amount of sunlight, typically around 5 – 10% of the clear sky value, still reaches the ground. The algorithm accounts for this by slightly increasing the global irradiance estimate in these conditions, preventing the estimation from being unrealistically low (Dagestad, 2005).

The Heliosat method (Heliosat-4) used in this study uses a different approach to estimate the SSI under both clear and cloudy skies instead of the classical cloud-index models used by its predecessors. It builds on radiative transfer models but does not solve the full radiative transfer equations directly in its core. Instead, it incorporates various simplifications, including using a look-up tables (LUT) generated by running a detailed radiative transfer model under different conditions to speed up the process. (Qu et al., 2017).

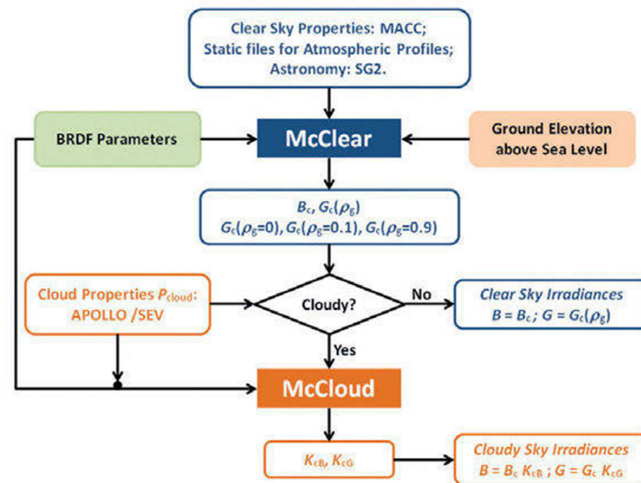


Figure 2.1.: Schematic structure of Heliosat-4 Method [image taken from Qu et al., 2017]

In this method, cloud masks based on a multi-spectral retrieval to detect clouds taken from APOLLO_NG. Heliosat-4 uses two independent models are used, one for clear-sky conditions (McClear) and one for cloudy conditions (McCloud) (see Figure 2.1). The McClear model estimates the Global SSI or Global Horizontal Irradiance (GHI) (G in Figure 2.1) and the Direct Normal Irradiance (BNI) (B in Figure 2.1). The Global SSI or GHI is comprised of the Direct Normal Irradiance component (BNI) and the Diffuse Horizontal Irradiance component (DHI). Under clear sky conditions GHI, BNI, and DHI are functions of the solar zenith angle (sza) and clear atmosphere properties, noted as P_c . GHI_c and DHI_c are also functions of the ground albedo ρ_g (Lefèvre et al., 2013). The clear-sky atmospheric properties include seven main variables related to the state

of the atmosphere under clear-sky conditions:

1. total column contents in ozone
2. total column contents in water vapor
3. elevation of the ground above mean sea level
4. vertical profile of temperature, pressure, density, and volume mixing ratio for gases as a function of altitude
5. aerosol optical depth (AOD) at $550nm$
6. Ångström coefficient
7. aerosol type

Cloud influence affects the amount of GHI, BNI, and DHI. McCloud refines the calculations of the irradiance, by adjusting the clear-sky global and direct irradiance (GHI_c and BNI_c) by considering the influence of clouds, using the parameters K_{cG} and K_{cB} , based on cloud properties (P_{cloud}) (see Equations 2.3 and 2.4). The cloud properties are discussed in detail in Qu et al., 2017.

For global SSI (GHI):

$$GHI = GHI_c(sza, \rho_g, P_c) \cdot K_{cG}(sza, \rho_g, P_c, P_{cloud}) \quad (2.3)$$

as well as for direct SSI (BNI):

$$BNI = BNI_c(\theta_S, P_c) \cdot K_{cB}(\theta_S, P_c, P_{cloud}) \quad (2.4)$$

McCloud is composed of two modules: an analytical module that computes K_{cB} as shown in Equation 2.5 and an abaci-based (or look-up tables) that computes K_{cG} as seen in Equation 2.6.

$$K_{cB} = \exp \left[\frac{-COD_{cloud}}{\cos(sza)} \right] \quad (2.5)$$

$$K_{cG}(sza, \rho_g, P_{c0}, P_{cloud}) = \frac{GHI(sza, \rho_g, P_{c0}, P_{cloud})}{G_c(sza, \rho_g, P_c)} \quad (2.6)$$

As it is seen in Equation 2.5, K_{cB} is a function of the cloud optical thickness (COD) and the solar zenith angle sza . For McCloud, cloud optical depth (COD) is assumed to be equivalent to the cloud optical thickness (COT) provided by APOLLO_NG. This assumption is necessary due to limited information on the complex 3D structure of clouds, and all retrieval processes for determining COT to date have relied on the assumption of homogeneous cloud conditions and/or made corrections for heterogeneity wherever feasible (Schroedter-Homscheidt et al., 2022). The other inputs necessary for McCloud are the cloud type and cloud coverage (mask) from APOLLO_NG. The cloud types are categorized into one of the four reference categories: low, medium, high, and thin ice cloud. Both are stored in a single dataset within the APOLLO_NG data described in 4.

This is the most fundamental cloud information needed from APOLLO_NG that is used as the main input by the Heliosat-4 algorithm in order to estimate SSI: cloud optical thickness (COT), Cloud Type, and Cloud Coverage).

Therefore, these are the main parameters that need to be adjusted by the parallax and shadow correction algorithm to better estimate SSI. This traces a clear path and shows the importance of this work. The geometrical problem of the parallax and shadow effects needs to be resolved, to shift the aforementioned parameters needed to run Heliosat-4 and estimate a better SSI.

3. Improved Geometrical correction

Now that the problem of the real position of a cloud in an image is clear, the question relies on how to correct the apparent and shadow position of a cloud in the image. Some basic assumptions and simplifications need to be established.

As it was mentioned, clouds have complex shapes and an image taken by a satellite can only go as far as the pixel size allows it. That is the main purpose of cloud masks. Each pixel will have a certain value of cloud optical depth as a measure of the presence of clouds within the area representing that pixel. And there needs to be a way of shifting those pixels to their real positions before applying the Heliosat Method, so that the SSI can be calculated.

The basis of the geometrical approach to this problem have been explored in Beyer et al., 1996 and further studied in Roy et al., 2023, but an improved geometrical correction is presented in this work.

3.1. Description of the improved geometrical correction

On any given point on Earth, the sun and satellite positions can be found given the correct time and coordinates, as shown in Figure 3.1. Given the fact that the observer is a geostationary satellite, its position remains fixed with respect to the earth. This simplifies the calculation of the viewing angles with respect to a given location.

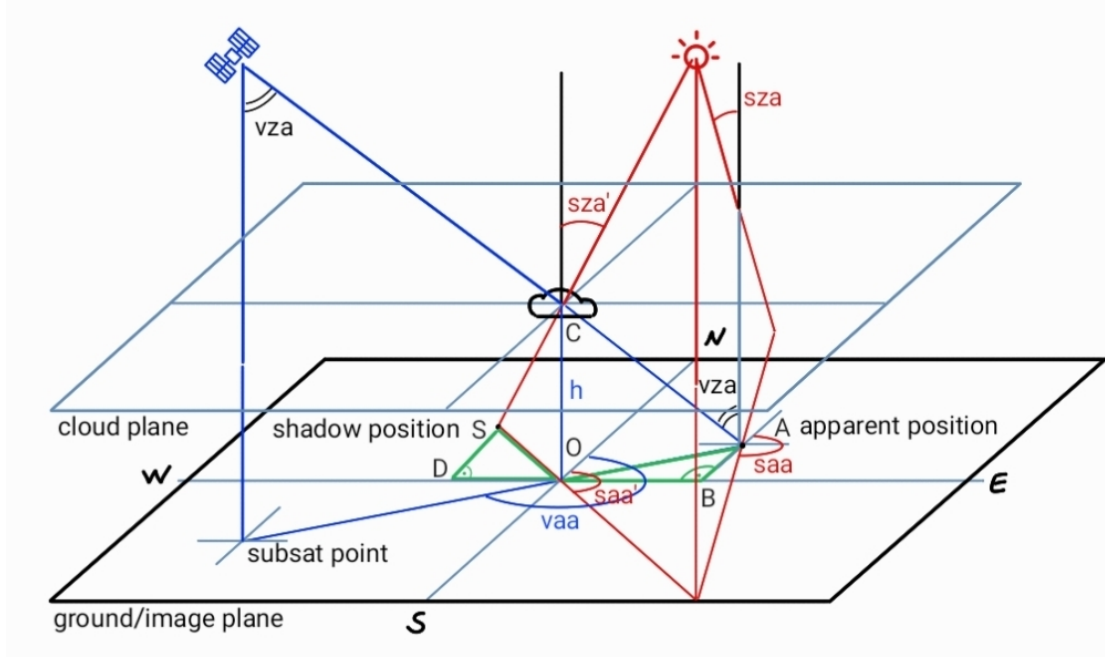


Figure 3.1.: Schematic representation of the relative geometry sun-pixel position-satellite

To calculate the parallax displacement of the cloud as seen by the satellite, the cloud height is needed. One way of knowing cloud height is having a special sensing device like a ceilometer, which are rare and expensive. Other possible ways of getting this information is estimating the cloud top height (CTH) using cloud top temperature (CTT) data and atmospheric temperature profiles.

For now, let us assume that the cloud height h is known. With it, both the parallax and the shadow displacements are functions of the cloud height. Both displacements (and its components) can be calculated using the viewing zenith angle (vza), viewing azimuth angle (vaa), solar zenith angle (sza) and solar azimuth angle (saa).

Assuming that the x-components (Δx) for each displacement is in the east/west direction and the y-components (Δy) for each displacement is in the south/north direction, the magnitude of each component was manually calculated using trigonometric functions as it was described in (Beyer et al., 1996):

For the apparent position:

$$\Delta x_{apparent} = h \cdot \tan(vza) \cdot \sin(vaa), \quad \Delta y_{apparent} = h \cdot \tan(vza) \cdot \cos(vaa) \quad (3.1)$$

as well as for the shadow displacement:

$$\Delta x_{shadow} = h \cdot \tan(sza) \cdot \sin(saa), \quad \Delta y_{shadow} = h \cdot \tan(sza) \cdot \cos(saa) \quad (3.2)$$

The satellite zenith and azimuth angles are calculated using Equations 3.4 and 3.5 (Roy et al., 2023), where the relative longitude between the SSP and the longitude of the location of the corresponding pixel ($\Delta lon = lon_{SSP} - lon$) :

$$elev = \arctan \left(\frac{\cos(\Delta_{lon}) \cdot \cos(lat) - 0.1512}{\sqrt{1 - \cos(\Delta_{lon})^2 \cdot \cos(lat)^2}} \right) \quad (3.3)$$

$$vza = 90^\circ - elev \quad (3.4)$$

$$vaa = \arctan \left(\frac{\tan(|rel_{lon}|)}{\sin(lat)} \right) \quad (3.5)$$

where vza is the viewing zenith angle and vaa is viewing azimuth angle.

Figure 3.1 shows an example of a pixel located northeast of the sub-satellite point (SSP) during an early time-of-day. In this example the sun is located southeast of the pixel. The solar angles of Point O differ from the solar angles of the original pixel where the apparent position of the cloud (Point A). This means, that for a correct calculation of the correct position of the cloud's shadow (Point S), the shadow correction must be performed using the solar angles of Point O . For this, the parallax correction should be applied in order to locate Point O . Only then the solar angles sza' and saa' can be determined. This changes Equations 3.2 slightly into 3.6. This is an important distinction not seen in Beyer et al., 1996 and Roy et al., 2023.

$$\Delta x_{shadow} = h \cdot \tan(sza') \cdot \sin(sza'), \quad \Delta y_{shadow} = h \cdot \tan(sza') \cdot \cos(sza') \quad (3.6)$$

Therefore, we perform the shadow correction in five main steps: (i) calculating the parallax displacement of each pixel, (ii) correcting the apparent cloud position, (iii) finding the new solar angles of those intermediate shifted points, (iv) calculating the shadow displacement relative to the intermediate points, and finally (v) locating the final cloud positions.

3.2. Calculation of Cloud Height

Cloud height is a key element that needs to be used on the geometrical correction. The question arises: How can the cloud height be calculated? Clouds have complex shapes that in reality has no unique top or bottom, but given its complexity, there is the need to model the cloud in an image.

Be that as it may, going back to Figure 1.1, there are two cloud distances from the ground up that could be determined: the cloud top height (CTH) and the cloud base height (CBH). There were a few possibilities that could be considered:

- Using a Ceilometer
- Employing Stereo-imaging using multiple sky cameras
- Working with atmospheric profiles and retrieving the Cloud Top Temperature (CTT) from the infrared channels of satellite images

3.2.1. Using ceilometers

Ceilometers are weather instruments that measure the height of clouds above the ground. It works by sending a narrow beam of light straight up and detecting when the light bounces off the bottom of the clouds. The time it takes for the light to return determines how far away the cloud base height is.



Figure 3.2.: Ceilometer Lufft CHM 15K [image taken from DLR Eye2Sky network]

These devices are typically used in Aviation, as they provide information about clouds, which helps pilots determine visibility which provides safety levels during takeoff, landing, and flight. Of course, they are widely used in Meteorology as well, to monitor cloud cover, assess atmospheric stability, and track weather patterns. A few of the stations in the Eye2sky Network (DLR Institute of Networked Energy Systems, 2021) have these devices, as seen in Figure 3.2, but they are rare and expensive. Ceilometers are also employed to study air quality, aerosol concentrations, and atmospheric phenomena such as fog and haze. Additionally, they are also used for wind energy applications, as they can monitor low-level clouds, which can impact wind turbine operations and energy production.

3.2.2. Using stereo-imaging

Another of the possibilities presented are stereo-imaging. This is a technique which uses two slightly different images taken from two slightly different viewpoints. This allows viewers to perceive depth and spatial relationships, just as human eyes (and other animals) perceive their surroundings, namely binocular vision.

A great example of how cloud height could be estimated is using another of the EnMet Group's projects that is being developed at the DLR: the Eye2Sky Network (Schmidt et al., 2022), which is a cloud camera network consists of 37 stations, with 28 currently in operation as of June 2021. Each station is equipped with a cloud camera, also known as an All-Sky-Imager (ASI), which captures images of the entire sky every 30 seconds. These cameras are fitted with fisheye lenses and ventilation/heating systems to ensure optimal image quality even in adverse weather conditions. The fisheye lens is able to capture the entire sky above the station. One of the cloud cameras can be seen below in Figure 3.3:

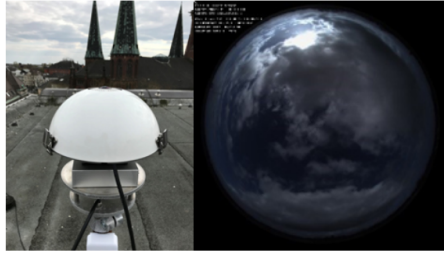


Figure 3.3.: Cloud camera Mobotix Q25 with its own ventilation and heating system
[image taken from DLR Eye2Sky network]

As the cloud cameras view the sky from the ground, two of them can determine the position and height of the cloud. This is in essence the problem of parallax: There is a difference in the apparent position of an object viewed along two different lines-of-sight of one or more observers. Similar to how a pencil in front of us would "change" location in our line-of sight if we close one eye and then the other. By using a geometrical approach similar to what is used in Astrophysics to find the distance of stars and other bright celestial bodies relative to Earth, two cameras in different places can determine the location of the same cloud, i.e. cloud base height (Lyu et al., 2021).

3.2.3. Using atmospheric profiles

There is an evident problem of using cameras or ceilometers for determining cloud height: There are limited devices that could be used locally and certainly not for any location in the satellite's field-of-view.

That brings us to the last approach: Working with atmospheric profiles and retrieving the Cloud Top Temperature (CTT) from the infrared channels of satellite images.

Atmospheric profiles are fundamental representations of how various atmospheric parameters change with elevation. These profiles provide essential information for applications in aviation, climate modeling, and atmospheric science research. For example, the International Standard Atmosphere (ISA) (Figure 3.4) is a widely used model of the Earth's atmosphere. It defines standard values for temperature, pressure, density, and other atmospheric parameters at different altitudes. This serves as a reference for aircraft performance calculations, aviation engineering, and atmospheric science research. The ISA profile is based on average atmospheric conditions derived from extensive measurements and observations collected from various locations around the globe.

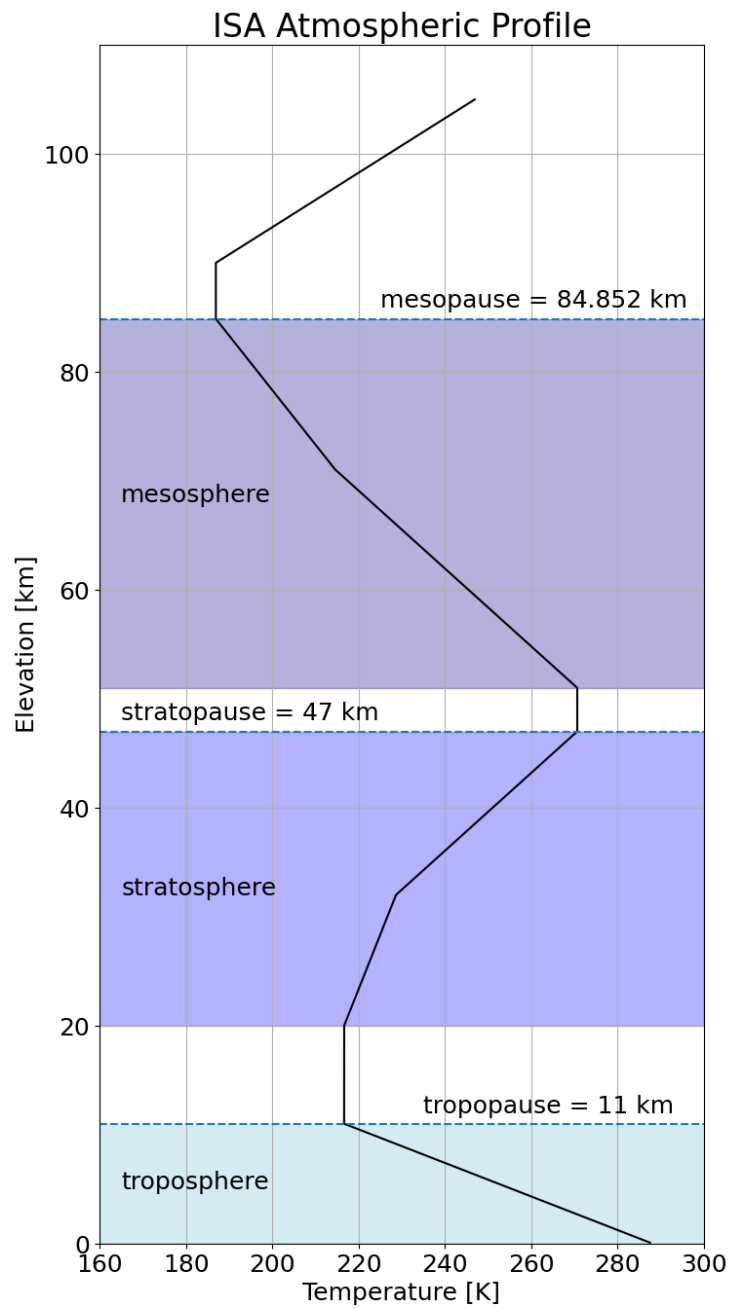


Figure 3.4.: ISA Atmospheric Temperature Profile

More detailed atmospheric profiles were developed by the AFGL and McClatchey (Anderson et al., 1986). The Air Force Geophysics Laboratory (AFGL) developed at-

mospheric constituent profiles independently to characterize the vertical distribution of various gases and aerosols in the atmosphere. These profiles though, provide a seasonal and regional distinction, which can be more useful, since atmospheric conditions can vary between different locations and seasons, e.g. there are different profiles between tropical and subarctic regions.

Last but not least, a different dataset of profiles was compiled by D. S. McClatchey and colleagues, which includes extensive measurements of atmospheric transmittance and constituent concentrations. The McClatchey dataset is widely used in atmospheric radiative transfer modeling, remote sensing algorithm development, and climate studies.

By using these atmospheric profiles as a sort of lookup-tables, Cloud Top Temperature values of a given cloud can be used to look for the probable height of the top of the cloud.

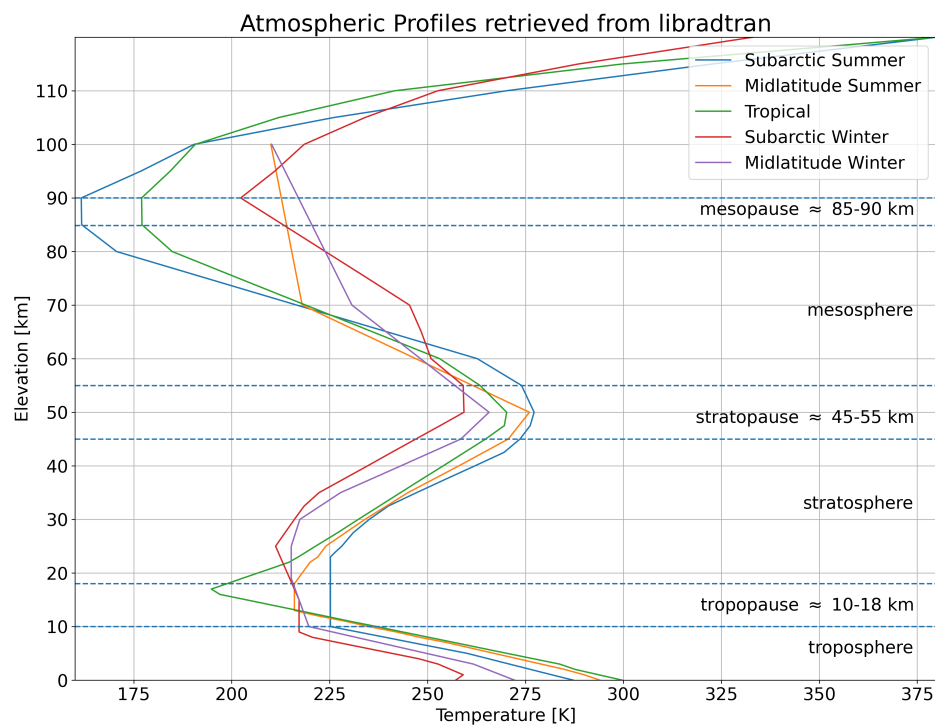


Figure 3.5.: AFGL and McClatchey Atmospheric Temperature Profiles

4. Data

The geometrical correction needs to be applied on existing historical data that will be used to estimate SSI. The necessary cloud information to be used by the Heliosat-4 algorithm to estimate SSI are: cloud optical thickness (COT), Cloud Type, and Cloud Coverage). These are the main parameters that need to be shifted by using the geometrical correction before applying Heliosat-4.

In order to calculate the parallax and shadow displacements, cloud height needs to be determined. The cloud information obtained from APOLLO_NG includes Cloud Top Temperature (CTT). Using this information alongside standard atmospheric temperature profiles, Cloud Top Height (CTH) can be estimated. The Satellite Application Facility on Climate Monitoring (CM-SAF) also provides ready-to-go instantaneous gridded CTT and CTH products. Therefore, three different approaches are used to ascertain which datasets are more suitable for the geometrical correction presented in this work.

A validation method is also determined to see if its implementation has a positive impact in the existing operation scheme of the CAMS radiation Service (CRS). A world-wide database of irradiance measurements developed by the OIE center of Mines Paris PSL are used as reference values.

Table 4.1 summarizes the datasets used in this work with the relevant details. Following sections describe each of them more thoroughly. The overview of the whole algorithm can be seen in Annex B.

Table 4.1.: Data sources and resolutions for various cloud and irradiance measurements

Data	Source(s)	Time res.	Spatial resolution
COT	APOLLO_NG	15-minutes	$3 \times 3km$ (3172×3172 px)
Cloud Type	APOLLO_NG	15-minutes	$3 \times 3km$ (3172×3172 px)
Cloud Coverage	APOLLO_NG	15-minutes	$3 \times 3km$ (3172×3172 px)
CTT	APOLLO_NG, CM-SAF	15-minutes	$3 \times 3km$ (3172×3172 px)
CTH	CM-SAF	15-minutes	$3 \times 3km$ (3172×3172 px)
GHI	libinsitu	1 minute	single-point per station

4.1. APOLLO_NG Data

APOLLO_NG retrieves cloud information, namely the cloud properties, using the method describe in Klüser et al., 2015. The cloud data used in this work is retrieved for the field-of-view of Meteosat-11 located over the Atlantic Ocean at latitude 0° and longitude 0° with a nominal spatial resolution of $3km \times 3km$ pixel size at the sub-satellite point. The datasets of each of these maps include: Cloud Top Temperature (CTT), Cloud Optical Thickness (COT), Cloud Coverage and Cloud Type. This time-series data also contain the geolocation of each pixel. This results in maps of 3712×3712 pixels.

Cloud Top Temperature (CTT) : Temperature in Kelvin of the detected cloud pixel as seen from the satellite.

Cloud Optical Thickness (COT) : a measure of the attenuation of light passing through the atmosphere due to scattering and absorption (value between 0-500)

Cloud Coverage : percentage of presumed cloudiness of pixel, with values ranging between 0% (clear pixel) and 100% (totally cloudy pixel)

Cloud Type : Classification of the pixels corresponding to the classification scheme:

- 1 = clear (no cloud)
- 5 = low-level cloud
- 6 = medium-level cloud
- 7 = high-level cloud
- 8 = thin ice cloud

In principle, the satellite takes 15 minutes to scan its whole field-of-view, so data can be read in 15-minute time steps. However, for this study a subset of data from the year 2021 during daylight slots was taken every two hours, that is: 0900, 1100, 1300, 1500, 1700. This is enough to cover a wide variety of cloud patterns and formations and see how parallax behaves during the day and avoid long computation times. The data was provided directly by the CAMS development team. Another important reason for using daylight time slots is the fact that low solar elevation angles presents challenges to estimate SSI for both satellite-based models and ground-based instruments. Damiani et al., 2018 highlights that satellite-derived irradiance estimates tend to have higher root mean square errors (RMSE) under low sun elevation conditions, primarily due to increased atmospheric path lengths and scattering effects. Additionally, Blaga et al., 2021 discusses how atmospheric aerosols can significantly impact irradiance measurements, especially when the sun is near the horizon, leading to discrepancies between modeled and observed data. This underscore the importance of considering low solar elevation angles in solar irradiance assessments to enhance the accuracy of both satellite-based models and ground-based measurements.

Figure 4.1 shows an example of the zoomed-in COT map around Oldenburg (latlon = (53,8)) of uncorrected data.

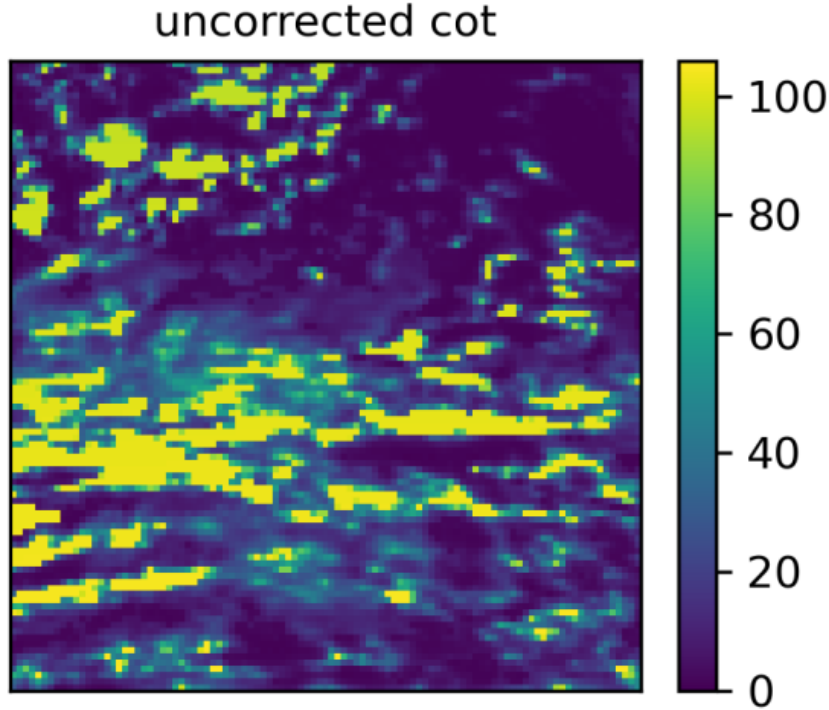


Figure 4.1.: Example COT data [timestamp: '2021-10-21T11:00']

The datasets mentioned (Cloud Optical Thickness (COT), Cloud Coverage and Cloud Type) are the main parameters that were shifted by the correction algorithm.

The cloud top temperature (CTT) from APOLLO_NG is used alongside atmospheric temperature profiles (described in Section 3.2.3) to derive cloud top height (CTH). During the first stages of the implementation, a few non-physical values of CTT were encountered, e.g. 0K (see Figure 4.2). These values might be presumably pointing to some software bug and was therefore reported back to the CAMS developer team. The improvement of this issue is outside the scope of this work and will be dealt by CAMS developers. A few other values would be out of the possible range of temperatures within the atmospheric temperature profiles used (see Section 3.2.3). In both cases, cloud top temperature values were interpolated to ensure the cloud top height taken for each pixel would be within the height ranges of the troposphere for that region. An example of the final interpolated values of CTT used for the calculation of CTH are shown in Figure 4.2.

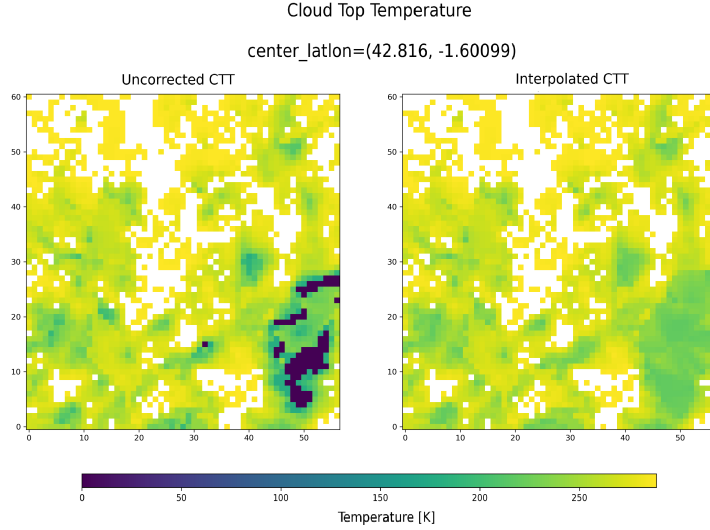


Figure 4.2.: Example CTT data with non-physical values encountered [timestamp: '2021-07-09T15:00']

The full timeframe of time series data (1-minute resolution) for the location of each station is downloaded prior to running the corrections. The main dataset that will be compared to is Global Horizontal Irradiance (GHI) data for each location.

4.2. Satellite Data from CM-SAF

The Satellite Application Facility on Climate Monitoring (CM-SAF) is part of the European operational satellite agency for monitoring weather (EUMETSAT) in an effort to develop, generate, archive and distribute high-quality satellite-derived products (EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF), 2024). Instantaneous gridded cloud top temperature (CTT) and cloud top height (CTH) values with the same resolution as APOLLO_NG datasets were retrieved from CM-SAF products.

CM-SAF products use different cloud masking than APOLLO_NG datasets. The algorithms for cloud detection of these products have been developed specifically for the AVHRR on board the current and future polar-orbiting NOAA and EUMETSAT Meteorological and Operational Weather Satellites (METOPs). Nevertheless, implementation of said algorithms could be performed on other satellites, provided that spectral channel differences are not too large (Dybbroe et al., 2005). As for APOLLO_NG, the CTT from CM-SAF corresponds to the temperature of the cloudy pixel in Kelvin as seen from the satellite. The CTH corresponds to the height of the top boundary of the detected cloud in kilometers. The data was obtained using the web interface of CM-SAF (EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF), 2024).

4.3. Differences between APOLLO_NG and CM-SAF Datasets

An important point here is made to both understand better the need for using different sources of the same data types, e.g. Cloud Top Temperature (CTT), and to understand the data itself that is used.

Since both data sources use different algorithms for cloud detection and cloud masking, the CTT data will naturally be different. The questions that arise are: how different are the two datasets? Will there be a significant difference in the estimation of the cloud height for either one of them? And will one of these datasets determine a better outcome of the overall geometrical correction with respect to the other? The first question is explored here, in order to better understand the sources. Later on, applying the geometrical correction to the different data sources will answer the remaining questions.

The differences between the retrieved CTT from CM-SAF and APOLLO_NG can be seen clearly in Figure 4.3 as a specific example of full disk map for the entire field-of-view of Meteosat. In this first comparison of a defined point in time, a few key differences are noticed. At the boundary conditions of the image, APOLLO_NG shows some problems, registering exceedingly low temperatures. This is somewhat expected, as the satellite viewing angles are far over the limit in which degradation of the quality in the retrieval of cloud properties is caused for APOLLO_NG (Qu et al., 2017). Another visible difference was described in Section 4.1, where non-physical values of CTT are seen in a few other pixels far from the boundary of the image. This is the reason behind the interpolation of said values.

Another important difference between these two datasets is the presence of clouds in the central region, detected by APOLLO_NG, but not for CM-SAF. These detected clouds are presumably a mixture between low-level clouds, dust clouds coming from the Sahara region in its direction to the east. This is usual above the Atlantic, and no meteorological stations are within this region, which could potentially have a big impact in the validation of the geometrical correction. A last noticeable difference is the temperature range of CM-SAF instantaneous CTT. CTT values for CM-SAF look slightly darker, meaning lower temperature of clouds, as seen in Figure 4.3. These differences are just a simple overview of the complete field-of-view images like the one in Figure 4.3, but a proper analysis of the differences needs to be done.

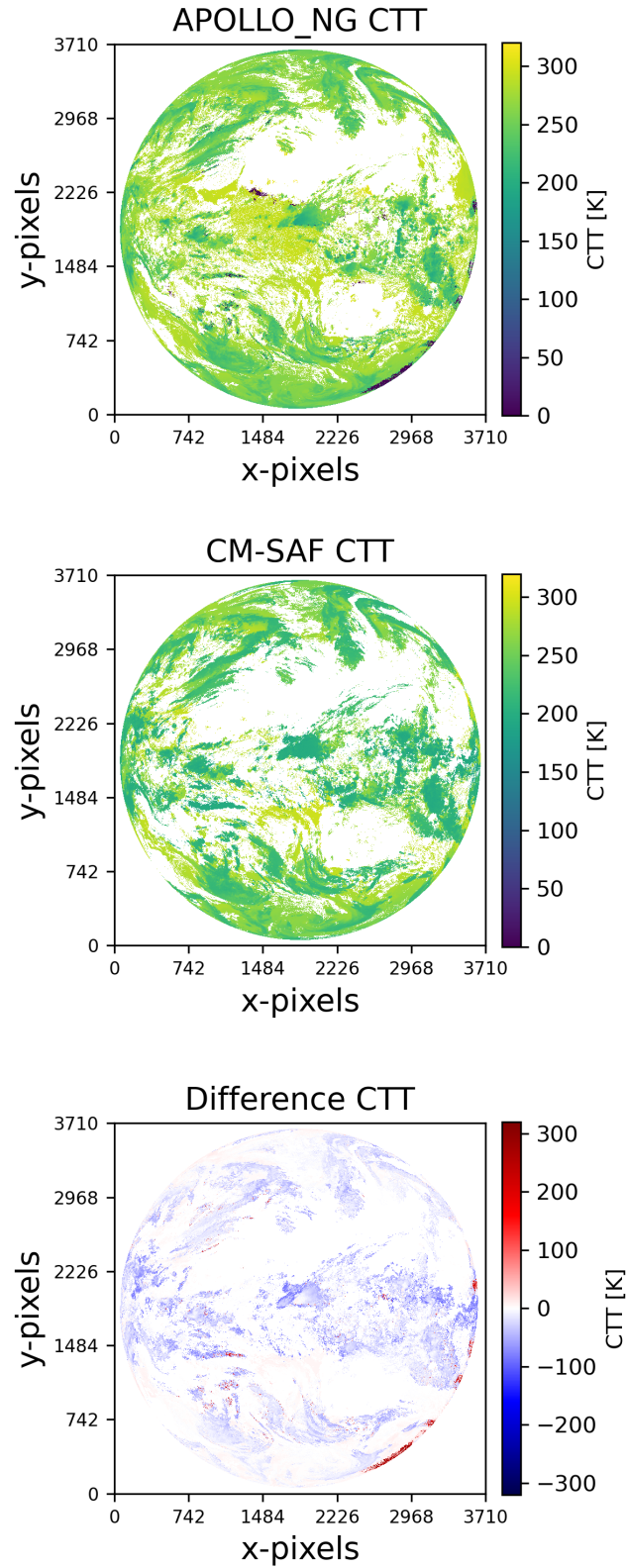


Figure 4.3.: CTT Datasets comparison between APOLLO_NG and CM-SAF [timestamp: '2021-04-12T13:00']

A time series overlook of 2021 data for the CTT values for the meteorological stations used in this work (listed in Section 4.5) is shown in Figure 4.5. This shows a clearer picture of the CTT differences in both datasets for all stations. This confirms the overall trend of lower CTT values for the CM-SAF dataset, where the difference between APOLLO_NG and CM-SAF values of CTT ($CTT_{APOLLO_NG} - CTT_{CM-SAF}$) is mainly positive. There are some negative peaks in the difference, because of the non-physical values (0 Kelvin) of CTT found in the APOLLO_NG dataset (see Figure 4.2). Another interesting observation in Figure 4.5 is the big differences in CTT values for the Budapest station, which might turn problematic for estimating the cloud top height during the implementation of parallax and shadow corrections, but this will be discussed in the results.

An look at the probability distribution of the cloud top temperature values of these stations for the whole time period of the data was explored. This also reveals a clear difference between the datasets. CM-SAF shows two distinctive peaks in the probability of the cloud top temperature: A low-temperature peak (at around 210K) of lower-level clouds, and a peak at around 270K. These low-level clouds are detected roughly within the same probability, although with APOLLO_NG is slightly higher. Nevertheless the low-temperature peak is not seen in APOLLO_NG. The probability of finding non-physical values of CTT for APOLLO_NG is visible, but the interpolation of CTT values will mitigate encountering such error.

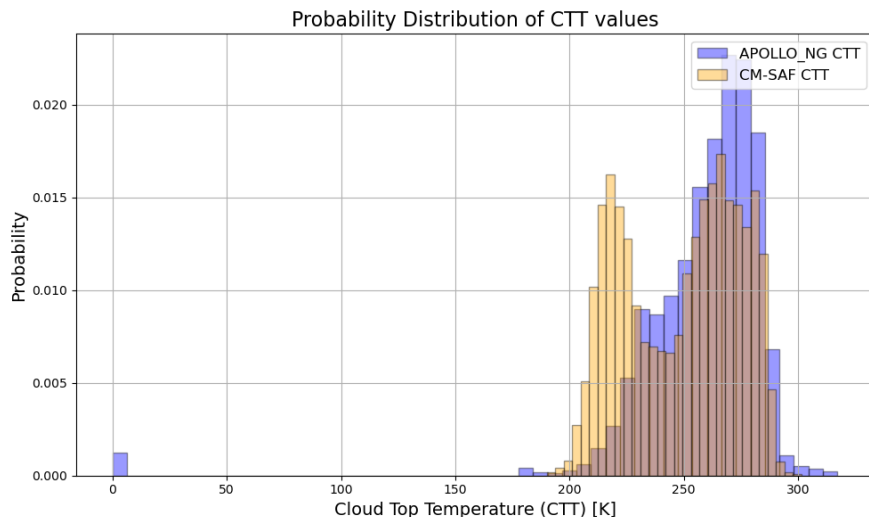


Figure 4.4.: Probability Distribution of CTT values

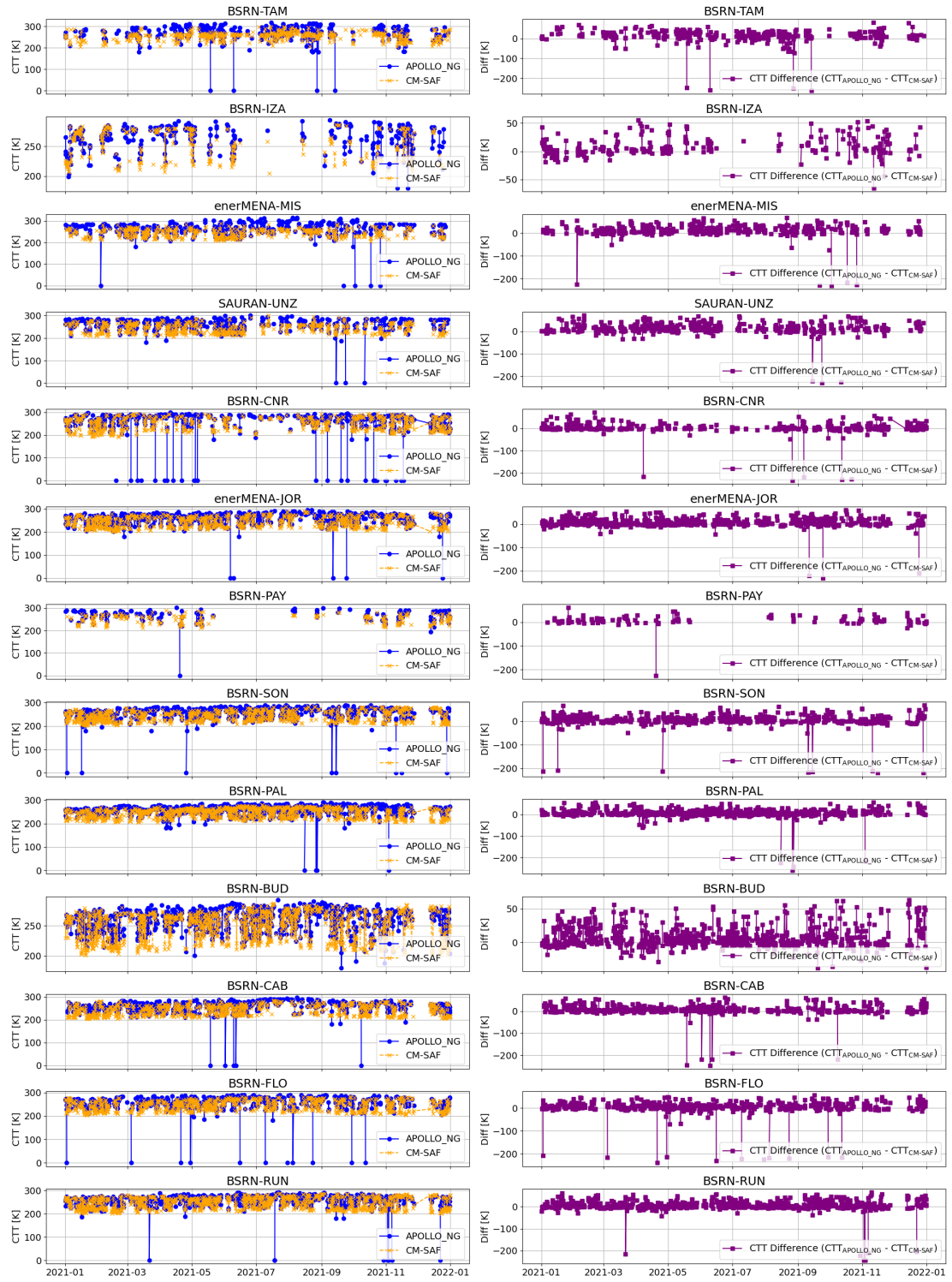


Figure 4.5.: Time series comparison between CTT datasets

4.4. Cloud Height

For estimating the cloud height, three different approaches based on available data are used in this work. The general workflow for any of the approaches presented follows roughly the same path, as seen in Figure 4.6. Once the correction is performed, the interpolated values of the Cloud Optical Thickness (COT), Cloud Coverage, and Cloud Type are extracted for each of the subsections selected for the study. These values are used to prepare as input for Heliosat-4, which will estimate the Solar Surface Irradiance (SSI), mainly the Global Horizontal Irradiance (GHI), which was used to compare against the uncorrected values of CAMS data and the ground observation data.

The APOLLO_NG images described at the beginning of this chapter for the year 2021 are used for this analysis. For each timestamp, Cloud Top Temperature (CTT), Cloud Optical Thickness (COT), Cloud Coverage, and Cloud Type are obtained from these maps. For computational efficiency, each correction is performed on a subsection around the location of each selected station is found and the correction method is applied for each zoomed-in region. Each method differs slightly from each other as described in this chapter. Each subsection is comprised of 17×17 pixels. The reason being, the angle for lower solar elevations and huge shifts in cloud displacement is enough, as it takes enough ground for the geometrical correction. One more reason for this is computational efficiency.

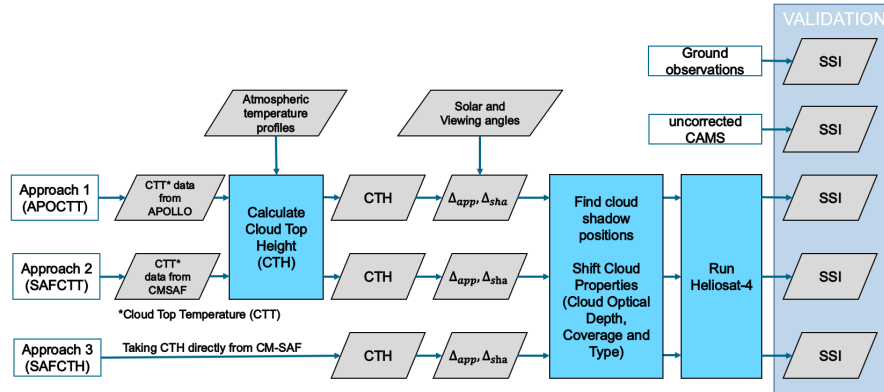


Figure 4.6.: Overview of the different correction approaches

4.4.1. Approach 1: Using Cloud Top Temperature Data from APOLLO_NG (APOCTT)

For each subsection, the pixel-wise satellite estimated cloud top temperature data from the APOLLO_NG maps is obtained. An appropriate atmospheric profile described in Section 3.2.3 is used. As described in Section 4.1 if there is the need of interpolating or removing flagged cloud top temperature (CTT) values, this is done prior to getting the corresponding cloud optical thickness (COT) values.

With the found CTH values, Equations 3.1 can be calculated to determine the distances (eastwards/westwards and northwards/southwards) for the parallax correction first. Using Haversine formulae, the estimated distances can be turned into new location points (latitude and longitude) for each pixel. For these new locations, new solar angles

(solar zenith angle (sza) and solar azimuth angle (saa) are calculated and with these, plus the already found cloud top height (CTH), is the shadow correction applied using Equations 3.6. This is a key difference between Beyer et al., 1996, which uses the solar angles for both the parallax and shadow corrections. Since the parallax correction essentially shifts the pixel to the location underneath the cloud (actual position, see Figure 3.1), the angles needed for the shadow correction should be for the location found after the parallax correction and not the apparent position of the cloud.

Since not all pixels are cloudy-pixels, not all pixels are shifted, so their position remains the same. This is important to ensure that all valid values within the region are taken in order to use them to interpolate them back to the original grid. Using a bilinear interpolation method, each of the parameters are shifted. The three main parameters that are shifted are Cloud Optical Thickness (COT), Cloud Coverage, and Cloud Type. These parameters are the main input for Heliosat-4.

4.4.2. Approach 2: Using Cloud Top Temperature Data from CM-SAF (SAFCTT)

The second method has the same structure as the method described above. The key difference is the source of the pixel-wise cloud top temperature (CTT) used for the cloud top height (CTH) estimation. CTT values from the CM-SAF datasets for the same timestamp are used. CM-SAF datasets correspond to the same satellite's field-of-view. This simplifies running operation of the algorithm since they have the same shape and resolution. This cloud top temperature dataset is then used for the cloud top height (CTH) calculation the same as with Approach 1. The main purpose behind this method is to assess the difference cloud masking from CM-SAF, and see whether this product shows a significant impact in the cloud correction method. This is shown more clearly in Figure 4.3.

It is important to note that the instantaneous cloud top height (CTH from CM-SAF is derived using cloud top temperature (CTT) and atmospheric temperature profiles from the Air Force Geophysical Lab (AFGL) (Dybbroe et al., 2005), similar to how it is presented in this work. There are some visible differences, mainly caused by the different cloud masking. This is the reasoning behind testing both datasets on the approaches explained below.

4.4.3. Approach 3: Using Cloud Top Height Data directly from APOLLO_NG (SAFCTH)

The last method avoids the use of cloud top temperature (CTT) data and atmospheric temperature profiles to estimate the cloud top height (CTH). The CM-SAF product provides instantaneous CTH of the same shape and resolution as the APOLLO_NG data. Nevertheless, it felt important to compare between the use of CTH given by CM-SAF with the CTH estimated using APOLLO_NG CTT Data. This CTH is used as the input for equations 3.1 and 3.6. The difference between the results for the estimated cloud top height (CTH) in Approach 1 and the CTH product from CM-SAF is shown in Figure 4.3 as well.

4.5. SSI Ground observations

A world wide database of irradiance measurements, collected as part of CAMS in the THREDDS server is used as the reference data for this work. A special library developed by a research group at center O.I.E of Mines-Paristech (“In Situ Energy Analysis”, n.d.) was built to transform in situ data time series) into a standard NetCDF format. This library takes solar radiation stations all around the world and high quality data of SSI is fetched, downloaded, examined and cleaned. The ground observation data is used as the reference of "true" values for comparing the current model with the cloud correction approaches done in this work. Therefore these ground observations need to be reliable and accurate. However, it can be prone to various errors, including instrumental malfunctions, human errors, and environmental interferences. A Quality Check (QC) of the raw ground observation data is performed in order to flag this inconsistencies and errors and clean the data that use for comparison. This is done prior to the validation in order to properly assess the validity of the cloud correction method. Using this library, the availability of radiation data can be sought. The stations retained for this assessment are described in Table 4.2.

Table 4.2.: Meteorological Stations retrieved from THREDDS Server

Name	Latitude [°N]	Longitude [°E]	Elevation [m]	Location
BUD	47.4291	19.1822	139.1	Budapest, Hungary
CAB	51.9711	4.9267	0.0	Cabauw, Netherlands
CNR	42.8160	-1.6010	471.0	Cener, Spain
FLO	-27.6047	-48.5227	11.0	Florianopolis, Brazil
INO	44.3439	26.0123	110.0	Magurele, Romania
IZA	28.3093	-16.4993	2372.9	Izana, Spain
PAL	48.7130	2.2080	156.0	Palaiseau, France
PAY	46.8150	6.9440	491.0	Payerne, Switzerland
RUN	-20.9014	55.4836	116.0	Reunion Island, France
SON	47.0540	12.9577	3108.9	Sonnblick, Austria
TAM	22.7903	5.5292	1385.0	Tamanrasset, Algeria
UNZ	-28.8529	31.8516	90.0	Uni Zululand, South Africa
JOR	30.1720	35.8180	1012.0	Ma'an, Jordan
MIS	32.8600	-4.1070	1107.0	Missour, Morocco
OUJ	34.6500	-1.9000	617.0	Oujda, Morocco

The stations selected for this study can be seen in Figure 4.7.

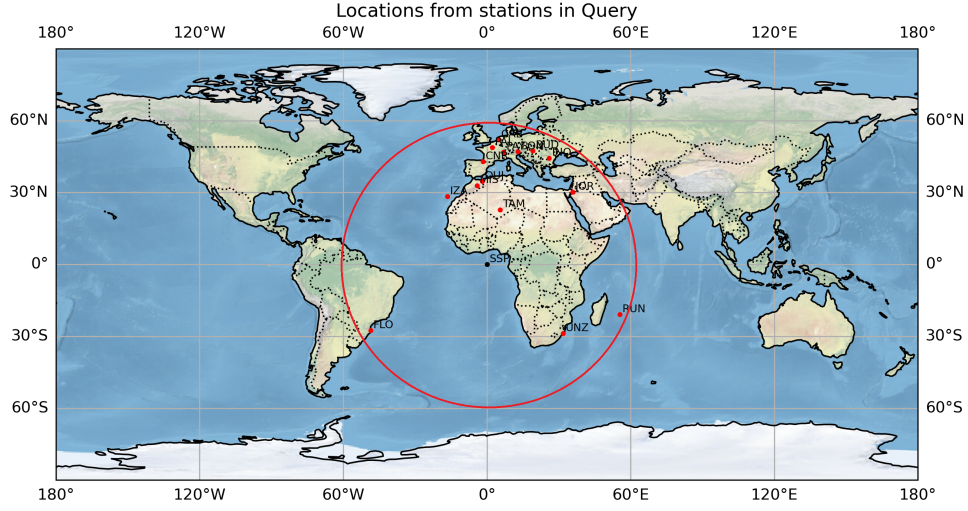


Figure 4.7.: Selected stations for the implementation

The data collected various solar irradiance data provided by “In Situ Energy Analysis”, n.d. The structure and detailed description of this data can be found in Driemel et al., 2018. The examination of these datasets is done by applying quality checks, which essentially perform a series of tests, in which atypical, physically impossible values of SSI are flagged and highlighted based on Forstinger et al., 2023 and Geuder et al., 2015. These cleaned values will be used as reference for comparison to the current CAMS model and the different cloud correction approaches investigated in this work. Data is stored in a QC Dashboard, a helpful visualization of all tests and flagged data for each station, is also available as an image file (as shown in Figure 4.8).

As an effort done to select suitable stations for validation and ensure the quality of the ground observation data, *libgndobs* library went into the development. *libgndobs* allows users to retrieve ground observation data using a query strategy in which the stations are filtered depending on user criteria, for example:

- all stations within a defined time range
- all stations inside a defined latitude/longitude domain
- all stations around a given radius (in kilometers) from a latitude-longitude coordinate
- all stations with a certain Köppen climate classification

The datasets for each station is retrieved using the in-house Python library *libgndobs*. QC is performed and flagged data is cleaned out, in order to use the most reliable reference data. SSI results obtained after the improved geometrical correction using all different approaches will then be validated using different error metrics described in Chapter 5.1.

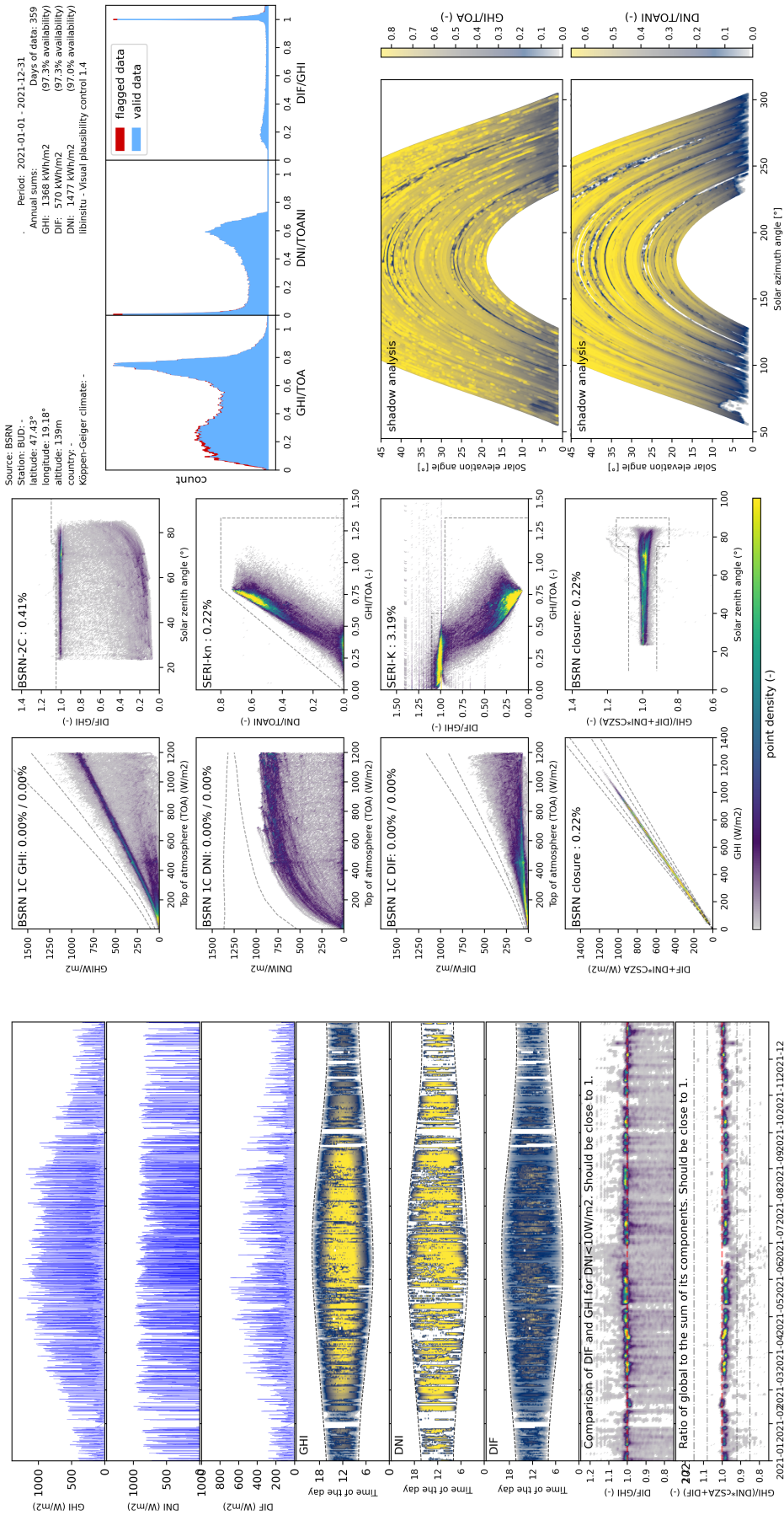


Figure 4.8.: Example QC on BSRN-BUD station [time range: '2021-01-01 - 2022-01-01']

5. Results and Analysis

The study utilized data from several ground meteorological stations distributed across the field-of-view of the MSG Satellite, whose sub-satellite-point (SSP) is located in Latitude $0^\circ N$ and Longitude $0^\circ E$. The reference data covered a range of meteorological and environmental parameters, which are critical for evaluating the satellite model's performance. In order to assess the performance of each of the correction methods described in Chapter 4.4, a comparison for all correction approaches is performed by analyzing the GHI values from the uncorrected and corrected datasets against the corresponding values from the ground measurement data. The evaluation focused on three key statistical metrics, both as absolute and relative values, described in Section 5.1. Following this, the results for these metrics is described and analyzed.

5.1. Evaluation Methodology

5.1.1. Absolute metrics

Mean Bias Error (MBE)

This metric indicates the systematic error of the predictions. It is calculated as the mean difference between the predicted values and the actual ground measurement values. A positive MBE indicates overestimation, while a negative MBE indicates underestimation.

$$\text{MBE} = \frac{1}{N} \sum_{i=1}^N (P_i - M_i)$$

where P_i is the predicted GHI value, namely the uncorrected CAMS GHI values, or each of the difference GHI values obtained from the three different correction approaches, M_i is the ground measured GHI value, and N is the number of observations.

Mean Absolute Error (MAE)

This metric provides the average magnitude of the errors in the predictions, without considering their direction. It is a measure of how close the predictions are to the actual measurements.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |P_i - M_i|$$

Root Mean Square Error (RMSE)

The RMSE measures the square root of the average squared differences between the predicted values and the actual measurements. It gives a higher weight to larger errors, providing a measure of the prediction accuracy.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - M_i)^2}$$

5.1.2. Relative metrics

In addition to absolute error metrics, relative metrics are crucial for understanding the performance of models in relation to the magnitude of the measured values. These metrics provide a normalized evaluation, making them useful for comparing errors across different scales. Relative error metrics help the comparison between results of different stations which are located in various regions. This is especially useful, because different regions receive different scales of incoming radiation. This is due to different geographical and climate conditions.

Relative Bias error(rel. MBE)

The relative mean bias error expresses the bias as a percentage of the mean ground measurement values. This helps to contextualize the bias in terms of the overall magnitude of the data.

$$\text{rel. MBE (\%)} = \frac{\text{MBE}}{\overline{M}} \times 100$$

where $\overline{M}$ is the mean of the ground measured GHI values.

Relative Mean Absolute Error (rel. MAE)

The relative MAE is a relative metric that indicates the average absolute error as a percentage of the actual measurements. This metric is especially useful for understanding the error in proportion to the magnitude of the data.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{P_i - M_i}{M_i} \right| \times 100$$

Relative MAE is particularly useful when dealing with datasets where the range of values is wide, as it normalizes the error to the size of the actual values.

Relative Root Mean Square Error (rel. RMSE)

The relative RMSE provides a normalized measure of the RMSE by dividing it by the range or the mean of the measured values, i.e. ground observations. This normalization facilitates the comparison of errors across datasets with different scales.

$$\text{rel. RMSE} = \frac{\text{RMSE}}{\overline{M}} \times 100$$

The relative RMSE expressed as a percentage is a widely used metric for comparing model performance across different datasets, making it easier to interpret the magnitude of the errors relative to the data.

Mean Percentage Improvement

The Mean Percentage Improvement (MPI) is a measure used to assess the improvement of each correction approach compared to the baseline model, in this case, the uncorrected CAMS model. This metric is particularly useful in quantifying the relative improvement in absolute or relative error metrics previously described: MBE, MAE, and RMSE. For any given station and correction approach, the MPI is calculated by taking the absolute value of the difference between the uncorrected CAMS error and the corrected model error, then normalizing it by the CAMS error, as shown in Equation 5.1.

$$\text{MPI}_{\text{station}} = \left(\frac{|\text{Uncorrected}_{\text{error}}| - |\text{Corrected}_{\text{error}}|}{|\text{Uncorrected}_{\text{error}}|} \right) \times 100 \quad (5.1)$$

Here:

- $\text{Uncorrected}_{\text{error}}$ represents the absolute error metric (MBE, MAE, or RMSE) for the CAMS model.
- $\text{Corrected}_{\text{error}}$ represents the absolute error metric for the selected correction method.

Equation 5.1 calculates the MPI for each station individually. A positive MPI value indicates an improvement, while a negative MPI value implies a worsening in the error metric relative to the CAMS model.

To obtain the overall MPI for each group of stations (e.g., best-performing, worst-performing), we average the MPI values across all stations in that group, as shown in Equation 5.2.

$$\text{MPI}_{\text{group}} = \frac{1}{N} \sum_{i=1}^N \text{MPI}_{\text{station}_i} \quad (5.2)$$

Here:

- N represents the number of stations in the group.
- $\text{MPI}_{\text{station}_i}$ is the MPI for each individual station i in the group.

Similarly, to compute the overall MPI across all groups, the same approach is applied, yielding a single aggregated improvement value for each approach. This aggregated MPI serves as an indicator of the average improvement that each method provides relative to the CAMS model, across all stations considered in the analysis.

In this way, the MPI allows for a direct comparison between methods, offering insight into the effectiveness of each correction technique in reducing MBE, MAE, and RMSE errors. As it was mentioned in chapter 4, the evaluation of the parallax and shadow correction method is based on data from ground stations used as reference values of Global

Horizontal Irradiance (GHI). The accuracy of the geometrical correction using each cloud estimation approaches is assessed using absolute and relative bias error (MBE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) against the same metrics for the current CAMS model. This chapter presents the relative error metrics, which are important for comparing the different stations. Relative metrics are particularly useful in the context of this work, to properly compare results across the different stations. This is because all stations have widely different locations, which make them have different climate conditions and therefore receive different incoming irradiation. Thus, normalization of SSI values help the comparison between these stations.

5.2. Overall relative error metrics

The results presented are sorted from distance to the sub-satellite-point (SSP). This is done in order to determine if larger shifting distances might profit more from the correction methods applied. The timeframe used for this study is the year 2021. The metrics presented in this section show the overall relative yearly metrics comparison. The overall absolute yearly metrics can be seen in Annex (C). Table 5.1 and Figure 5.1 shows the results for the relative error metrics.

Table 5.1.: Relative Metrics for 2021 Data

Stations	Relative MBE (%)				Relative MAE (%)				Relative RMSE (%)			
	CAMS	APOCTT	SAFCTT	SAFCTH	CAMS	APOCTT	SAFCTT	SAFCTH	CAMS	APOCTT	SAFCTT	SAFCTH
TAM	-5.11	-4.76	-4.82	-4.74	9.96	9.33	9.47	9.35	13.77	12.83	13.04	12.90
IZA	-11.93	-7.97	-8.06	-8.02	14.27	10.85	10.99	11.00	21.34	14.17	14.42	14.40
MIS	-4.36	-3.96	-3.89	-3.88	10.86	10.21	10.22	10.29	17.82	16.68	16.84	16.98
OUJ	-0.54	0.77	1.14	1.07	13.32	12.11	12.48	12.58	18.58	17.02	17.50	17.47
UNZ	3.47	1.80	2.35	2.34	16.67	19.04	19.10	19.01	23.14	25.46	26.01	25.98
CNR	-0.93	1.68	2.08	1.92	16.53	16.51	16.53	16.80	23.49	23.92	24.00	24.64
JOR	-5.87	-5.80	-5.77	-5.79	7.35	7.23	7.19	7.21	13.04	12.16	12.10	12.18
PAY	-1.35	2.07	1.87	2.23	18.04	19.38	19.47	19.21	29.29	32.36	32.34	31.89
SON	-25.67	-29.61	-29.37	-29.77	46.23	49.50	49.53	49.45	66.07	70.57	70.64	70.57
PAL	-2.97	0.71	0.40	0.28	21.77	22.68	22.88	22.84	35.24	36.33	37.01	37.08
INO	-3.22	-1.67	-1.80	-2.08	13.45	12.69	12.98	12.90	20.28	19.30	19.99	20.12
BUD	-1.00	2.02	2.05	2.32	16.29	16.42	16.53	16.66	28.30	29.17	29.38	29.33
CAB	-2.96	-0.51	-0.78	-0.48	20.85	20.45	20.17	20.53	33.41	32.72	32.45	32.82
FLO	-0.88	1.69	1.52	1.73	19.37	20.69	20.79	20.69	32.73	34.65	34.77	34.69
RUN	-7.58	11.61	10.13	9.47	26.99	30.76	30.56	30.34	43.50	49.24	49.13	49.02

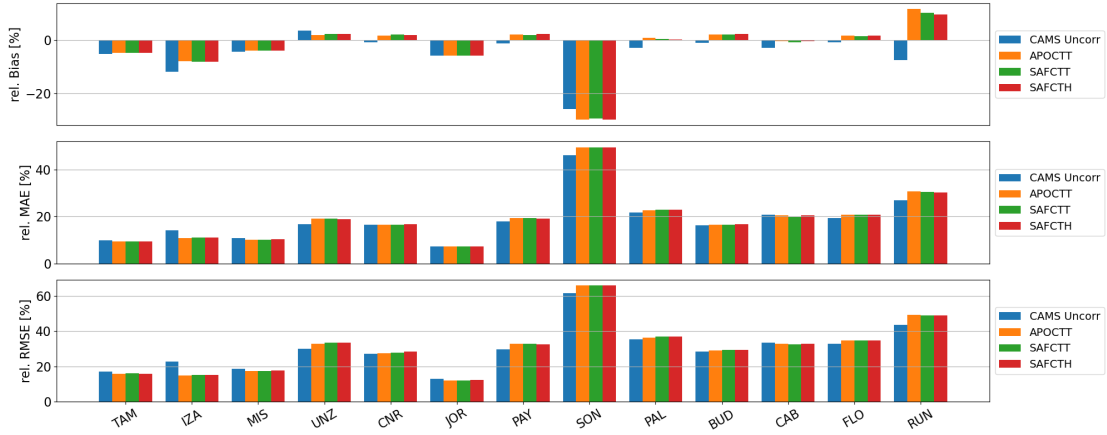


Figure 5.1.: Relative error comparison for 2021 data

It is evident that the different stations exhibit varying degrees of error across the correction methods employed. Sonnblick shows the highest relative errors across all metrics and methods, with a particularly high relative MBE of approximately -29.77% and relative RMSE reaching up to 70.64% , indicating significant discrepancies in estimation of Global Horizontal Irradiance (GHI). Nevertheless, Jordan consistently displays the smallest errors, with relative MBE values around -5.87% and the lowest relative MAE and RMSE values, suggesting more accurate and reliable predictions. Among the different cloud height estimation approaches, Approach 1 (APOCTT) generally reduces relative MBE and RMSE more effectively compared to CAMS, especially noticeable in stations like Izana, where this first approach reduces the relative RMSE from 21.34% to 14.17% . However, some stations such as University of Zalaland (UNZ) and Payerne show increased errors in relative MAE and RMSE with APOCTT, indicating that the performance of error reduction methods can vary significantly depending on the station. This examination of relative error metrics highlights the importance of station-specific analysis to optimize predictive accuracy and reliability of the parallax and shadow corrections.

Of course, there are key points here. It is important to note that the parallax and shadow effects are difficult to account for because of the nature of the main parameter needed for the correction: the cloud height. The main objective of these results is to assess whether the estimation of cloud top height (CTH) using cloud top temperature (CTH) data alongside standard atmospheric profile from climatological data is enough to tweak the current model and help it predict SSI more accurately. This requires many assumptions, mainly given by the fact that the atmospheric profiles are stationary and cloud formation is ever changing. Other assumptions are already described in chapter 3, but it is important to mention: the geometrical correction sees the cloud as a flat surface and therefore its projection as seen by the satellite is also flat. These projections are done on the image plane and does not yet include any digital elevation models (DEMs). This is relevant, since complex terrain, e.g. mountainous regions, seems to play a significant role in properly estimating the parallax and shadow displacements. This can be seen especially on the worst-performing station: Sonnblick, located in the alps with a high elevation.

A key point here is that the improved geometrical correction acts between the retrieval

of cloud information from APOLLO_NG and the application of Heliosat-4 to estimate SSI. Therefore, the improvement of the CRS model that includes geometrical correction (independent from the cloud height estimation approach) is subject to the limitations of both tools, but a deeper dive will be discussed later. But before this, the seasonal performance will be presented in Section 5.3.

5.3. Relative error metrics over shorter time periods

An important analysis to be seen in this study is the performance of the geometrical correction using the different cloud height approaches on shorter time frames within the yearly data used. For this, periodic performance of the relative error metrics is analyzed for the same stations, except Oujda (OUJ) and Magurele (INO). The reason being, these two stations have data available for the first and second half of 2021. These stations were taken out of this analysis. The results for the remaining stations were obtained both for the relative metrics previously discussed in Section 5.2. The data is divided into three-month periods such as JFM (January–February–March), AMJ (April–May–June), JAS (July–August–September), and OND (October–November–December). These will be referred in this work as seasons. Keep in mind that these are standard seasons used in climatology, meteorology, and other sectors, given the nature of the data analyzed, e.g. datasets for 2021 in this work.

Figure 5.2 shows the performance of the different approaches compared to the uncorrected CAMS model for the different stations along the different seasons. The presented plot provides a comprehensive view of the accumulative error metrics (MBE, MAE, and RMSE) across the four seasons (JFM, AMJ, JAS, and OND). Taking a look at these results, some temporal trends could be seen. The data reveals significant seasonal variations in the relative metrics across different periods, with some stations showing marked improvement or deterioration in specific seasons.

The results indicates that seasonal variability plays a crucial role in the performance of the parallax and shadow corrections using the different cloud height estimation approaches. Predictable solar conditions during Summer (Season JAS) lead to the smallest errors for all approaches. The different approaches show significant improvement in MBE for most stations during this season, although an interesting behavior is seen where MBE for Reunion Island shifts to positive values for the correction results, while the uncorrected CAMS model’s MBE is negative. Highly variable solar radiation during winter, coupled with frequent cloud cover, poses significant challenges. While the approaches reduce errors compared to the uncorrected CAMS model, RMSE values remain relatively high, highlighting the difficulty of accurately modeling GHI during this season. This season contributes to most of the negative MBE for the most troublesome of the stations (Reunion Island).

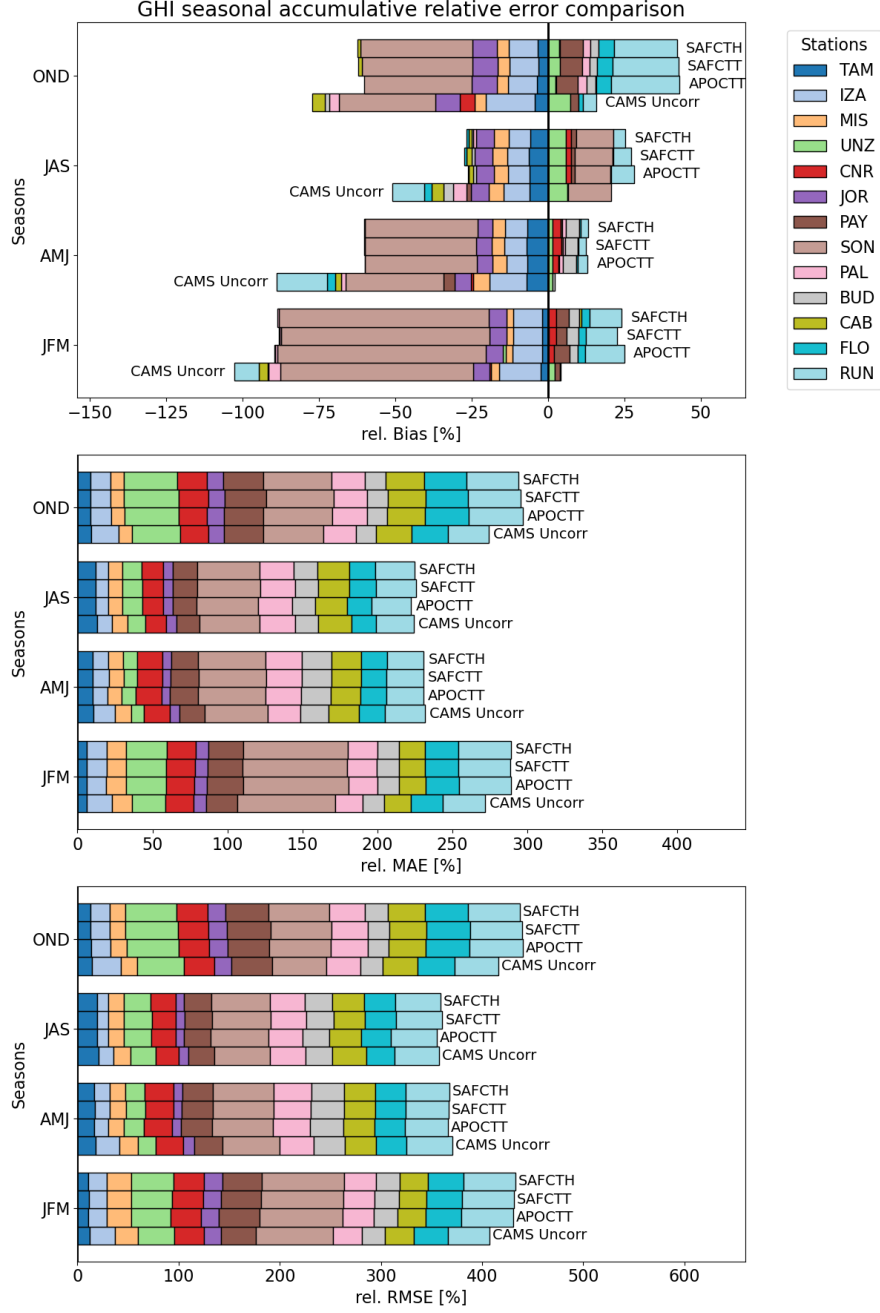


Figure 5.2.: Accumulative Relative Error Metrics per Season

The uncorrected model often shows a stronger negative bias, particularly in Winter and Spring. The performance gap between the uncorrected model and the approaches is most significant in Winter and Autumn, seasons with higher variability in cloud formation. Despite the higher variability of sky conditions during Spring, all correction approaches show improvement compared to the uncorrected model. A key insight is taken from this: the effectiveness of the correction approaches is highly season-dependent. The second approach (SAFCTT) consistently provides the best overall accuracy (MBE and MAE), while the last approach (SAFCTH) is more adept at reducing extreme deviations

(RMSE), particularly during the challenging Winter and Autumn seasons. This can be explained by taking a look at the probability distribution of Cloud Types for each station, shown in Figure 5.3.

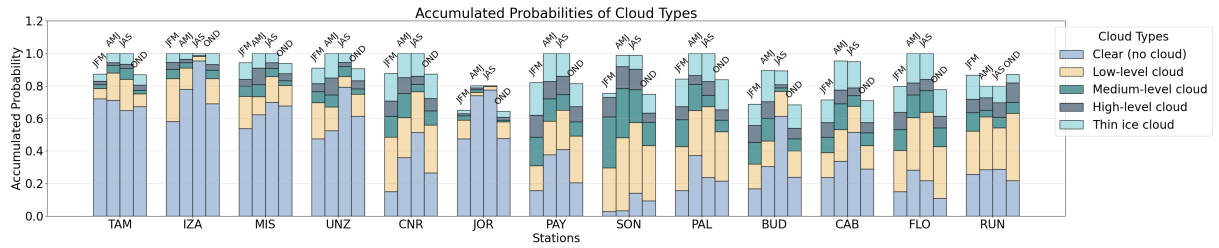


Figure 5.3.: Probability distribution of Cloud Type across seasons

Figure 5.3 is an important indicator of how the other parameters, besides Cloud Optical Depth (COD) or Clout Top Height (CTH), are vital for applying a good geometrical correction. Stations with higher probabilities of encountering different cloud types show more degradation. This highlights more how Reunion Island and Sonnblick tend to be the worst-performing, especially in Sonnblick, where clear skies have the smallest probability during all seasons. Another problem that might be discussed in Qu et al., 2017, where some clear-sky instants in early morning or late afternoon in winter can sometimes be mistaken as clouds. This means that when retrieving cloud information, the geometrical correction will move pixels where this situation happens, essentially moving parameters including Cloud Type, estimating SSI incorrectly and therefore increasing the errors during these conditions.

5.4. Correction performance

The previous analysis has given a optimistic outcome of what the correction methods can do to the current model, but as it was mentioned, there are some stations that did not show any significant improvement, in some cases showed worse results than the uncorrected model. The overall results of both the absolute and relative error metrics were used to classify the stations based on their performance.

Stations where the correction methods show improvement (reduction in error metrics) compared to the uncorrected CAMS model were classified as 'improved stations'. These stations demonstrate the positive impact of the improved geometrical correction, where some of the cloud estimation approaches perform better than others. A few other cases where the approaches show improvement in some of the metrics, but not all, were classified as 'partially improved stations'. Finally, there are some stations where the correction methods do not show any improvement in any of the absolute or relative error values compared to the uncorrected CAMS model, this were categorized as 'degraded stations'. The sections below summarize the percentage improvements (or lack thereof) for each station and method.

5.4.1. Improved Stations

The stations where the geometrical correction show general improvement in the estimation of SSI are: Tamanrasset (TAM), Izana (IZA), Missouri (MIS), Jordan (JOR), Tamanrasset (INO), and Cabauw (CAB). Their locations are shown in Figure 5.4

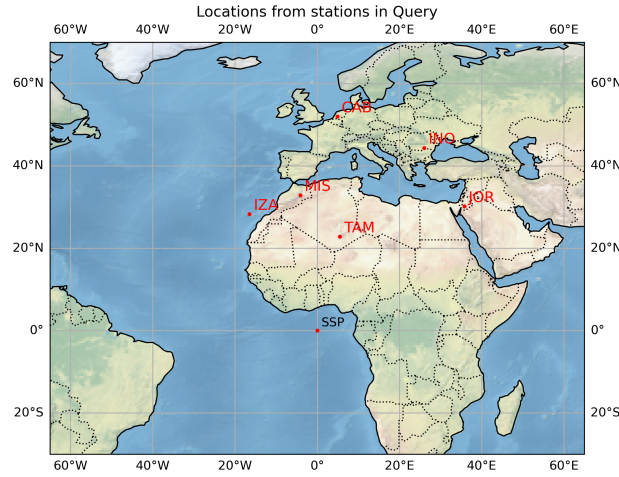


Figure 5.4.: Locations of best-performing stations

These stations exhibited positive improvements across all metrics. Figure 5.5 show the estimated mean percentage improvement of MBE, MAE, and RMSE for these stations. This category highlights the effectiveness of the correction methods applied to these locations, demonstrating consistent improvements across all error metrics, making these stations benchmarks for a potential successful implementation of this method in the current CAMS model.

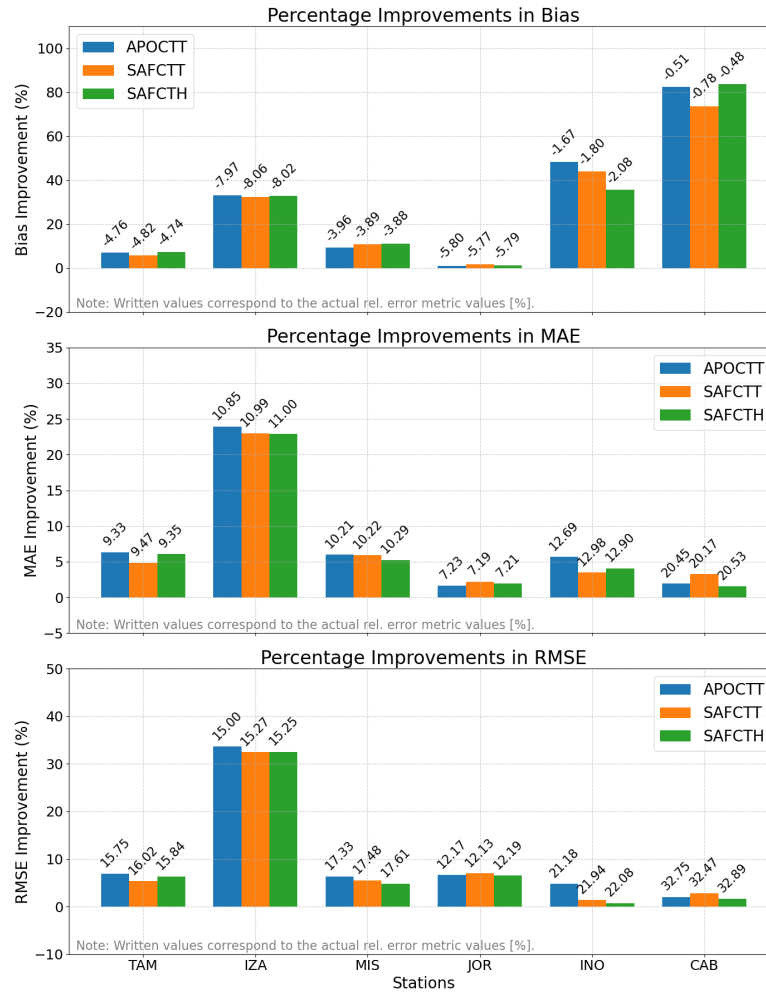


Figure 5.5.: Percentage improvement for 'Improved' stations

TAM station shows consistent improvements with APOCTT, achieving 6.98% in absolute and relative MBE, 6.32% in absolute and relative MAE, and 6.92% in absolute and relative RMSE. This looked promising, since the Tamanrasset station is located the closest to the sub-satellite-point (SSP) of the MSG-satellite. Izana had the highest improvements, particularly with APOCTT, where absolute and relative MBE improved by 33.18%, MAE by 23.95%, and RMSE by 33.64%. Missouri shows moderate but positive improvements across all metrics, with the highest being in MBE improvement: 11.17% with SAFCTH. Jordan station showed small but consistent improvements in all metrics, with the highest being in absolute and relative RMSE: 7.01% with the second approach (SAFCTT). Magurele station exhibits substantial improvements, especially in MBE with APOCTT showing a 48.22% improvement, while also showing positive gains in MAE and RMSE. Cabauw demonstrates high improvements in MBE, with SAFCTH showing 83.76%, along with positive MAE and RMSE improvements. This seems to indicate that the geometrical correction performs better the closer better closer to the SSP. Of course the parallax effect is less significant in this situation, but the shadow effect is present throughout all day, so the shadow correction is indeed acting favorably with respect to the uncorrected model.

5.4.2. Partially Improved Stations

The stations seen in Figure 5.6 where some improvements were observed—OUJ, UNZ and PAL—showed mixed results, with improvements in either RMSE or MBE, but not both. Figure 5.7 shows the results for this special category of stations. OUJ station showed a peculiar trend with degradation in MBE but positive improvements in MAE and RMSE. Despite the negative MBE improvements (-42.82% to -110.96%), it showed positive MAE improvements up to 9.06% and RMSE improvements up to 8.44% .

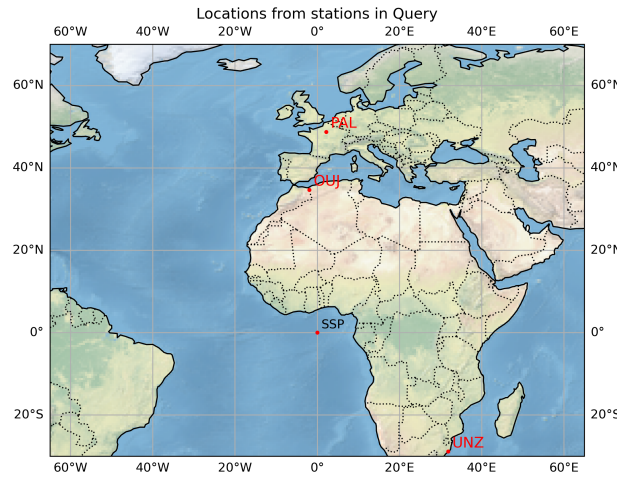


Figure 5.6.: Locations of 'Partially Improved' Stations

Oujda is part of the stations in the enerMENA network from libinsitu. Desert-type stations are particularly challenging for estimating SSI from CRS because of presence of desert dust aerosol and other factors that make desert stations have the greatest errors in Heliosat-4 (Qu et al., 2017). This is also an issue with the meteorological station itself, because a significant portion of ground observation data is flagged as unreliable, making a proper assessment of the improvement of the model more difficult.

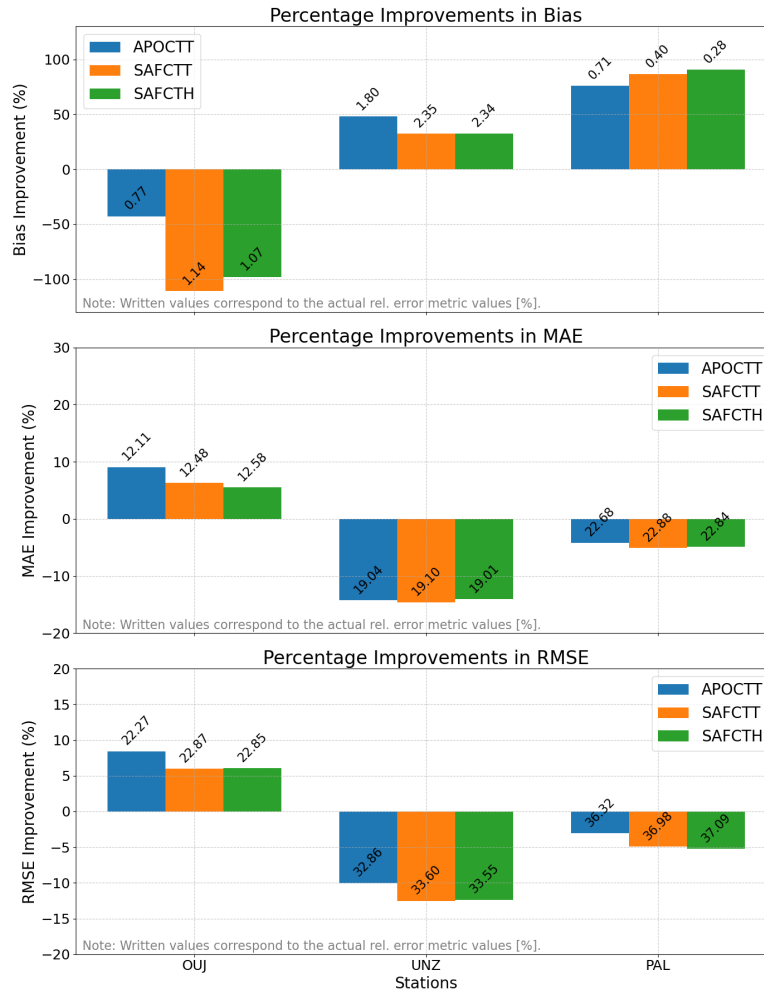


Figure 5.7.: Percentage improvement for OUJ, UNZ and PAL stations

UNZ exhibited significant improvements in MBE, particularly with APOCTT (48.16%), but showed negative improvements in MAE and RMSE, indicating an increase in errors in these metrics. PAL showed substantial improvements in MBE, up to 90.62% with SAFCTH, but had slight negative improvements in MAE and RMSE, indicating some inconsistencies in error reduction. These two stations are located far from the sub-satellite point, e.g. greater satellite viewing angles, where degradation of the quality in the retrieval of cloud properties is caused by large satellite viewing angles (Qu et al., 2017).

5.4.3. Degraded Stations

The worst-performing stations: CNR, PAY, SON, BUD, FLO, and RUN, exhibited degradation across all metrics (see Figure 5.9). The last two were a particular case that needed to be tested, because the initial postulated theory, was that the bigger the distance from the sub-satellite-point (SSP) would benefit more from the parallax corrections. Although this might be true for locations like CAB station, which showed a significant reduced MBE and an improvement in MAE and RMSE, the geometrical

correction explained in this study seems to be beneficial up to a point where Earth's curvature does not penalize the geometrical scheme described in detail in Chapter 3. Furthermore, just as with the previous group, greater satellite viewing angles exhibit greater errors because of degradation of the quality of cloud properties obtained by APOLLO_NG.

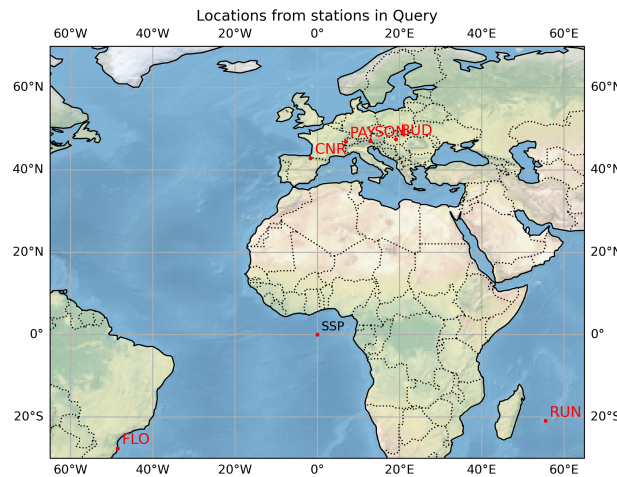


Figure 5.8.: Locations of 'degraded' stations

CNR shows the most significant decrement, with MBE deteriorating by up to -124.82% on SAFCTT) and similar negative trends in MAE and RMSE. PAY also exhibits significant degradation, with MBE worsening by -64.96% for SAFCTH as the worst case) and negative changes in MAE and RMSE, indicating overall poor performance. SON has consistently poor improvements, with MBE changes up to -15.94%) and similarly negative trends in MAE and RMSE. BUD shows substantial decline across all metrics, with MBE deteriorating between -102.59% and -132.63% for all methods) and similar negative trends in MAE and RMSE. FLO showed degradation across all metrics, with MBE worsening by -96.31% , MAE by -7.32% , and RMSE by -6.10%). RUN demonstrate deterioration, with MBE changing by -53.10% with the first approach (APOCTT) as the worst case and -24.90% with SAFCTH as a best case, and MAE and RMSE also showing negative trends.

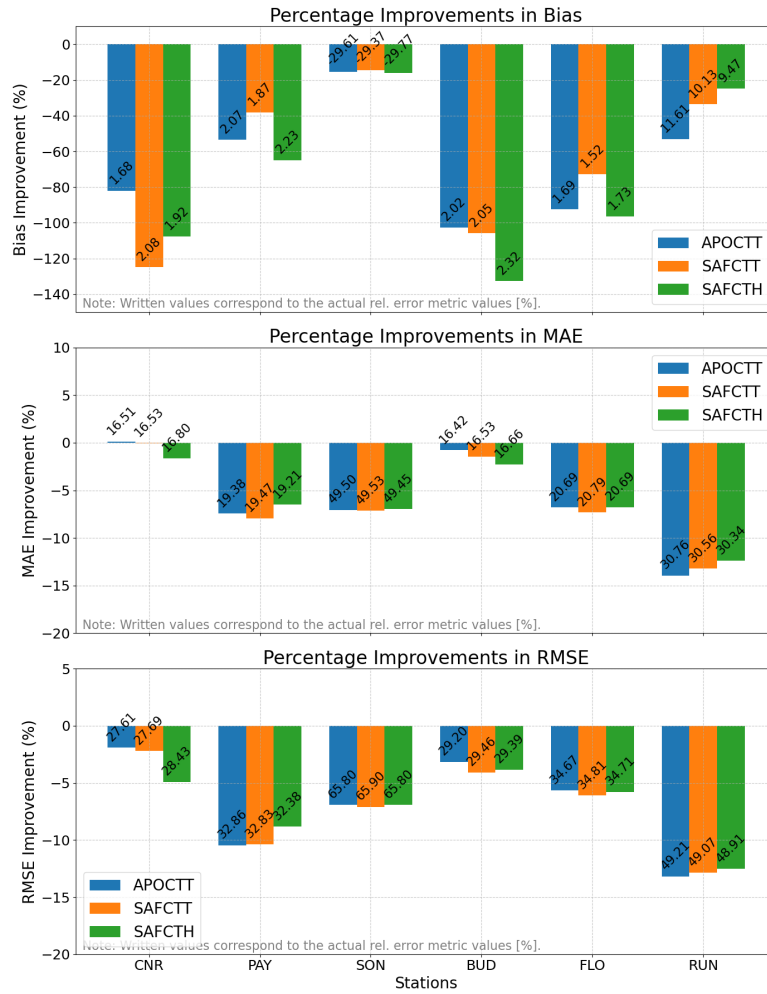


Figure 5.9.: Percentage improvement for degraded stations

The locations of these stations are shown in Figure 5.8. Florianopolis and Reunion Island are located almost on the edge of the field-of-view of the satellite. This affects the geometrical correction, because the curvature of the Earth becomes significant, more clearly seen in Figure 5.10. Another remark is the fact that because of the significance of the planet's curvature, clouds are no longer seen directly from the top, but partially from the side. This has two distinctive effects: (i) the cloud optical depth (COD) is detected differently by the satellite. (ii) estimation of the cloud top height (CTH) will be over or underestimated depending on the thickness of the cloud.

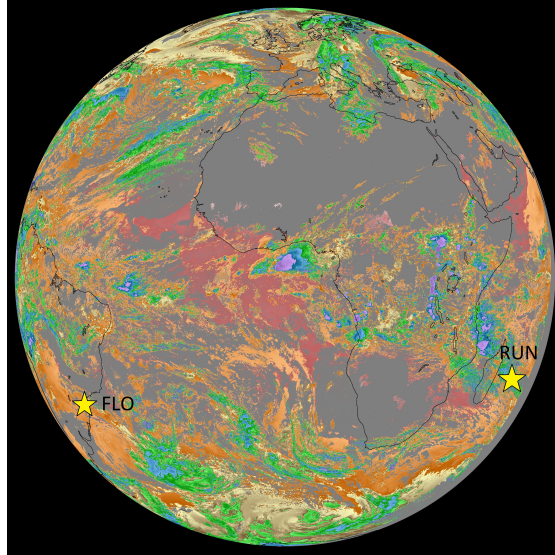


Figure 5.10.: Locations of FLO and RUN stations

This can also explain the shift in the direction of MBE for both cases, as it goes from underestimation for the current CAMS model for FLO and RUN to an overestimation, more significantly for the latter.

5.4.4. General performance of the correction methods

Now that the performance of the parallax and shadow correction for all cloud estimation approaches were analyzed for each station separately which provided key insights of the capabilities for different locations, the average performance from all correction methods across all stations is estimated for the full dataset (yearly data for 2021). The percentage improvement for all methods is shown in Table 5.2 below.

	Approach1	Approach2	Approach3
MBE	0.7085	-8.2123	-9.5285
MAE	1.6115	0.9831	0.8469
RMSE	2.5538	1.4838	1.2708

Table 5.2.: Average performance metrics for all correction methods (full data)

Approach 1 (APOCTT) exhibit an average minor improvement of the MBE (0.7085%, demonstrating that this method slightly reduces the systematic error compared to the uncorrected CAMS method. the other approaches showed a significant negative values of the improvement, meaning an increased MBE with respect to CAMS, Approach 3 (SAFCTH) being the most substantial (-9.5285%). These results indicate that while APOCTT offers a slight improvement, while the other approaches (SAFCTT and SAFCTH) lead to a considerable worsening of MBE, increasing systematic errors rather than correcting them.

APOCTT results in a modest MAE improvement of 1.61%, indicating a slight reduction in the average error magnitude, while the other two methods shows smaller

improvements (0.98% and 0.85% respectively), reflecting a smaller reduction in the average error. These findings suggest that while all methods offer some improvement in MAE, APOCTT leads in reducing the average error magnitude, with Methods 2 and 3 provide considerably less improvement.

Lastly the relative RMSE is a critical metric, particularly useful for understanding the consistency of the methods in handling large errors. APOCTT shows the highest improvement (2.55%), indicating a modest reduction in the overall error variance. SAFCTT shows an intermediate improvement (1.48%), reflecting a smaller reduction in larger errors. Finally, SAFCTH demonstrates the smallest improvement with 1.27%, suggesting it is the less effective in minimizing the impact of larger errors. These results show that all methods contribute to reducing larger errors, with APOCTT is in all cases the best option for potentially implement the satellite and shadow correction methods into the current CAMS model. This is promising, since using this method, ensures the use of more reliable data for irradiance estimation, as it does not use external datasets for the calculation or retrieval of cloud top height (CTH), since the cloud top temperature (CTT) is within the operating chain of CAMS at DLR. Overall, APOCTT shows a better improvement for the 'good' category stations and less degradation among the worst results.

Another key point is the limitation of the CRS model, both for the cloud detection algorithms from APOLLO_NG and estimation of SSI by the Heliosat Method. There can be several reasons for these limitations. This can include inaccuracies due to different factors, e.g. limited spectral bandwidth of the instruments from the satellite, available aerosol and cloud properties, or degradation of the quality in the retrieval of cloud properties is caused by large satellite viewing angles. Other important factors can be underestimation of aerosol optical depth (AOD) in the presence of desert dust aerosol (Qu et al., 2017). These limitations of the CRS model itself affect many of the stations used in this work, some more than others.

5.5. Possible reasons for the lack of improvement

One of the main limitations of the geometrical corrections is that it cannot account for the variability in elevations of each pixel. The geometrical corrections (independently from the cloud height estimation approach) will shift the surrounding pixels positions differently. This is because the geometrical correction exposed in Chapter 3, as well as it is described in Beyer et al., 1996 and Roy et al., 2023, cannot project the apparent and shadow displacements for complex terrains. The geometrical correction shifts the pixels within the image plain. Cloud top height is estimated with respect to the mean sea level, since all estimation approaches use standard atmospheric profiles. The problem arises when the cloud pixel is located in regions where there are significant elevation differences around the clouded pixels, that will affect the projection due to parallax or shadow effects. The displacement of pixels can therefore be over or underestimated in these situations, essentially changing their calculation using equations 3.1 and 3.6. Four possible scenarios can be seen in Figure 5.11. These are naturally simplistic illustrations of the possible scenarios that can occur because of the elevation differences.

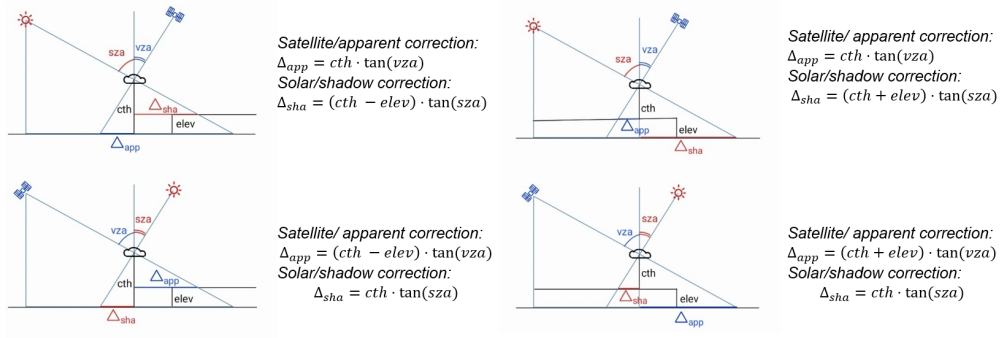


Figure 5.11.: Schematic representation of limitations of the geometrical correction

A clear illustration is shown in Figure 5.12 for a specific time step as an example. This figure shows the locations of the original gridded points, namely the apparent cloud positions (in blue) and their final cloud locations after the correction has been made (in red). The maximum displacement is around $70km$. This is a significant distance for considerable changes in elevations, which is the reason why the improved geometrical correction is performed with $17pixels \times 17pixels$, enough to cover a good amount of area for displacements to be calculated. And this also illustrates the issue with complex terrain. Thus, complex terrain around the locations of some of the degraded station was investigated.

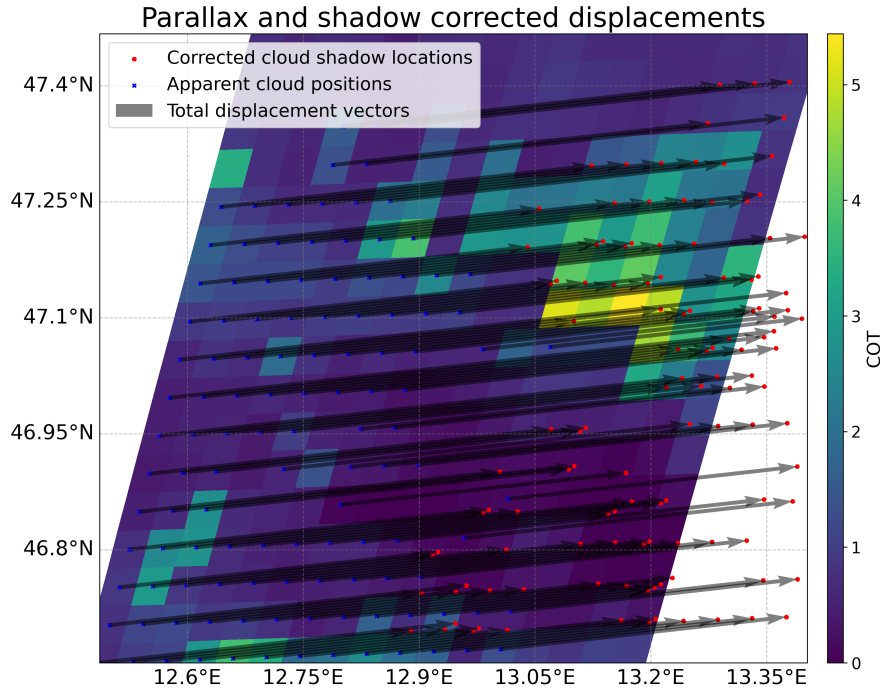


Figure 5.12.: Total parallax and shadow displacements for Sonnblick Station [Times-tamp: 2021-10-12T15:00]

The northwest-side of Budapest, where BUD station is located has a ridge which

can be changed the cloud height on the pixels taken near this station (See Figure 5.13). This could affect the estimation of shifted distances which in turn account for a bad approximation of the interpolation of Cloud Optical Thickness (COT) for this station, and in turn the Global Horizontal Irradiance (GHI) for all correction methods.

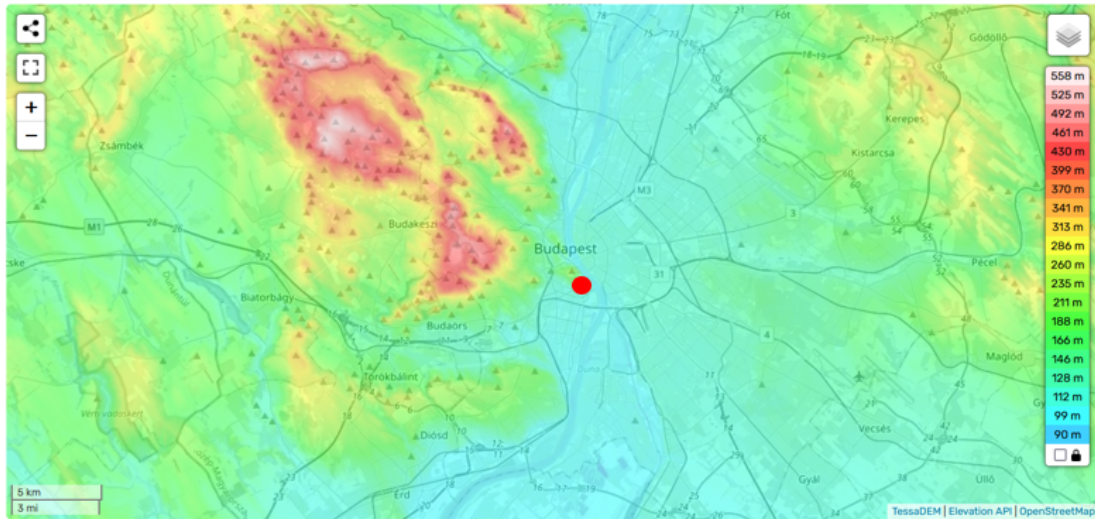


Figure 5.13.: Topographic map of BUD station

The region around Cener station (CNR) located near Pamplona, Spain, as seen in Figure 5.14, shows a slightly more complex terrain. It has an elevation of 471.m (see Table 4.2) and is surrounded by mountains in all directions.

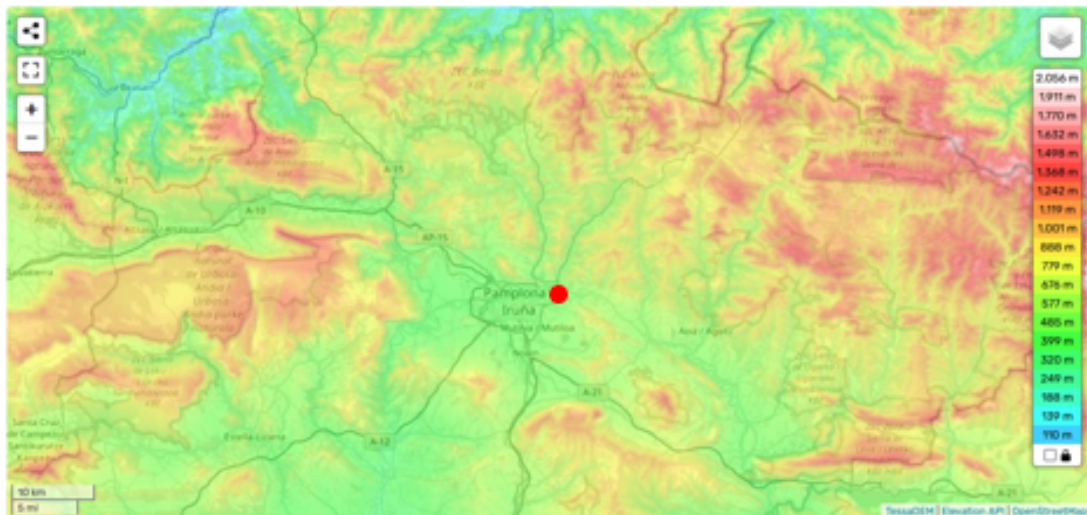


Figure 5.14.: Topographic map of CNR station

Payerne station (PAY) is also located on a moderate elevation (491.m, see Table 4.2) on a slope down a mountainous region to the south-east (See Figure 5.15). This can affect the shadow corrections during early morning, as the sun is to the east.

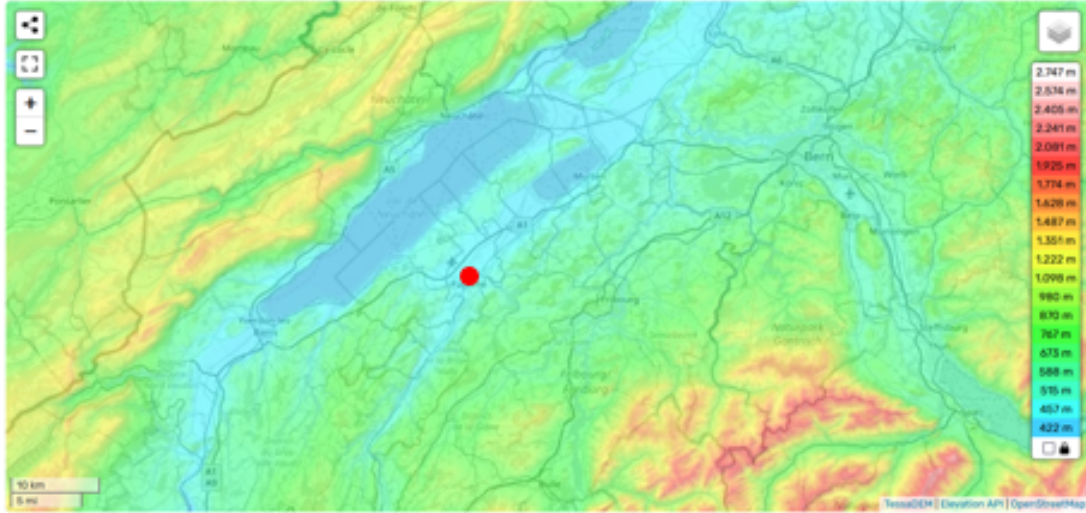


Figure 5.15.: Topographic map of PAY station

The Sonnblick station is located 3108.9 meters above sea level, making it the station with the highest elevation. Additionally, it has one of the most complex terrain, as it can be seen in Figure 5.16. The displacements in Figure 5.12 is from exactly this region. It is worth noting that the Izana station has also a high elevation (2372m), although the conditions surrounding the station are completely different, as Izana station is located on an island, while Sonnblick has far more intricate geography. A good hypothesis could be the mixture of high elevation and complex terrain what makes Sonnblick one of the worst-performing stations.



Figure 5.16.: Topographic map of SON station

An important point is made here. To further improve the model, a Digital Elevation Model (DEM) could be integrated into the algorithm, to account for the elevation differences between the non-shifted pixel and its surrounding, but requires a deeper discussion

for future implementation.

5.6. Final remarks

This study has provided a comprehensive analysis of the performance of the three different cloud estimation approaches to apply parallax and shadow correction, for the possible implementation of the current CAMS model. By examining the performance of these approaches across multiple stations and different time periods throughout the year 2021, valuable insights were noticed on how these corrections behave under various seasonal conditions and how they compare to each other. Optimistic results could be observed, which highlights the effect parallax and shadow effects have on satellite-based irradiance models. Thus, these improvements are quite small, so there is no huge influence in the actual operations of CAMS for the moment. Nonetheless, since the improved geometrical correction was performed on low-resolution gridded-data ($3km \times 3km$ nominal pixel resolution), a more significant positive impact of the improved parallax and shadow corrections on higher pixel resolution is to be expected.

5.6.1. Key Findings

The analysis revealed that while the uncorrected CAMS model serves as a reliable baseline, implementing the parallax and shadow correction methods exposed can have a positive impact, as it exhibits some improvement in the estimation of Global Horizontal Irradiance (GHI) compared to the current CAMS model. However, these implementation has limitations: the geometrical approach detailed in Section 3, similar to those of Beyer et al., 1996, and further developed by Roy et al., 2023, is more effective for pixels near the sub satellite point (SSP). This is evident on the stations that have a distance lower than 30000 km from the SSP, where all display significant improvement with respect to the uncorrected CAMS model. The outer-most stations, where Earth's curvature plays a significant roll from the perspective of the satellite, show a non-favorable results. A possible solution could be the implementation of a spherical trigonometry correction method like that of Bieliński, 2020. Other errors on the correction methods might be due to the complex terrain around those locations, which can under- or overestimate GHI, caused by the wrong projection of the parallax and shadow effects on that terrain. A digital elevation model (DEM) could be implemented to assess this limitation. Using higher-resolution maps from the latest generation satellites (3rd generation) might also help improve GHI estimation on more refined image grid.

The study also highlighted the significant impact of seasonal variations on model performance. In general, the geometrical correction (using any of the cloud height estimation approach) tended to perform better during summer and spring months, with lower MBE and RMSE, while the winter months posed greater challenges. This seasonal sensitivity suggests that any future development of these models should consider incorporating seasonal adjustments to improve accuracy. Moreover, the consistent under-performance at certain stations like SON and CNR indicates that local environmental factors, may be influencing the correction models. Understanding these local effects could be crucial for refining the models further.

Specific adjustments could be implemented, e.g. by setting a maximum range to

apply the correction. By analyzing unique terrain and environmental characteristics of each station, models could be tailored to implement corrections in locations where they currently struggle. This approach would likely involve a combination of empirical adjustments and machine learning techniques to identify and correct for local anomalies.

Finally, further validation and testing of these correction methods across a wider range of geographic locations and longer time periods would be beneficial. Expanding the dataset to include longer temporal spans would help ensure that the correction model is robust across different conditions.

5.6.2. Concluding Thoughts

This study has provided a solid foundation for understanding parallax and shadow effects on satellite-based irradiation modeling. It also helps understand the strengths and weaknesses of the methods presented. The insights gained from this analysis will be valuable in guiding the development of more accurate and reliable correction methods. By focusing on seasonal adjustments, station-specific refinements, and thorough validation, future models can build on the progress made here, ultimately contributing to more precise solar radiation forecasting and better-informed decision-making in fields such as renewable energy and climate science.

6. References

- Anderson, G., Clough, S., Kneizys, F., Chetwynd, J., & Shettle, E. (1986). Aflg atmospheric constituent profiles (0.120km), 46.
- Beyer, H. G., Costanzo, C., & Heinemann, D. (1996). Modifications of the heliosat procedure for irradiance estimates from satellite images. *Solar Energy*, 56(3), 207–212. [https://doi.org/10.1016/0038-092X\(95\)00092-6](https://doi.org/10.1016/0038-092X(95)00092-6)
- Bieliński, T. (2020). A parallax shift effect correction based on cloud height for geostationary satellites and radar observations. *Remote Sensing*, 12(3), 365. <https://doi.org/10.3390/rs12030365>
- Blaga, R., Calinoiu, D., Stefu, N., Boata, R., Sabadus, A., Paulescu, E., Pop, N., Mares, O., Bojin, S., & Paulescu, M. (2021). Quantification of the aerosol-induced errors in solar irradiance modeling. *Meteorology and Atmospheric Physics*, 133(4), 1395–1407. <https://doi.org/10.1007/s00703-021-00815-z>
- Cano, D., Monget, J. M., Albuissou, M., Guillard, H., Regas, N., & Wald, L. (1986). A method for the determination of the global solar radiation from meteorological satellite data. *Solar Energy*, 37(1), 31–39. [https://doi.org/10.1016/0038-092X\(86\)90104-0](https://doi.org/10.1016/0038-092X(86)90104-0)
- Cebecauer, T., Šúri, M., & Gueymard, C. (2011). Uncertainty sources in satellite-derived direct normal irradiance: How can prediction accuracy be improved globally, 20–23.
- Ceilometer. (2024). <https://www.lufft.com>
- Copernicus Atmosphere Monitoring Service. (2024). Solar radiation supplementary products [Accessed: 2024-05-08].
- Dagestad, K.-F. (2005). Estimating global radiation at ground level estimating global radiation at ground level from satellite images. *Doctor Scientiarum Thesis in Meteorology*.
- Damiani, A., Irie, H., Horio, T., Takamura, T., Khatri, P., Takenaka, H., Nagao, T., Nakajima, T. Y., & Cordero, R. R. (2018). Evaluation of himawari-8 surface downwelling solar radiation by ground-based measurements. *Atmospheric Measurement Techniques*, 11(4), 2501–2521. <https://doi.org/10.5194/amt-11-2501-2018>
- DLR Institute of Networked Energy Systems. (2021). Eye2Sky cloud camera network. *DLR Institute of Networked Energy Systems*. https://www.dlr.de/ve/en/desktopdefault.aspx/tabid-3780/6257_read-53178/
- Dribssa, E., Coglian, E., Lavagno, E., & Petrarca, S. (1999). A modification of the heliosat method to improve its performance. *Solar Energy*, 65(6), 369–377. [https://doi.org/10.1016/S0038-092X\(99\)00014-6](https://doi.org/10.1016/S0038-092X(99)00014-6)
- Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., Denn, F. M., Duprat, T., Fukuda, M., Grobe, H., Haeffelin, M., Hodges, G., Hyett, N., Ijima, O., Kallis, A., Knap, W., Kustov, V., Long, C. N., Longenecker, D., ... König-Langlo, G. (2018). Baseline surface radiation network (bsrn): Structure and data description (1992–2017). *Earth System Science Data*, 10(3), 1491–1501. <https://doi.org/10.5194/essd-10-1491-2018>

-
- Dybbroe, A., Karlsson, K.-G., & Thoss, A. (2005). Nwcsaf avhrr cloud detection and analysis using dynamic thresholds and radiative transfer modeling. part i: Algorithm description. *Journal of Applied Meteorology*, 44(1), 39–54. <https://doi.org/10.1175/JAM-2188.1>
- EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF). (2024). Cmsaf operational products [Accessed: 2024-10-29].
- Forstinger, A., Wilbert, S., Jensen, A., & Kraas, B. (2023). Iea-pvps worldwide benchmark of modelled solar irradiance data.
- Geuder, N., Wolfertstetter, F., Wilbert, S., Schüler, D., Affolter, R., Kraas, B., Lüpfer, E., & Espinar, B. (2015). Screening and flagging of solar irradiation and ancillary meteorological data. *Energy Procedia*, 69, 1989–1998. <https://doi.org/10.1016/j.egypro.2015.03.205>
- Golombek, J. (2010). Satellite analema plotter. Retrieved September 28, 2023, from <http://www.golombek.com/analema.html>
- Hammer, A., Kühnert, J., Weinreich, K., & Lorenz, E. (2015). Short-term forecasting of surface solar irradiance based on meteosat-seviri data using a nighttime cloud index. *Remote Sensing*, 7(7), 9070–9090. <https://doi.org/10.3390/rs70709070>
- Herzog, J., Fiedler, H., Schildknecht, T., & Hinze, A. (2018). Using conjunction analysis methods for manoeuvre detection – application to optical observations. <https://doi.org/10.7892/boris.126436>
- In Situ Energy Analysis [Accessed: 2024-02-19]. (n.d.).
- Klüser, L., Killius, N., & Gesell, G. (2015). Apollo_ng – a probabilistic interpretation of the apollo legacy for avhrr heritage channels. *Atmospheric Measurement Techniques*, 8(10), 4155–4170. <https://doi.org/10.5194/amt-8-4155-2015>
- Kuhn, P., Wilbert, S., Prah, C., Garsche, D., Schüler, D., Haase, T., Ramirez, L., Zarzalejo, L., Meyer, A., Blanc, P., & Pitz-Paal, R. (2018). Applications of a shadow camera system for energy meteorology. *Advances in Science and Research*, 15, 11–14. <https://doi.org/10.5194/asr-15-11-2018>
- Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Schroedter-Homscheidt, M., Hoyer-Klick, C., Arola, A., Benedetti, A., Kaiser, J. W., & Morcrette, J.-J. (2013). McClear: A new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmospheric Measurement Techniques*, 6(9), 2403–2418. <https://doi.org/10.5194/amt-6-2403-2013>
- Lyu, B., Chen, Y., Guan, Y., Gao, T., & Liu, J. (2021). Estimating geo-referenced cloud-base height with whole-sky imagers. *Earth and Space Science*, 8(8). <https://doi.org/10.1029/2019EA000944>
- Meirink, J. F., Roebeling, R. A., & Stammes, P. (2013). Inter-calibration of polar imager solar channels using seviri. *Atmospheric Measurement Techniques*, 6(9), 2495–2508. <https://doi.org/10.5194/amt-6-2495-2013>
- Miller, S. D., Rogers, M. A., Haynes, J. M., Sengupta, M., & Heidinger, A. K. (2018). Short-term solar irradiance forecasting via satellite/model coupling. *Solar Energy*, 168, 102–117. <https://doi.org/10.1016/j.solener.2017.11.049>
- Pagola, I., Gastón, M., Bernardos, A., & Fernández-Peruchena, C. (2014). A combination of heliosat-1 and heliosat-2 methods for deriving solar radiation from satellite images. *Energy Procedia*, 57, 1037–1043. <https://doi.org/10.1016/j.egypro.2014.10.088>
- Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., & Hoff, T. E. (2010). Validation of short and medium term operational solar radiation forecasts in the

-
- us. *Solar Energy*, 84(12), 2161–2172. <https://doi.org/10.1016/j.solener.2010.08.014>
- Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., Klüser, L., Lefèvre, M., Saboret, L., Schroedter-Homscheidt, M., & Wald, L. (2017). Fast radiative transfer parameterisation for assessing the surface solar irradiance: The heliosat-4 method. *Meteorologische Zeitschrift*, 26(1), 33–57. <https://doi.org/10.1127/metz/2016/0781>
- Rigollier, C., Lefèvre, M., & Wald, L. (2004). The method heliosat-2 for deriving short-wave solar radiation from satellite images. *Solar Energy*, 77(2), 159–169. <https://doi.org/10.1016/j.solener.2004.04.017>
- Roy, A., Hammer, A., Heinemann, D., Schroedter-Homscheidt, M., Lünsdorf, O., & Lezaca, J. (2023). *Revisiting parallax and cloud shadow correction in satellite-based solar irradiance estimation*. <https://doi.org/10.2139/ssrn.4630664>
- Schmidt, T., Stührenberg, J., Blum, N., Lezaca, J., Hammer, A., & Vogt, T. (2022). A network of all sky imagers (asi) enabling accurate and high-resolution very short-term forecasts of solar irradiance. *21st Wind & Solar Integration Workshop (WIW 2022)*, 372–378. <https://doi.org/10.1049/icp.2022.2778>
- Schroedter-Homscheidt, M., Azam, F., Betcke, J., Hanrieder, N., Lefèvre, M., Saboret, L., & Saint-Drenan, Y. (2022). Surface solar irradiation retrieval from msg/seviri based on apollo next generation and heliosat-4 methods. *Meteorologische Zeitschrift*, 31(6), 455–476. <https://doi.org/10.1127/metz/2022/1132>
- Schroedter-Homscheidt, M., Betcke, J., Breitkreuz, H.-K., Hammer, A., Heinemann, D., Heinicke, S., Lorenz, E., Petrak, S., Holzer-Popp, T., & Wald, L. (2000). Energy-specific solar radiation data from msg: The heliosat-3 project. *Proceedings of the 3rd MSG RAO Workshop*, 52–57.
- Topographic map [Accessed: 2024-08-07]. (n.d.).

A. Haversine formula

The Haversine formula is used to calculate the distance d between two points with latitude and longitude on the surface of a sphere. The formula is given by:

$$d = R \cdot c \quad (\text{A.1})$$

where:

R is the radius of the Earth (mean radius is commonly used, $R \approx 6371.0$ km), c is the central angle between the two points, calculated using the Haversine function:

$$\text{haversine}(\theta) = \sin^2\left(\frac{\theta}{2}\right) \quad (\text{A.2})$$

The central angle θ is determined by the latitudes ϕ_1 and ϕ_2 and longitudes λ_1 and λ_2 of the two points:

$$\theta = \arcsin\left(\sqrt{\sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2\left(\frac{\Delta\lambda}{2}\right)}\right) \quad (\text{A.3})$$

where:

$\Delta\phi = \phi_2 - \phi_1$ is the difference in latitudes, $\Delta\lambda = \lambda_2 - \lambda_1$ is the difference in longitudes.

The distances calculated using the Haversine formula are approximate and are based on a spherical Earth model. For more accurate results, especially for long distances, more complex formulas like the Vincenty formulae are often used.

B. Implementation

The algorithm is divided in four main phases:

1. Data collection
2. Data processing
3. Application of Heliosat-4
4. Validation

B.1. Data collection

The datasets used for the data collection phase are more clearly described in Chapter 4. Using the *libgndobs* library, a query is ran for the time range and stations required. Processing the query retrieves the ground measurement data from each station selected and the same timerange and the location of each station (stored in the metadata of each file, i.e. the latitude-longitude points of each station) is used to download the corresponding CAMS data of those locations.

Using the time range, i.e. 2021 for this study, APOLLO_NG map files were prepared every two hours. Map files contain primary variables needed (COT, Cloud Type, coverage) and location for center of each pixel. Missing timestamps (called **dts**) in Figure B.1) are filtered out in all datasets. There can be some files that were not available, so missing timestamps are filtered out. For each of the timestamps within the time range, satellite image data from APOLLO_NG is read, where the gridded pixel-wise COT and cloud properties are stored.

Additionally, a scan time must be added to properly fit the parameters needed for the cloud correction for running Heliosat-4.

When a satellite captures an image of the Earth, the timestamp of the image is recorded at start of scan (timestamp of line 0). The actual timestamps of the following lines will have an specific timestamp offset with respect to the image timestamp. This is because the satellite records the reflected light for each pixel corresponding to each region of the Earth at slightly different times as it progresses. It measures the counts of the reflectivity from the pixels of the downside of the image (southernmost pixel-lines) upwards. As a result, the image is not a simultaneous snapshot but rather a composite of data points collected over a span of time. Each segment of the image corresponds to a specific time when that particular section was scanned by the satellite, leading to a temporal gradient from the bottom to the top of the image. This offset needs to be calculated to get real pixel timestamp and correct variable estimations.

B.2. Data processing

The ground observation data was used as the reference of "true" values for comparing the uncorrected CAMS data with each of the cloud correction approaches done in this work.

The Cloud Top Temperature (CTT) data is used alongside the Atmospheric temperature profiles to estimate the Cloud Top Height (CTH). Before this calculation, a linear interpolation of CTT values was performed, to ensure that the CTT for each timestamp and pixel is within the Troposphere, where clouds will be located. This ensures that a reasonable CTH is extracted for the Atmospheric profile curves used depending on the region and time-of-year (see Figures 3.5 and 4.3).

The satellite zenith and azimuth angles were estimated using equations 3.4 and 3.5. The parallax correction is ran first, in order to find the actual location of the cloud (see Figure 3.1) using the Haversine method described in Annex A. The solar angles of the new shifted pixels were calculated using the *esy.solar* library from DLR and the haversine method is used once more. Finally, the final location points and their shifted parameters (Cloud Optical Thickness (COT), Cloud coverage and cloud type) were used to interpolate the values at the exact locations of the stations. This data is then stored as corrected map data, needed for the next phase.

B.3. Application of Heliosat-4

In order to run the Heliosat-4 method, Heliosat-4 is ran. Two main files were needed for this tool: a config file with the information of the paths of the corresponding input and output files, value of solar constant in W/m^2 and coefficients for circumsolar pyrheliometer adaptation, and the time series input file in CAMS format. The CAMS data downloaded previously for each station (see Section 4) was copied as an input file and the shifted parameters were replaced in the corresponding timestamps of each file. The tool outputs the parallax/shadow corrected Global Horizontal Irradiance (GHI) data, which was validated in the next phase.

B.4. Validation

Each of the correction approaches must be run and GHI results must be stored in their corresponding h5 files. For each of the stations and approaches, absolute and relative error metrics (see section 5.1) are compared against the uncorrected CAMS GHI results. The metrics are plotted and stored in separate files.

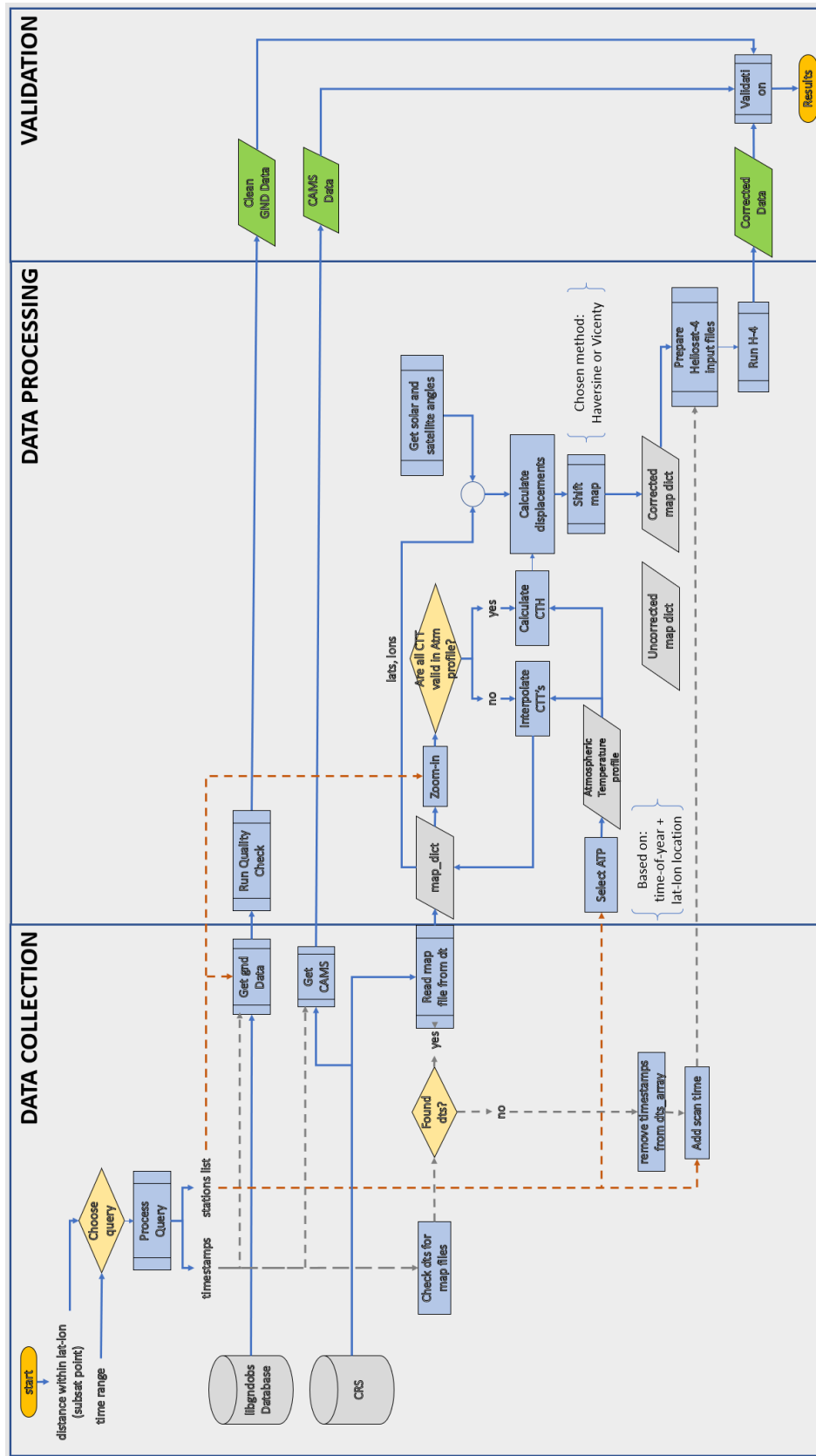


Figure B.1.: Overview of the implementation algorithm

C. Overall absolute error metrics

Absolute error metrics provide critical insight into the accuracy of the CAMS satellite-based model and the performance of the corrections methods of this work. These metrics show the real scale of the errors from the uncorrected CAMS model and each of the correction models developed, giving a clear indication of how far off the predictions are from the actual values on average. Figure C.1 shows the absolute errors for each station using the uncorrected model (CAMS) and three correction approaches. Table C.1 summarizes this metrics in numerical form.

Table C.1.: Absolute Metrics for 2021 Data

Stations	Absolute MBE (W/m^2)				Absolute MAE (W/m^2)				Absolute RMSE (W/m^2)			
	CAMS	APOCTT	SAFCTT	SAFCTH	CAMS	APOCTT	SAFCTT	SAFCTH	CAMS	APOCTT	SAFCTT	SAFCTH
TAM	-34.26	-31.87	-32.27	-31.74	66.71	62.49	63.48	62.66	113.36	105.52	107.29	106.13
IZA	-81.88	-54.71	-55.32	-55.03	97.94	74.49	75.41	75.51	155.18	102.98	104.82	104.69
MIS	-25.58	-23.22	-22.82	-22.72	63.66	59.84	59.90	60.31	108.46	101.60	102.48	103.20
OUJ	-2.59	3.70	5.47	5.13	63.97	58.18	59.96	60.44	116.83	106.97	109.86	109.76
UNZ	13.92	7.21	9.41	9.37	66.85	76.39	76.61	76.23	119.78	131.80	134.76	134.58
CNR	-3.98	7.24	8.94	8.26	71.05	70.97	71.08	72.21	116.48	118.71	119.05	122.20
JOR	-38.06	-37.64	-37.41	-37.57	47.66	46.88	46.60	46.73	84.57	78.94	78.64	79.07
PAY	-5.21	8.00	7.20	8.60	69.60	74.79	75.12	74.12	114.78	126.79	126.69	124.92
SON	-102.14	-117.79	-116.85	-118.42	183.93	196.92	197.05	196.73	244.79	261.78	262.16	261.79
PAL	-10.54	2.54	1.43	0.99	77.32	80.54	81.27	81.12	125.14	128.99	131.31	131.70
INO	-14.54	-7.53	-8.13	-9.36	60.64	57.20	58.51	58.17	100.31	95.51	98.92	99.56
BUD	-3.90	7.91	8.03	9.08	63.77	64.26	64.69	65.21	110.76	114.27	115.30	115.01
CAB	-9.61	-1.67	-2.53	-1.56	67.82	66.52	65.61	66.77	108.71	106.52	105.63	106.97
FLO	-3.83	7.37	6.61	7.51	84.26	89.97	90.43	89.96	142.70	150.78	151.40	150.95
RUN	-32.40	49.61	43.27	40.47	115.33	131.44	130.57	129.62	185.73	210.23	209.63	208.98

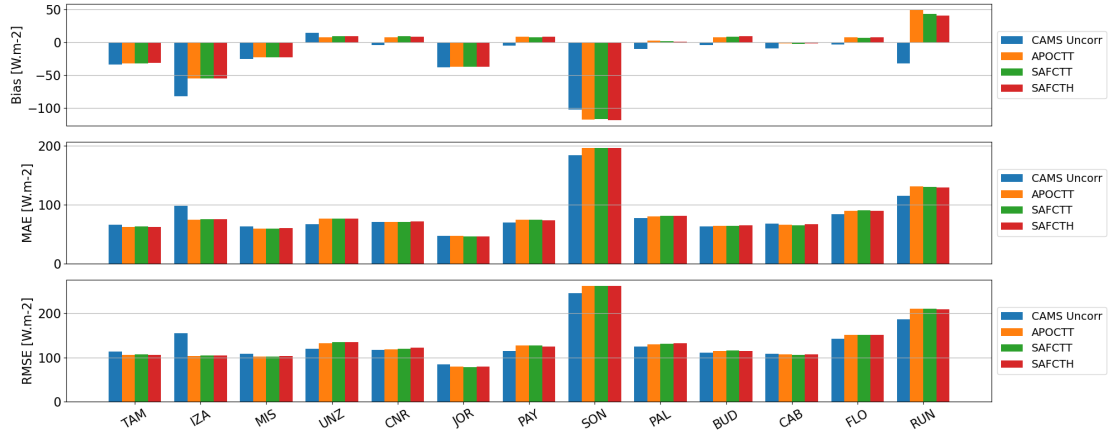


Figure C.1.: GHI absolute error comparison for 2021 data

Distinct patterns of performance across different stations and methods are observed. Sonnblick station (SON) exhibits the largest errors across all absolute metrics and methods, with an absolute MBE reaching $-118.42W/m^2$, and a substantial absolute RMSE of $262.16W/m^2$, indicating significant inaccuracies for this particular station. Conversely, Jordan station (JOR) consistently shows the smallest errors, with an absolute MBE around $-37.57W/m^2$ and the lowest absolute RMSE of $78.64W/m^2$, demonstrating more reliable GHI estimations. APOCTT generally improves the performance in reducing absolute errors compared to the uncorrected CAMS model, particularly noticeable in the Izana station (IZA), where APOCTT reduces the absolute RMSE from $155.18W/m^2$ to $102.98W/m^2$. However, some stations like Budapest (BUD) and Cener (CNR) experience increased errors with Approach 1, suggesting that the method's effectiveness can vary significantly depending on the station. Detailed analysis of absolute error metrics highlights the importance of station-specific evaluations to optimize model performance, revealing areas for potential improvement and ensuring more accurate and dependable predictions across different data points.

D. Seasonal Absolute Error Metrics

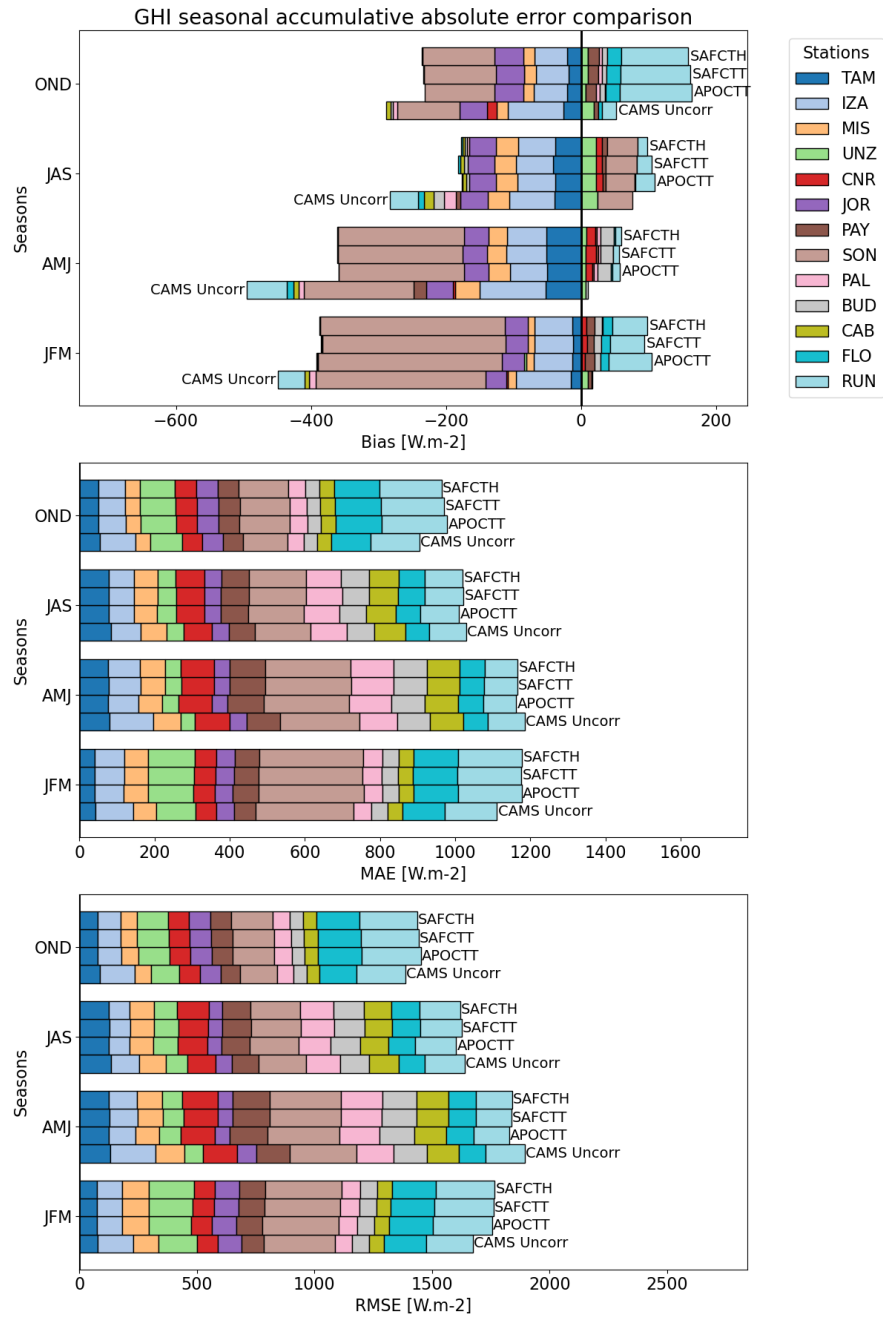


Figure D.1.: Seasonal Absolute error metrics