https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

MuscaDet - A multiscale and uncertainty-aware detector of cloud-free periods in time series of 1 min measurements of solar irradiance at ground level

Yves-Marie Saint-Drenan, Lucien Wald, Cyril Voyant, Mireille Lefèvre, Yehia Eissa

Mines Paris, Université PSL, Centre Observation Impacts Energie (O.I.E.), 06904 Sophia Antipolis, France

Correspondence to: Yves-Marie Saint-Drenan (yves-marie.saint-drenan@minesparis.psl.eu)

Abstract. Detecting cloud-free periods in time series of measurements of solar irradiance at ground level is essential in many domains, from meteorology to management of photovoltaic systems, including verification of so-called clear-sky models such as the McClear service of the Copernicus Atmosphere Monitoring Service. However, current cloud-free detectors call themselves on clear-sky models to discard measurements that are not close enough to the model outputs, introducing potential biases that pose a problem for the quarterly verification of the McClear service carried out in our institute on behalf of Copernicus. Accordingly, a new multiscale detector called MuscaDet was devised that is based on the correlations between measurements and a locally fitted Lambert-Beer-type function computed over a cascade of temporal windows of different lengths. This approach ensures that the detection remains independent of any predefined clear-sky model while still leveraging the dependence of the irradiance on the solar zenithal angle in cloud-free conditions. Measurement uncertainty and natural atmospheric variability are accommodated through a dynamic tolerance depending on the solar zenithal angle. MuscaDet is designed to operate on 1 min measurements of irradiance for the specific requirements of evaluation of clear-sky models. It needs measurements of the irradiance and its direct or diffuse component to discard overcast conditions. For its verification, time series spanning several years were collected at seven stations of the Baseline Surface Radiation Network covering various climates. Visual inspections of the measurements and outputs of MuscaDet show that the detector is accurate even in cases of short-lived isolated clouds or ramping cloudy events. Most cloud-free instants are detected and no visually-detected clouds are falsely noted as cloud-free by MuscaDet. A comparison with the currently used cloud-free detector indicates that MuscaDet provides a more balanced statistical distribution of solar zenithal angles, thus improving the verification of the McClear service. MuscaDet is easy to implement and needs no external input except the solar zenithal angle which is easily computed using existing libraries.

1 Introduction

This paper deals with the solar irradiance received by a horizontal surface at ground level and spectrally integrated over the entire solar wavelength range. It is noted E hereafter and is typically estimated by satellite-based observations and numerical weather models or measured at ground-based radiometric stations. In the later cases, pyranometers are widely used to measure E, and, when equipped with a shadow ball or ring, its diffuse component D. In addition, the direct component B_N received

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

from a small solid angle centered to the position of the Sun and received by a surface perpendicular to the solar beam (Blanc et al., 2014), is typically measured with pyrheliometers.

A cloud-free, or cloudless, instant refers to an instant when no cloud is present over the portion of the sky vault observable by an instrument or an observer. An equivalent term often used in the literature is “clear-sky”, that may be ambiguous because it may be confused with the clean atmosphere which refers to an atmosphere with no aerosols. A sequence of consecutive cloudless instants defines a cloud-free period. The detection of cloud-free periods is essential in studies of solar radiation (Bright et al., 2020; Gueymard et al., 2019; Lefèvre et al., 2013) and is also relevant to many other applications, such as quality control of pyranometric measurements and studies of atmospheric composition, photovoltaic systems, hydrology, and agriculture.

Cloud-free detectors are algorithms developed to detect cloud-free periods from time-series of irradiance measurements. This task is challenging because irradiance variability arises not only from atmospheric factors, such as clouds, aerosols, and water vapor, but also from instrument-related effects, including noise and thermal offsets, that are not necessarily cloud-induced (Vuilleumier et al., 2014). Numerous detectors have been proposed in the literature, each optimized for the specific application it targets.

Among applications requiring cloud-free detectors, a key motivation for the present work is the verification of clear-sky models. Such models estimate E, B_N and D in cloud-free conditions. They require inputs such as solar geometry, atmospheric composition (aerosols, water vapor, ozone), and surface altitude. Prominent examples include SOLIS (Mueller et al., 2004), REST2 (Gueymard, 2008) and McClear (Gschwind et al., 2019; Lefèvre et al., 2013). Since verification consists in comparing qualified ground-based measurements against model outputs, the accuracy of the verification of any clear-sky model depends critically on the detection of cloud-free conditions in the time series of measurements and the choice of the cloud-free detector is itself a source of uncertainty. This was particularly illustrated by Wandji Nyamsi et al. (2023) who reported that evaluations of McClear conducted by different works at the same stations but using different cloud-free detectors can lead to noticeably different performance metrics.

The need for accurate cloud-free detection has led to the development of numerous approaches, most of which rely on the output of a clear-sky model (Bright et al., 2020; Gueymard et al., 2019 and references herein). However, such model-dependent detectors pose a problem when the objective is to verify a clear-sky model. Bright et al. (2020) further demonstrated that many cloud-free detectors exhibit characteristic errors at multiple temporal scales, likely arising from measurement noise, variability in atmospheric composition, cloud-related fluctuations at different scales, and normalization issues.

A new multiscale cloud-free detector – called MuscaDet – is proposed to address these limitations. Unlike many approaches that rely on the output of a clear-sky model, MuscaDet is inspired by the work of Long and Ackerman (2000) and utilizes a local fit of a simple parameterization describing the dependence of E on the solar zenithal angle θ_s under cloud-free conditions. Besides the use of this parameterization, the innovation lies in the use of a multiscale approach and the accounting for uncertainties arising from both sensor measurements and atmospheric variability.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

Designed specifically for the purpose of verifying clear-sky models, our method is not intended to detect every cloud-free instant but to ensure a high confidence in detected cloud-free periods. As a result, it may exclude some actually cloud-free measurements in order to maximize confidence in the selected cloud-free periods, although this rejection is expected to remain minimal.

This work focuses on detecting cloud-free conditions using primarily measurements of E, whereas measurements of B_N or D are needed to exclude overcast cases. The proposed method is therefore designed for stations measuring E and at least one of its components—either E and B_N, or E and D. It also assumes accurate time-stamping of the measurements and minimal levelling errors of the pyranometers.

Finally, a comparison is conducted between the outputs of the cloud-free detector of Lefèvre et al. (2013), currently used in the operational verification of the McClear service of the Copernicus Atmosphere Monitoring Service, and those of MuscaDet. This comparison aims at assessing the impact of the choice of one of these detectors and at evaluating the possible benefits of using MuscaDet in the operational verification.

The remainder of this paper is organized as follows. Section 2 reviews the state of the art in cloud-free detectors and highlights the need for further development. Section 3 presents the measurements used for development and verification of MuscaDet, along with the calculation of astronomical quantities. Section 4 described MuscaDet in detail. Results of the verification are analysed in Section 5. Section 6 deals with the comparison between the outputs of the cloud-free detector of Lefèvre et al. (2013) and those of MuscaDet. Finally, Section 7 summarizes key findings, discusses the limitations of MuscaDet and outlines potential directions for future work.

2 Related works

A large fraction of the cloud-free detectors proposed in the literature—including most recent ones—rely on the output of a clear-sky model as a reference for detecting cloud-free conditions (Alia-Martinez et al., 2016; Ellis et al., 2018; Inman et al., 2015; Polo et al., 2009; Quesada-Ruiz et al., 2015; Reno and Hansen, 2016; Xie and Liu, 2013; Zhang et al., 2018). Although these approaches have demonstrated good performance, they have a methodological limitation in the context of the McClear service within the Copernicus Atmosphere Monitoring Service (CAMS) and for which our institute is responsible for its quarterly verification. Using a detector calling itself on McClear or any other clear-sky model to select the ground-based measurements that serve as reference to validate McClear introduces a circular dependency that may bias the results. Ensuring methodological independence therefore requires a detector that does not rely on clear-sky model outputs.

Only a few of the 22 cloud-free detectors reviewed by Bright et al. (2020) are independent of the outputs of a clear-sky model. Specifically, the approaches of Batlles et al. (2000), Ineichen et al. (2009), Lefèvre et al. (2013), Long and Ackerman (2000), and Perez et al. (1990b) satisfy this criterion, highlighting the limited availability of model-independent detectors suitable for verifying clear-sky models.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

Perez et al. (1990b) proposed a classification of sky conditions based on three variables: solar zenithal angle θ_s, sky clearness index ε, and sky brightness index. Cloud-free conditions are detected when ε, a function of B_N, D, and θ_s, exceeds 6.2. Building on a similar rationale, Ineichen et al. (2009) used a modified clearness index K_T' (Perez et al., 1990a), which accounts for atmospheric optical effects via air mass, and defined cloud-free conditions when $K_T' > 0.65$. Batlles et al. (2000) extended the approach based on thresholding of the clearness index K_T by introducing a condition on the diffuse fraction K - defined as the ratio of D to E. Empirical thresholds were applied to both indices, with values dependent on θ_s.

Long and Ackerman (2000) proposed a more data-driven approach: E is normalized using a cosine-based power-law fitted via robust regression, followed by multiple tests on the fitted parameters, the diffuse component, and short-term variability in irradiance. Lefèvre et al. (2013) combined an upper limit on K with a test on the temporal stability of K_T' based on a thresholding of the standard deviation of the measurements within a moving-time-window, and strict criteria on data availability to ensure high-confidence detection of cloud-free conditions.

A key distinction among the methods described above is how they account for the effect of solar geometry on irradiance measurements. To illustrate this, daily profiles of measurements of E and D, recorded at the Carpentras station on 15 July 2014, 30 April 2014, and 3 April 2014, are shown for typical cloud-free, broken-cloud, and overcast days in the upper row of Fig. 1 along with the horizontal irradiance at the top of the atmosphere (E_0). The lower row presents normalized quantities commonly used in cloud-free detection: the clearness index K_T, the modified clearness index K_T', and the Perez sky clearness index ε, with ε divided by 10 for the sake of the comparison.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

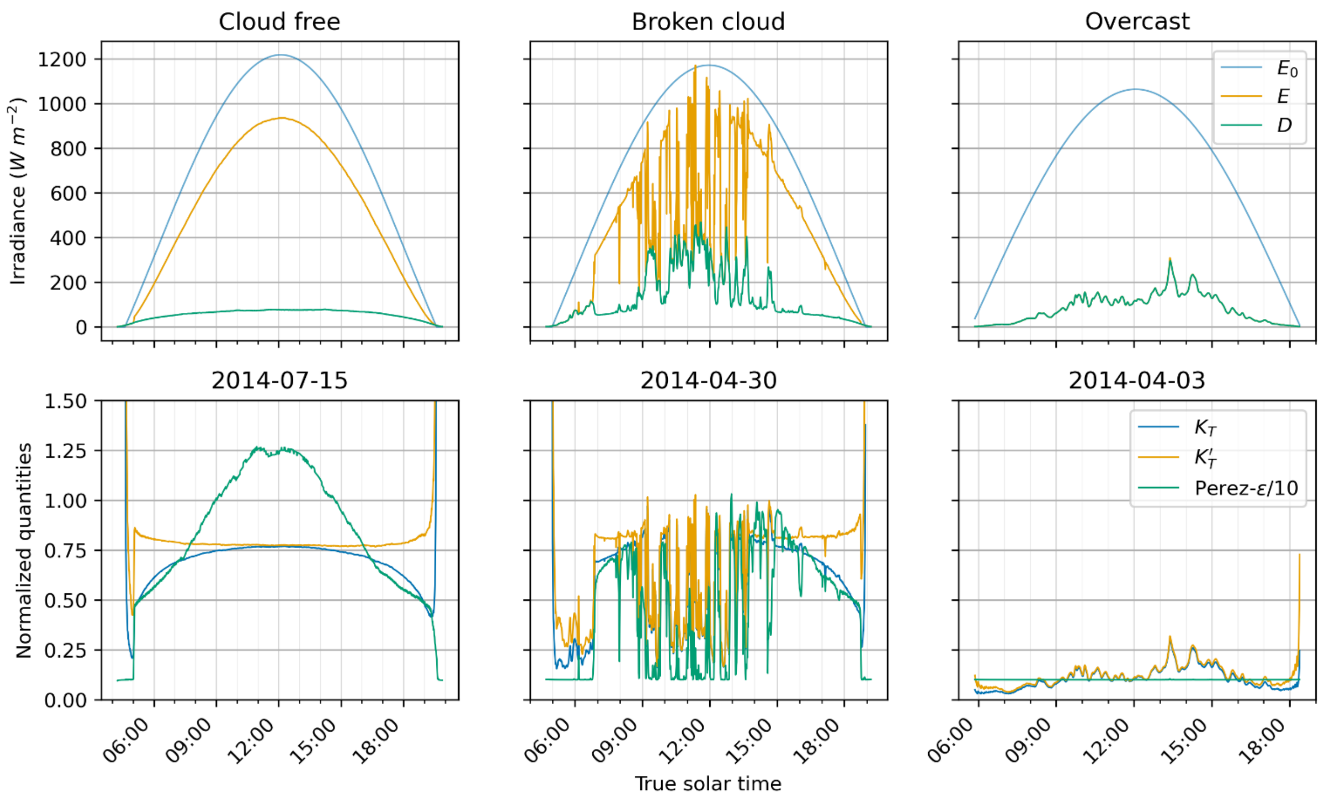

Figure 1: Daily profiles of E and D recorded at Carpentras, along with E_0, for typical cloud-free, broken-cloud and overcast days (upper plots from left to right) and corresponding daily profile of normalized quantities used in several cloud-free detectors (lower plots)

In the three examples shown in Fig. 1, the normalized variables K_T, K_T', and ε exhibit high values during the cloud-free day (left plot) and low values during the overcast day (right plot). Their intra-daily variations reflect cloud occurrence during the broken-cloud day (middle plot), confirming their relevance for detecting cloud-free periods. During the cloud-free day, each index exhibits a residual daily cycle linked to its sensitivity to air mass, whereas an ideal cloud-free indicator should remain constant throughout the day. Among the three, this sensitivity is strongest for ε, intermediate for K_T, and weakest for K_T'. These examples clearly illustrate how the use of K_T' reduces the dependence on θ_s compared to K_T or ε, and highlight the importance of such normalisation in cloud-free detection. However, as shown in the first example in Fig. 1, even the modified clearness index does not fully remove this dependence, since the relationship between θ_s and E remains sensitive to atmospheric conditions—particularly aerosol load and larger air mass. This limitation partly explains the growing use of clear-sky models in more recent approaches, as they provide a more accurate normalisation by explicitly accounting for radiative transfer under specified atmospheric conditions.

The detector proposed by Long and Ackerman (2000) offers an elegant way to address this issue by fitting a power-law function to E as a function of θ_s. However, the use of robust regression adds practical complexity, particularly for operational

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

applications. Moreover, as noted by Lefèvre et al. (2013), our objective — to detect cloud-free periods with high confidence, even at the expense of occasional false negatives (i.e. predicted cloudy while cloud-free, or equivalently missed cloud-free instants) — differs from that of Long and Ackerman (2000), who focused solely on detecting cloud-free situations without this requirement.

Beyond the works introducing a cloud-free detector, reviews of existing detectors have been provided by Gueymard et al. (2019) and Bright et al. (2020), who highlighted several recurring issues. In particular, they identified systematic patterns of false positives (i.e. predicted cloud-free while cloudy) and false negatives occurring across different temporal scales. Their analyses suggest that certain cloud situations—such as short-lived isolated clouds—are best detected using short time windows, whereas others—such as ramping events or persistent overcast conditions—are better captured over longer periods. These findings motivate our development of a multiscale detection strategy capable of accounting for the diversity of cloud dynamics. Finally, Bright et al. (2020) reported that false negatives are frequent in arid climates, where high surface albedo and rapidly changing aerosol conditions (e.g. dust events) may interfere with cloud-free detection. We hypothesize that additional sources of variability — such as instrumental noise, thermal offsets — may also contribute to the observed fluctuations in irradiance and be incorrectly interpreted as cloud-induced variability. This motivates the inclusion of an uncertainty term in our detection framework to reduce the risk of misclassification due to non-cloud-related variability in E.

Our algorithm Muscadet was initially developed in 2024 for the verification of the hourly means of cloud-free irradiance provided by the reanalyses ERA5, JRA-3Q, and MERRA-2 using 28 stations belonging to the Baseline Solar Radiation Network (Saint-Drenan and Wald, 2025). In this work, MuscaDet and its verification were presented so briefly that it did not allow for proper implementation by third parties. The present article offers a more precise description of MuscaDet and its verification and aims to facilitate its use by others.

3 Data

3.1 Selected sites and periods

The Baseline Solar Radiation Network (BSRN) offers high-quality measurements of E, B_N and D made every 1 min (Ohmura et al., 1998), noted respectively E_{meas}, $B_{N,meas}$ and D_{meas}. For this study, we use the same five BSRN stations as Bright et al. (2020) representing different climatological conditions: a tropical site with high rainfall (Brasilia), a mid-latitude site with frequent cloud cover (Lerwick), a high-latitude station with an extreme yearly cycle in solar zenithal angle (South Pole), a desert site affected by dust events (Tamanrasset), and the Sioux Falls station, extensively used in previous cloud-free detection studies. In addition to these, we include the Carpentras, Payerne and Palaiseau BSRN stations, for which extensive knowledge has been acquired during our routine activities in the verification of the McClear service. They are located in mid-latitude temperate regions and are characterized by moderate cloud cover. No additional plausibility check has been made on the measurements.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

At all stations, 1 min measurements from 2014 to 2016 were collected except Brasilia, for which the period is 2009–2011. A summary of the stations and measurement periods is provided in Table 1.

Table 1 : BSRN stations: names, countries, geographic location, periods of measurements, and Köppen-Geiger climates (Beck et al., 2018)

Site name (Country)	ID	Latitude	Longitude	Elevation asl (m)	Period	Köppen–Geiger climate
Lerwick (UK)	LER	60.1389°	-1.1847°	80	2014-01 / 2016-12	Cfd
Palaiseau (France)	PAL	48.7130°	2.2080°	156	2014-01 / 2016-12	Cfb
Payerne (Switzerland)	PAY	46.8123°	6.9422°	491	2014-01 / 2016-12	Cfb
Carpentras (France)	CAR	44.0830°	5.0590°	100	2014-01 / 2016-12	Csa
Sioux Falls (USA)	SXF	43.7300°	-96.6200°	473	2014-01 / 2016-12	Dfa/Dfb
Tamanrasset (Algeria)	TAM	22.7903°	5.5292°	1385	2014-01 / 2016-12	BWh
Brasilia (Brazil)	BRB	-15.6010°	-47.7130°	1023	2009-01 / 2011-12	Aw
South Pole (Antartic)	SPO	-89.9830°	-24.7990°	2800	2014-01 / 2016-12	EF

3.2 Astronomical quantities and others

Several astronomical quantities were calculated with the SG2 algorithm (Blanc and Wald, 2012) using the time in UT and the geographic coordinates at each station. These include the solar zenithal angle, θ_s, and the irradiance at the top of the atmosphere, E_0. The computation of the latter adopts the nominal total solar irradiance (TSI) of 1361 W m^{-2} recommended by Mamajek et al. (2015) where the TSI is the yearly mean of the irradiance at the top of the atmosphere at normal incidence over a solar cycle. For illustration purposes, the verification protocol described in Section 5.1 uses estimates of irradiance provided by the McClear service that generates 1 min time series of irradiance and its diffuse component under cloud-free conditions, noted respectively $E_{McClear}$ and $D_{McClear}$, consistent with the measurements.

4 Method

MuscaDet consists of four main steps: (i) fitting the modified Lambert–Beer (MLB) function to E_{meas}, yielding $f_{MLB}(\theta_s)$; (ii) dynamically defining a criterion to quantify the agreement between $f_{MLB}(\theta_s)$ and E_{meas}, which constitutes the core of the detector; (iii) selecting a tolerance that accounts for the expected variability of E_{meas} due to atmospheric fluctuations and measurement noise; and (iv) implementing a multiscale framework within the algorithm. These steps are described in Sections 4.1–4.4, while Section 4.5 provides an overview of how they are integrated into the complete method.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

4.1 Fit of the modified Lambert-Beer function to E_{meas}

The first step of MuscaDet consists of fitting a function relating E_{meas} and θ_s in cloud-free conditions. The approach is similar to that of Long and Ackerman (2000), but the power-law formulation is replaced by a MLB function (Mueller et al., 2004), defined as:

$$f_{MLB}(\theta_s) = E_0 \, exp[-b \, cos(\theta_s)^{-a}] \tag{2}$$

Here, a and b are estimated by applying a linear regression onto E_{meas}, using a linearized form of Eq. (2) obtained via a double-logarithmic (log–log) transformation.

$$log\left(log\left(\frac{E_0}{E_{meas}}\right)\right) = log(b) - a \, log(cos(\theta_s)) \tag{3}$$

The MLB function is fitted over moving time windows $[t - \Delta T, t + \Delta T]$ of varying lengths $2\,\Delta T + 1$, centered over the time stamp of the measurement under concern and discretized at a 1 min increment. To ensure the validity of the above expression, only E_{meas} strictly positive and less than E_0 are used to estimate the MLB parameters. To avoid fits during fully overcast periods, the parameters a and b are estimated only if at least 30 % of the measurements in both the preceding and following halves of the window satisfy the inequality $0 < E_{meas} < E_0$ and for which the corresponding $B_{N,meas}$ is strictly greater than 10 W m^{-2} and K less than 0.9. These criteria are designed to filter out persistently overcast conditions, although it is possible that several E_{meas} affected by transient or partial cloud cover may be included in the fitting procedure. The upper limit for K adopted here is greater than that in Lefèvre et al. (2013), who set it to 0.3. This choice is motivated by the observed occurrence of high K at large θ_s or under cloud-free conditions in highly turbid atmospheres (Li et al., 2025). Measurements not satisfying one of these conditions are flagged as “possibly cloudy”. Measurements fulfilling all these conditions are denoted as “suitable” and are used to fit the MLB function.

Once persistently overcast periods are excluded, a poor agreement between E_{meas} and $f_{MLB}(\theta_s)$ is expected within the moving window in the presence of clouds. Conversely, in cloud-free conditions, the MLB function should accurately describe the profile of E_{meas}; this is the key point of MuscaDet.

4.2 Criterion for cloud-free decision accounting for measurement uncertainty and atmospheric variability

Several statistical metrics have been tested to evaluate the agreement between E_{meas} and $f_{MLB}(\theta_s)$, namely the root-mean-square difference (RMSD), the mean absolute difference (MAD), their normalized counterparts (nRMSD and nMAD), as well as the Pearson correlation coefficient ρ and the coefficient of determination R^2. The Pearson correlation coefficient was adopted here to compare the temporal pattern of E_{meas} and $f_{MLB}(\theta_s)$ as its bounded and dimensionless nature facilitates the selection of an upper threshold.

Applying a strict threshold to the correlation coefficient may, however, lead to false negatives when irradiance exhibits limited temporal variability, such as near solar noon or at high latitudes. In such cases, the variance of the measurements may be

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

dominated by noise or small-scale atmospheric fluctuations, resulting in artificially low correlation coefficients despite actual cloud-free conditions. To mitigate this effect, a tolerance parameter σ_{tol} is introduced, which accounts for both measurement noise and natural atmospheric variability. As detailed in Appendix A, the expression for the Pearson correlation coefficient is reformulated to explicitly include σ_{tol}, leading to the following decision criterion. For a given interval $[t-\Delta T, t+\Delta T]$, the instant t is declared cloud-free if:

$$\frac{Cov(E_{Meas}, f_{MLB}(thetas))}{\sqrt{max(0, Var(E_{meas}) - \sigma_{tol}^2)\, Var(f_{MLB}(\theta_s))}} > \rho_{min} \tag{4}$$

where ρ_{min} is an adjustable parameter empirically set to 0.999 after sensitivity testing.

It can be observed that Eq. (4) reduces to the standard Pearson correlation coefficient when $Var(E_{meas}) \gg \sigma_{tol}^2$. Conversely, when $Var(E_{meas}) < \sigma_{tol}^2$, the denominator becomes null, resulting in an undefined (formally infinite) value and thus a systematically satisfied test. This highlights the role of σ_{tol}^2as a tolerance threshold.

To apply Eq. (4), the variances of E_{meas} and $f_{MLB}(\theta_s)$, as well as their covariance, are calculated using all measurements within the time window, including those not suitable for the fitting procedure.

4.3 Formulation of the tolerance

The tolerance term σ_{tol} accounts for the combined effect of instrumental noise and intrinsic atmospheric variability on the difference between E_{meas} and E_{MLB} under cloud-free conditions. The latter arises from temporal fluctuations in atmospheric properties within the considered time window, in particular aerosol optical depth (AOD) and total column water vapour. The magnitude of the resulting variability in E_{meas} depends on θ_s. Therefore, σ_{tol} is parameterized as a linear function of $cos(\theta_s)$:

$$\sigma_{tol} = c_1 + c_2\, cos(\theta_s) \tag{5}$$

The two parameters have been empirically set to c_1=1 W m^{-2} and c_2=4 W m^{-2} based on exploratory tests conducted using measurements from the Palaiseau station. The validity of these parameters at other stations was assessed through a systematic qualitative evaluation based on visual inspection of model outputs (see Sect. 5).

The tolerance is intended to address a limitation of the decision criterion used in MuscaDet, which relies on the dependence of E on θ_s, a dependence that weakens around solar noon, i.e. at low θ_s values. The proposed formulation relaxes the test for low solar zenithal angles, when the effect of atmospheric variability on E becomes more pronounced than that of θ_s, while remaining conservative under other conditions.

Although this expression has shown satisfactory results in practice, it remains empirical. Future improvements could involve incorporating a formal model of the uncertainty of pyranometer measurements, as well as an explicit estimation of variations in atmospheric composition. Such refinements would provide a more physically grounded definition of σ_{tol}, better aligned with the conceptual framework developed throughout this work.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

4.4 Implementing a multiscale framework

To account for the wide range of temporal dynamics associated with cloud cover, MuscaDet is implemented within a multiscale framework. This approach enables the detector to remain sensitive to both short-lived transients and more persistent cloud structures. Specifically, the method is applied to moving windows of varying lengths, corresponding to different temporal scales.

The shortest timescale must be short enough to detect rapid cloud transitions, such as isolated cumulus clouds, yet long enough to ensure that the correlation remains statistically meaningful. A minimum window length of 15 min (ΔT= 7 min) was therefore selected as a compromise between sensitivity and sufficient sample size for reliable correlation estimation. Conversely, the longest temporal scale must be sufficiently long to capture slowly evolving cloud patterns. Extending the window too far, however, could increase the likelihood of significant variations in atmospheric composition (e.g., aerosols, water vapor) during the period, thereby reducing the validity of the fitted MLB function. A maximum window length of 121 min (ΔT= 60 min) was thus chosen as a suitable upper bound.

To cover intermediate scales, the selected temporal resolutions are inspired by the dyadic scaling commonly used in time–frequency analysis. Accordingly, MuscaDet evaluates four window lengths of 15, 31, 61, and 121 min, corresponding to temporal scales approximately following a 2^n progression.

The principle of multiscale detection is that an instant t is classified as cloud-free only if it satisfies the detection criterion at all four scales simultaneously. This multiscale condition ensures a conservative classification, minimizing the risk of false positives while increasing confidence in the detected cloud-free periods.

4.5 Summary of MuscaDet

MuscaDet consists of a series of cascading quasi-dyadic multiscale filters applied to 1 min measurements for each day in the time series.

The MLB function is fitted over moving time windows $[t - \Delta T, t + \Delta T]$ of varying lengths 2 ΔT+1, centered over the time stamp of the measurement under concern. Several constraints are applied prior to the fit. Only E_{meas} in this interval that are strictly positive and less than E_0 are retained to ensure the validity of Eq. 3. Further constraints are added to avoid fitting during fully overcast periods: only measurements for which the corresponding $B_{N,meas}$ is strictly greater than 10 W m^{-2} and K less than 0.9 are selected. An instant t, and accordingly the corresponding E_{meas}, is flagged as "*suitable*" if it satisfies all four criteria above. The fit of the MLB function is realized if an instant is "suitable" and if at least 30 % of the measurements in both the preceding and following halves of the window are also "suitable". Otherwise, the instant is considered "*possibly cloudy*," and the algorithm proceeds to the next instant.

The agreement between E_{meas} and $f_{MLB}(\theta_s)$ over the interval is then assessed. To this end, the variances $Var(E_{meas})$ and $Var(f_{MLB}(\theta_s))$, as well as the covariance $Cov(E_{meas}, f_{MLB}(\theta_s))$, are computed using all measurements in the same interval, including measurements that are not suitable for the fitting procedure. In parallel, σ_{tol} is determined based on θ_s at time t,

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

using Eq. (5). The instant t is finally flagged as “cloud-free” for this ΔT if the condition defined in Eq. (4) is satisfied, otherwise the instant is considered "possibly cloudy".

The procedure produces flags "cloud-free" / "possibly cloudy" at each instant t for each of the four windows (ΔT = 7, 15, 30, and 60 min), resulting in four flags at each t. MuscaDet finally flags as "cloud-free" only the instants for which the four flags are "cloud-free". The different steps of MuscaDet are summarized in the flowchart shown in Fig. 2.

MuscaDet requires at least 30% suitable measurements within time windows of length ΔT before and after each instant. As a result, detections cannot occur in the immediate vicinity of sunrise and sunset. The corresponding temporal offset depends not only on ΔT but also on additional filtering criteria, notably thresholds on K and $B_{N,meas}$, which are not satisfied at high solar zenithal angles.

The Python implementation of MuscaDet is openly available on GitHub (see Code availability section), ensuring transparency and reproducibility while providing an efficient implementation of its core computational steps.

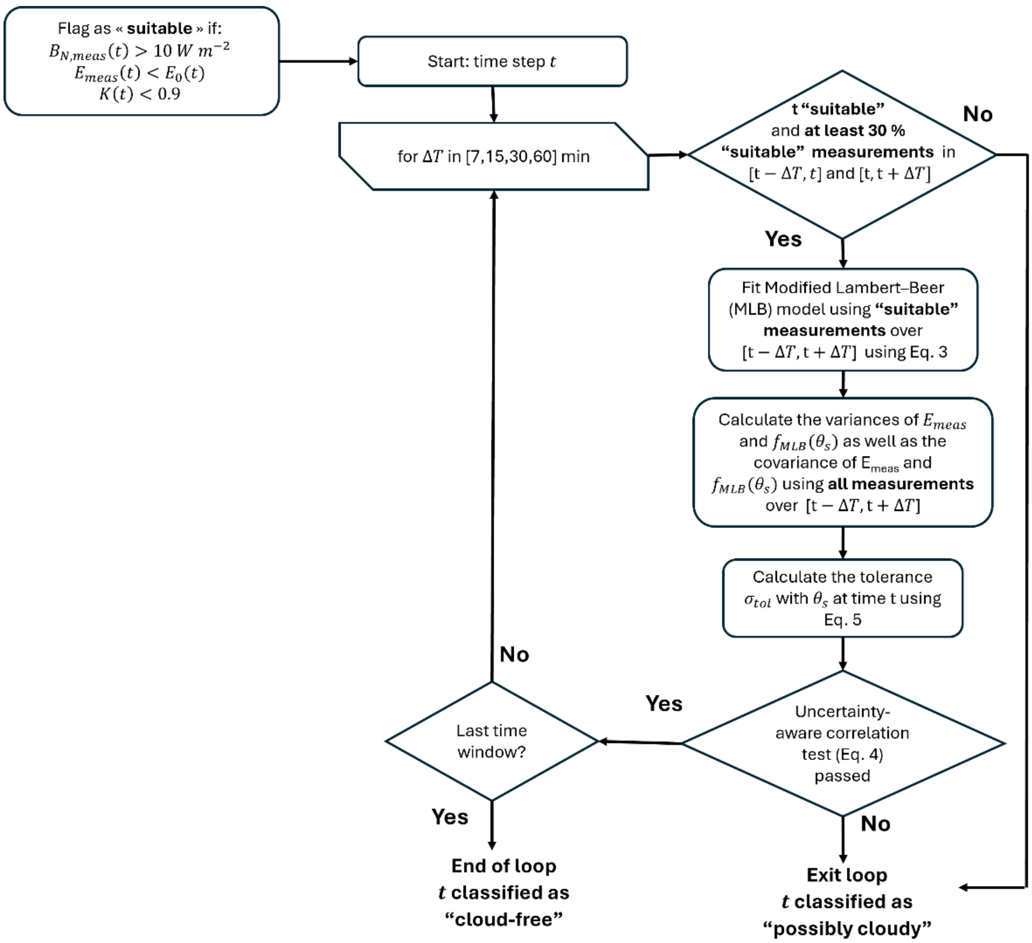

Figure 2 : Flowchart of MuscaDet. Rounded rectangles represent processing steps; diamonds indicate decision points. The method is applied across multiple timescales. An instant is classified as cloud-free only if detection criteria are met at all scales.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

5 Verification of the Performance of MuscaDet

5.1 Verification protocol

Verification of cloud-free detectors may be performed using independent references such as collocated sky imagers (Bright et al., 2020; Gueymard et al., 2019; Long and Ackerman, 2000). In the absence of such instruments at the studied BSRN sites, and given the current inadequacy of sky camera data for the automatic separation of cloudy and cloud-free periods (Gueymard et al., 2019), the verification of Muscadet was performed through visual inspection of the measurements by the authors. As Gueymard et al. (2019) concluded that there is currently no known suitable verification database for detectors of cloud-free instants, quantitative methods to evaluate the performance of any detector lack appropriate statistical significance. In any case, and for further reference, we follow the evaluation approach of Bright et al. (2020) and provide all our results as supplementary material.

MuscaDet was subjected to several verification steps during its development (Saint-Drenan and Wald, 2025). In a first step, MuscaDet was applied to time-series provided by the McClear service. All McClear outputs were correctly identified as cloud-free by MuscaDet, thus confirming the absence of false negatives (missed cloud-free instants) in well-defined cloud-free conditions—a limitation sometimes reported for other approaches, such as Lefèvre et al. (2013) (see Sect. 6). This assessment showed that McClear outputs are reliably flagged as cloud-free, except at higher solar zenithal angles (see Sect. 5.2), thereby providing an initial assessment of the performance of MuscaDet under ideal conditions. In a second step, cases of cloud-free skies including the slow appearance of thin clouds were simulated; MuscaDet was applied to the outcomes and correctly labelled the cloudy instants. Finally, MuscaDet was applied to randomly selected days in the BSRN data set. Its outcomes as well as those of the detectors of Lefèvre et al. (2013) and Reno and Hansen (2016) were visually inspected for a further check of the proper functioning of MuscaDet.

This paper presents a more extensive verification of the outcomes of MuscaDet which is applied to the whole data sets of the 8 stations. Graphs were drawn for each day exhibiting the daily profiles of E_{meas} on which were superimposed MuscaDet outcomes. Each graph was visually inspected. As already mentioned, MuscaDet is not intended to detect every cloud-free instant but to ensure a high confidence in detected cloud-free periods. Consequently, false negative errors (missed cloud-free instants) are considered less critical than false positive errors (cloud-free detections while cloud contaminated). Avoiding false positives is therefore the primary objective, while keeping false negatives as low as reasonably achievable so that valuable cloud-free measurements are not unnecessarily discarded. This was kept in mind by the authors when visually inspecting the graphs.

The visual inspection shows generally consistent behaviour of MuscaDet across stations. Key observations from this inspection are summarized in the following sections for five categories of representative cloudy and atmospheric conditions: (i) cloud-free days, (ii) low-amplitude fluctuations in irradiance in a globally cloud-free day, (iii) isolated dips in irradiance induced by passing clouds, (iv) transitions between cloud-free and cloudy periods, and (v) cloudy days. A dedicated analysis is also provided for the South Pole station (SPO), a high-latitude site where the specific daily evolution of θ_s poses a particular

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

challenge for cloud-free detectors. The use of this station as a stress test follows the proposal of Bright et al. (2020) for assessing detector robustness.

For each category, representative examples from different days or stations are presented and discussed. This verification evaluates the extent to which the objective of MuscaDet is achieved and identifies current limitations. The insights gained from this analysis inform potential future improvements of Muscadet, discussed in the concluding section.

5.2 Cloud-free days

Fig. 3 presents the outputs of MuscaDet for two days identified as cloud-free during the visual inspection: 2014-05-04 at Carpentras (left panel) and 2009-07-30 at Brasilia (right panel). In each panel, E_{meas} is plotted as a function of time throughout the day. Dots indicate values of E_{meas} at time steps identified as cloud-free instants by MuscaDet.

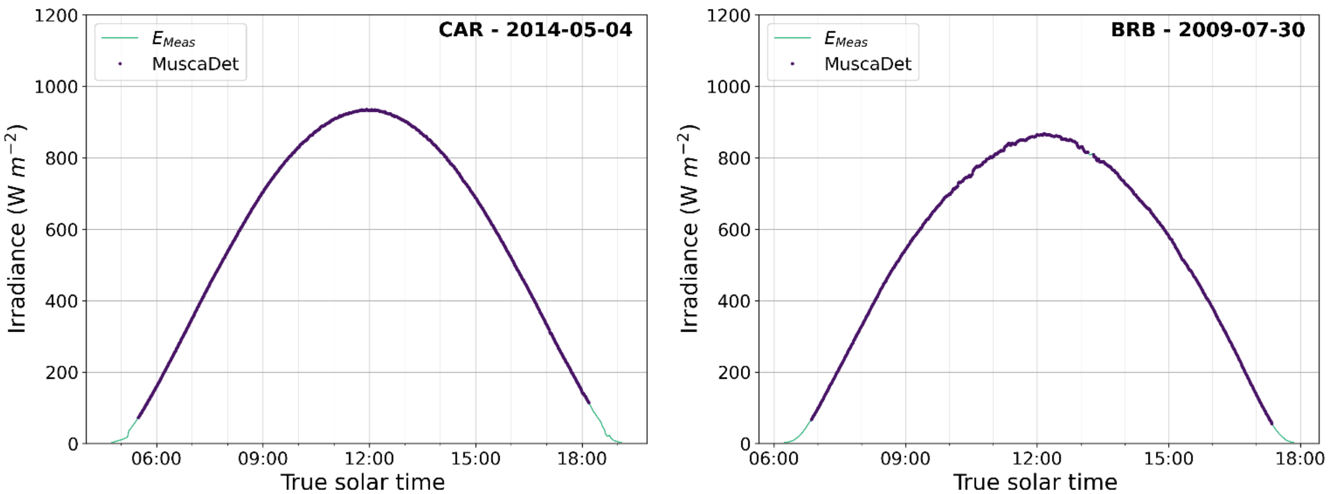

Figure 3: Measured irradiance (E_{meas}, green line) for two cloud-free days: 2014-05-04 at Carpentras (CAR, left) and 2009-07-30 at Brasilia (BRB, right). Blue dots indicate instants identified as cloud-free by MuscaDet.

In both examples, the daily profile of E_{meas} follows a smooth sine-like shape. Nearly all values throughout the day are classified as cloud-free, except for short periods in the early morning and late afternoon, as discussed in Section 4.5.

A common limitation of cloud-free detectors was reported by Bright et al. (2020) who showed that the performance of several detectors tends to degrade at high solar zenithal angles, often resulting in fewer detections near sunrise and sunset. As illustrated in Fig 3, visual inspection indicates that MuscaDet does not exhibit such behaviour, apart from the specific situations discussed above. This robustness likely results from the adopted strategy, which avoids normalizing E_{meas} to remove its dependence on θ_s (e.g. using K_T or its modified version K_T') and instead compares E_{meas} directly with a MLB function fitted to E_{meas}.

Although cloud-free conditions prevail on both selected days, one may note differences in the daily profile of E_{meas}. At Carpentras, the profile is very smooth while small fluctuations can be seen at Brasilia, notably between 11:00 and 13:00. It remains unclear whether these variations arise from changes in atmospheric variables such as water vapour, aerosol load, or thin cirrus clouds, or from defects in instrumentation. Given the small amplitude of the observed fluctuations, these instants

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

are classified by MuscaDet as cloud-free because of the adopted tolerance-aware detection criteria. The authors are unable to make a definitive judgment on the causes of these fluctuations.

5.3 Low-amplitude fluctuations in irradiance in a globally cloud-free day

While MuscaDet performs well under ideal cloud-free conditions, small fluctuations around the theoretical daily profile of E in cloud-free conditions are frequently observed. These variations may result from subtle atmospheric effects, such as thin cirrus clouds or changes in atmospheric content, but they could also arise from instrumental or datalogger issues. Although they do not necessarily indicate significant cloud cover, these fluctuations may lead to misclassification. Careful examination of such cases is therefore important to assess the sensitivity of the detection method and its robustness under realistic cloudy and atmospheric conditions.

Two representative examples of such cases are presented in Fig. 4: 2014-03-13 at Palaiseau (left panel) and 2014-02-26 at Tamanrasset (right panel). E_{meas} is plotted as a function of time, and cloud-free instants detected by MuscaDet are marked by dots. To facilitate visual inspection, measurements of the diffuse component D_{meas} are also shown, together with the McClear estimates $E_{McClear}$ and $D_{McClear}$.

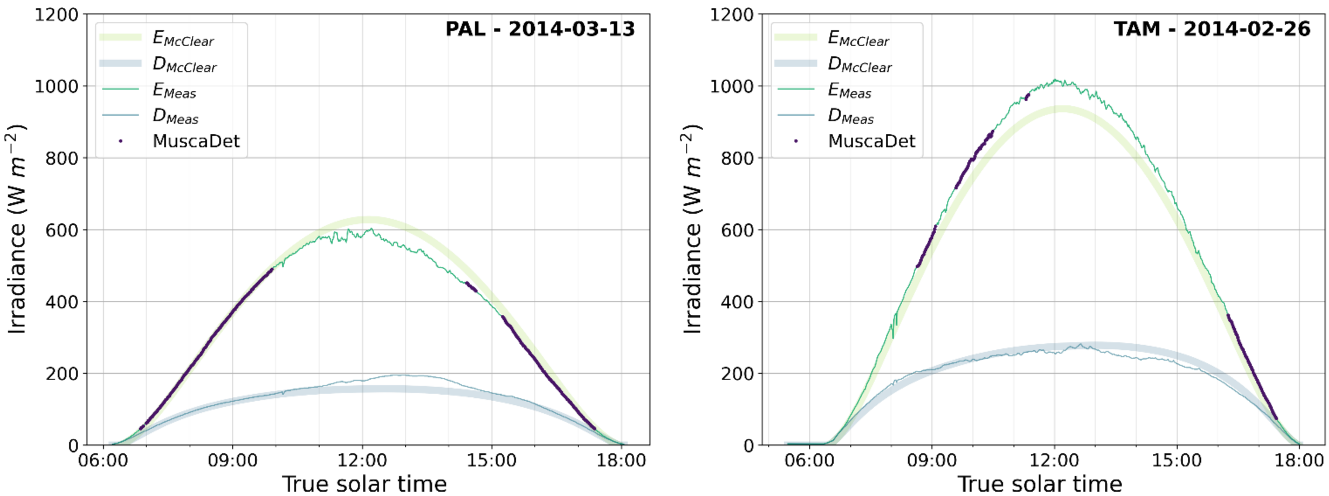

Figure 4: Measured irradiance (E_{meas}, green line) for two days with low-amplitude fluctuations in E_{meas} during globally cloud-free conditions: 2014-03-13 at Palaiseau (PAL, left) and 2014-02-26 at Tamanrasset (TAM, right). Light green and light blue lines denote $E_{McClear}$ and $D_{McClear}$, respectively. Blue dots indicate instants identified as cloud-free by MuscaDet.

Even though the overall daily profiles of E_{meas} and D_{meas} in the two examples from Fig. 4 may initially appear indicative of cloud-free conditions—suggesting a high number of false negatives—a closer examination shows that low-amplitude fluctuations in E_{meas} explain why several instants are classified as "potentially cloudy" by MuscaDet.

In the first example (PAL, 2014-03-13), the period from 11:00 to 15:00 is flagged as "likely cloudy" by MuscaDet. High-frequency variations of small amplitude are observed in E_{meas} which does not follow the daily profile of $E_{McClear}$. In addition, one notes an increase in D_{meas} with respect to $D_{McClear}$ indicating the likely presence of clouds. In the second example (TAM,

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

2014-02-26), high-frequency, low amplitude variations in E_{meas} are observed almost all day. E_{meas} is often greater than $E_{McClear}$ while D_{meas} is often less than $D_{McClear}$, especially from 10:00 to 16:00. This may denote changes in scattering and/or absorption, which may be due to clouds or variations in atmospheric constituents with a possible greater contribution of the diffuse component with respect to the McClear model. The presence of clouds cannot be ruled out. In line with the objective of MuscaDet to confidently identify cloud-free instants, these instants are not classified as cloud-free.

If future analyses confirm that such fluctuations occur under truly cloud-free conditions, potential users could consider relaxing the detection threshold, particularly for short timescales.

5.4 Isolated dips in irradiance induced by passing clouds

Visual inspection of the data sets reveals a frequent situation in which the daily profile of E_{meas} generally follows that of a typical cloud-free day, except for brief, short-lived dips, which are defined as rapid decreases in irradiance followed by a quick recovery within a few minutes. These dips are caused by the attenuation of B_N by small clouds. In such cases, not only the dip period must be classified as “possibly cloudy” but also the preceding and following intervals, since the passing small cloud responsible for the dip is present in the sky before and after the dip event, thus influencing the diffuse component D_{meas}.

Two examples illustrating such cases are shown in Fig. 5: measurements at Lerwick on 2014-03-23 (left panel) and at Carpentras on 2014-08-27 (right panel). As in Fig. 4, the daily profiles of E_{meas} and MuscaDet detection are drawn together with those of D_{meas}, $E_{McClear}$ and $D_{McClear}$, to facilitate visual inspection.

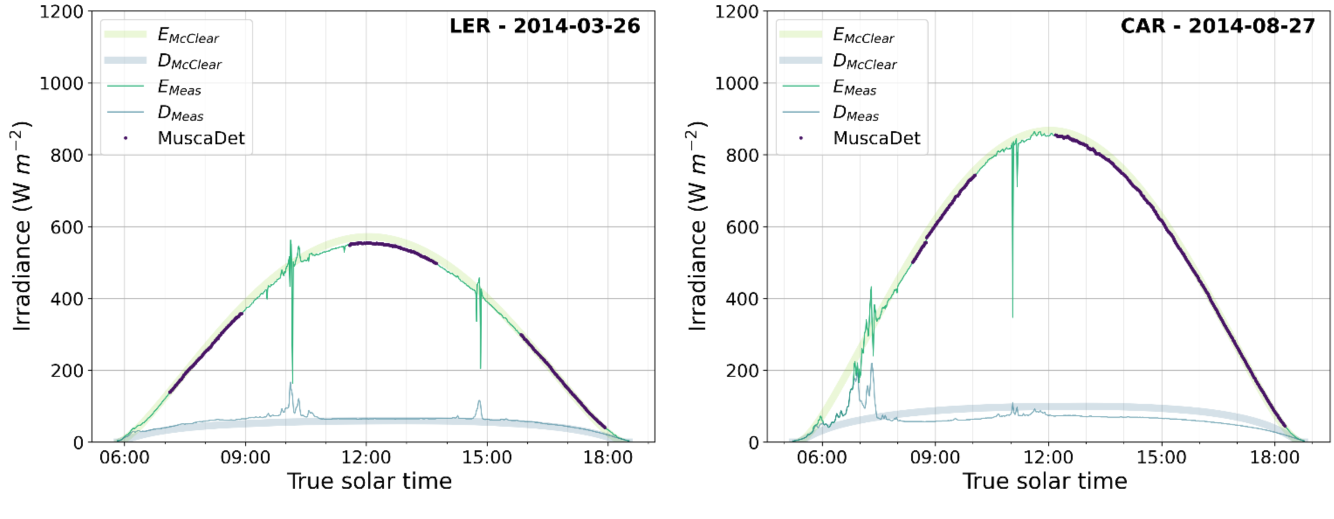

Figure 5: Measured irradiance (E_{meas}, green line) for two days with cloud-induced isolated dips in irradiance causing short departures from cloud-free conditions: 2014-03-26 at Lerwick (LER, left) and 2014-08-27 at Carpentras (CAR, right). Light green and light blue lines denote $E_{McClear}$ and $D_{McClear}$, respectively. Blue dots indicate instants identified as cloud-free by MuscaDet.

In the first example (LER, 2014-03-23), two distinct dips are observed around 10:00 and 15:00. These periods as well as a time buffer of roughly 60–80 min surrounding the dips are flagged as "possibly cloudy" by MuscaDet. The use of a maximum temporal scale of 121 min (=2ΔT + 1) ensures a minimal period of ΔT = 60 min between an instant affected by clouds and the

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

nearest cloud-free detection. A similar behaviour is observed in the second example (CAR, 2014-08-27), where a single dip is preceded and followed by a buffer of approximately 60 min during which instants are flagged as "possibly cloudy".

Irradiance dips are typically accompanied by increases in diffuse component relative to measurements obtained under truly cloud-free conditions, often starting before the dip and persisting afterwards (see e.g. D_{meas} at 11:00 and 13:00 in CAR, 2014-08-27). During these periods, E_{meas} remains close to clear-sky values, whereas D_{meas} is significantly enhanced compared to expected cloud-free levels. Such measurements should therefore be excluded when verifying diffuse component from clear-sky models.

The multiscale approach introduces a time buffer that effectively excludes these periods of higher diffuse component. This behaviour is particularly evident for the second dip in the first example (LER, 2014-03-23) and for the single dip in the second example (CAR, 2014-08-27). These results confirm that the algorithm behaves as intended: it not only excludes the dips themselves but also applies a temporal buffer around them, thereby preventing the misclassification of moments immediately before or after a dip as cloud-free.

5.5 Transitions between cloud-free and cloudy periods

Another behaviour observed in the data sets is the transition between cloud-free and cloudy conditions and reciprocally. As in the case of isolated dips in irradiance, a temporal buffer must be applied before and after the cloudy regime to exclude instants when clouds are present in the sky. Such transitions are illustrated in Fig. 6 through two examples at Carpentras: 2014-04-15, and 2014-08-14.

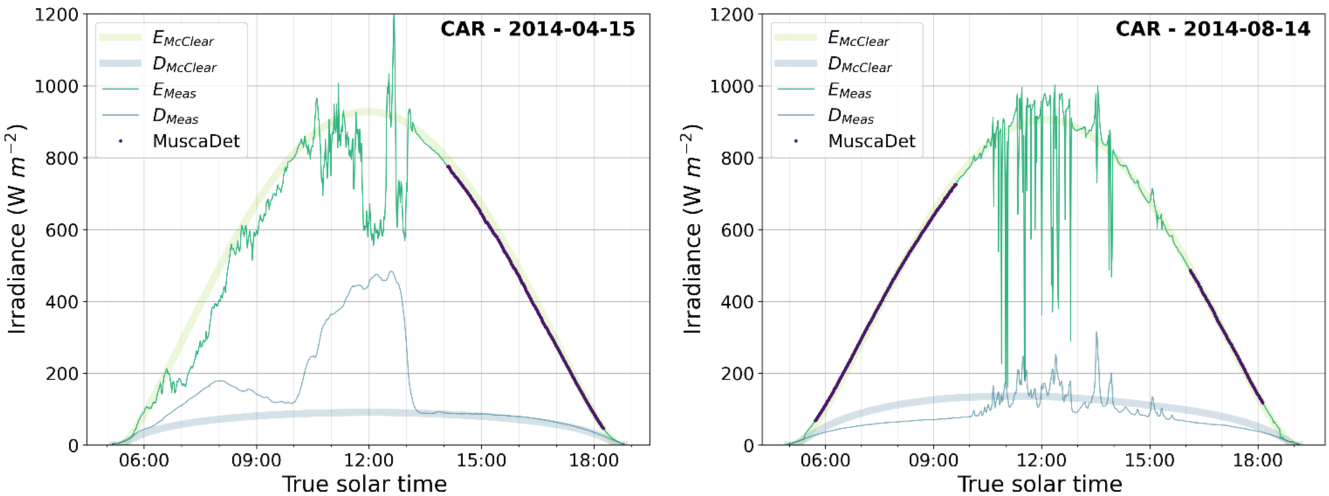

Figure 6: Measured irradiance (E_{meas}, green line) for two days with transitions between cloudy and cloud-free conditions at Carpentras (CAR): 2014-04-15 (left) and 2014-08-14 (right). Light green and light blue lines denote $E_{McClear}$ and $D_{McClear}$, respectively. Blue dots indicate instants identified as cloud-free by MuscaDet.

In the first example (CAR, 2014-04-15), fluctuations in E_{meas} are observed between 05:00 and 14:00 that may be related to varying cloud cover as also suggested by high values of D_{meas} with respect to $D_{McClear}$. These fluctuations disappear around

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

13:00, after which the irradiance profiles become consistent with cloud-free conditions. Expectedly, all instants after 14:05 are flagged as "possibly cloudy" in relation with the size of the 121 min time window. In the second example (CAR, 2014-08-14), variability in E_{meas} is observed between 10:40 and 14:00, likely due to scattered clouds. Cloud-free conditions are detected in the morning until 09:40, corresponding to a 60 min buffer before the onset of variability, and cloud-free conditions are again detected at 16:10—130 min after the end of the main cloudy period. This longer delay may result from residual variability, including a noticeable increase in E_{meas} around 15:10.

5.6 Cloudy days

Visual inspection confirms that MuscaDet behaves efficiently under persistently cloudy conditions whether overcast or highly variable. Two illustrative examples are shown in Fig. 7: Lerwick on 2014-10-23 and Brasília on 2009-08-03. In the first example (LER, 2014-10-23), the daily profile of E_{meas} is fairly flat, typical of overcast conditions. This is confirmed by the fact that $B_{N,meas}$ remains equal to zero throughout the day. As MuscaDet applies the detection test only to instants where $B_{N,meas}$ exceeds 10 W m^{-2}, no cloud-free instants are detected during this day. In the second example (BRB 2009-08-03), the daily profile of E_{meas} is fairly close to that of cloud-free conditions in the morning though highly variable and is very chaotic after 13:00, typical of a broken cloud cover. Though $B_{N,meas}$ is greater than 10 W m^{-2}, MuscaDet does not detect any instants as cloud-free due to pronounced variability in E_{meas}.

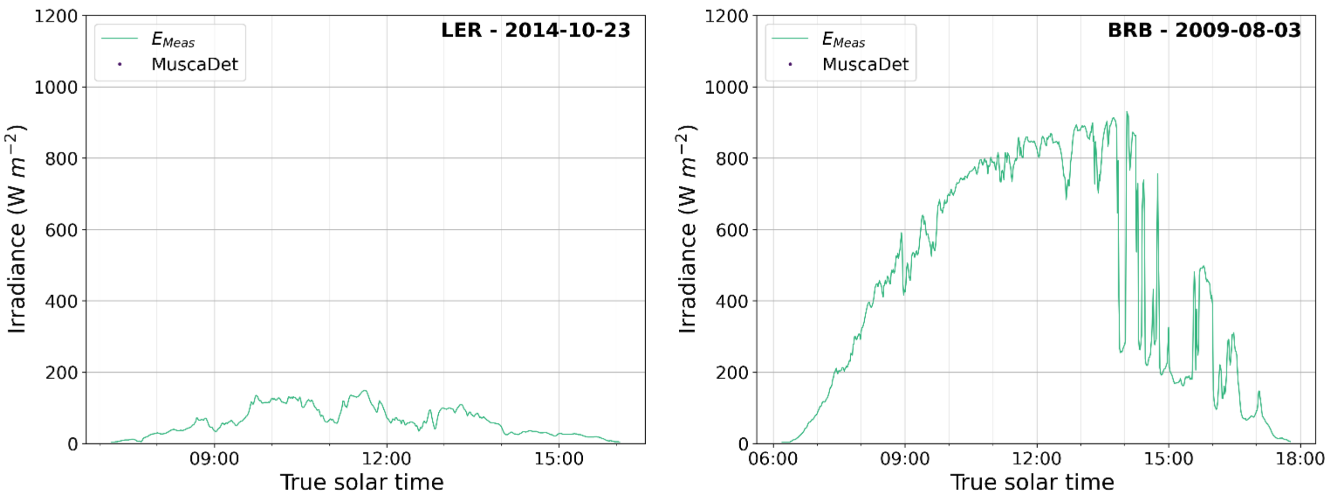

Figure 7: Measured irradiance (E_{meas}, green line) for two cloudy days: 2014-10-23 at Lerwick (LER, left) and 2009-08-03 at Brasilia (BRB, right). Blue dots indicate instants identified as cloud-free by MuscaDet.

5.7 Verification at the South Pole Observatory station

The visual inspection was finally conducted at the SPO station, where the behaviour of E_{meas} is specific due to the high latitude, providing a particularly challenging test for MuscaDet. Two representative days are shown in Fig. 8: a cloud-free day (2014-01-21, left panel) and a day characterised by a cloudy period followed by cloud-free conditions (2014-11-09, right

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

panel). In the first example (SPO 2014-01-21), the daily profile of E_{meas} remains nearly constant throughout the austral summer day. This behaviour is characteristic of high latitudes, where the dependence of E_{meas} on θ_s is weak. Despite this limited dependence, MuscaDet is nevertheless able to classify most of the day as cloud-free, owing to the tolerance term included in the detection criterion. In the second example (SPO 2014-11-09), MuscaDet correctly classifies cloudy periods as "possibly cloudy" and detects the subsequent cloud-free period, demonstrating consistent performance under these challenging high-latitude conditions.

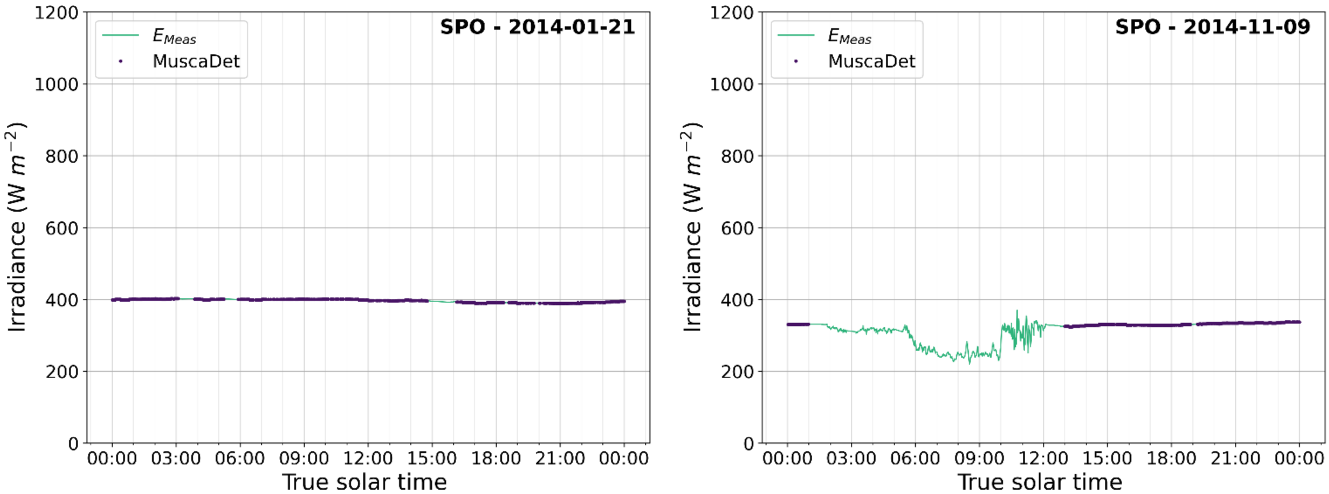

Figure 8: Measured irradiance (E_{meas}, green line) for two days at the South Pole Observatory station (SPO): 2014-01-21 (left) and 2014-11-09 (right). Blue dots indicate instants identified as cloud-free by MuscaDet.

Some false negatives were identified at the SPO station during the visual inspection. A representative case is shown in Fig. 9, for 2014-01-03. Although the day appears visually consistent with cloud-free conditions— exhibiting very limited variability in E_{meas}—MuscaDet does not identify correctly cloud-free conditions between 04:00 and 09:00.

A zoomed view of this interval (Fig. 9, right panel) reveals small fluctuations in E_{meas} that exceed the detection thresholds, despite not being clearly associated with cloud presence. These variations may reflect natural atmospheric variability, such as changes in aerosol or water vapour content. This case highlights a limitation of the current threshold configuration and suggests that further refinement could improve robustness to atmospheric variability while maintaining the conservative character of the detection approach.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

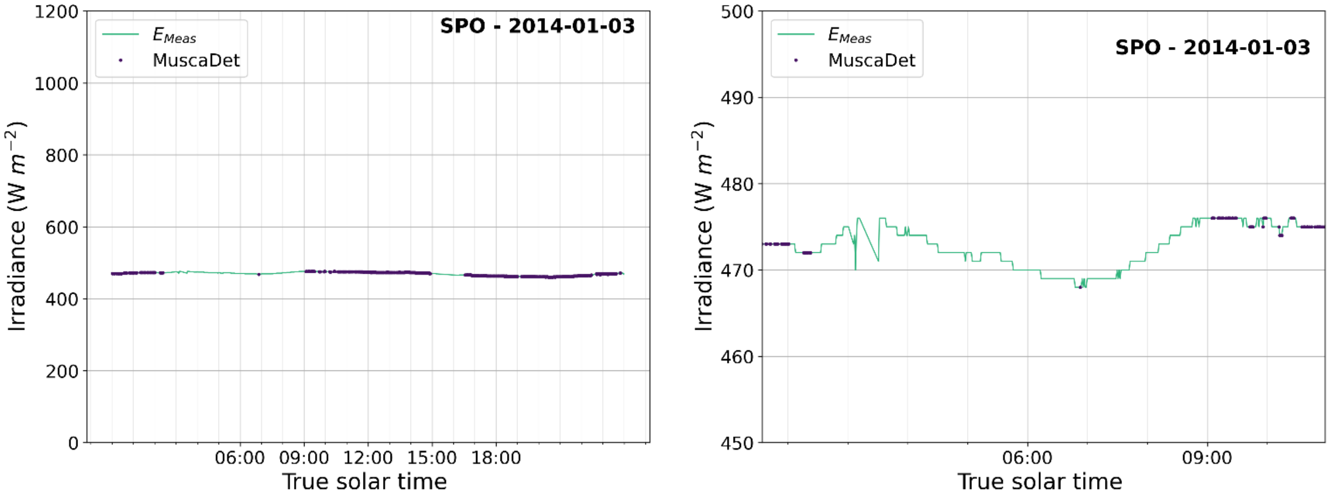

Figure 9: Measured irradiance (E_{meas}, green line) at the South Pole Observatory station on 2014-01-04. Left panel: full-day. Right panel: excerpt between 02:00 and 10:00, highlighting the period during which false negatives are observed. Blue dots indicate instants identified as cloud-free by MuscaDet.

6 Comparison with the cloud-free detector of Lefèvre et al. (2013) and implication for the verification of the McClear service

6.1 Comparison with the cloud-free detector of Lefèvre et al. (2013)

Among the few cloud-free detectors being independent from the outputs of a clear-sky model as discussed in Section 2, that of Lefèvre et al. (2013) is particularly relevant here, as it is currently used for the quarterly verification of the CAMS McClear service. For the sake of the simplicity in reading, this detector is simply called Lefèvre in this Section 6. Hence, this Section aims at comparing the results of the verification of the McClear service when using Lefèvre or MuscaDet.

Consequently, the results presented in the following may differ from those reported in the CAMS quarterly reports. These differences are partly explained by the fact that the CAMS protocol includes quality controls for measurements, unlike the present study where no additional controls are applied, and partly by the summarization of measurement, which is 10 min in the CAMS protocol and 1 min in the present study.

Both detectors were applied to the set of measurements considered in this study, except for SPO, where the McClear service is unavailable due to the lack of SRTM elevation data at high latitudes. Fig. 10 exhibits the number of cloud-free instants produced by each detector as a function of θ_s at each station. Results from MuscaDet are shown in blue, while those from Lefèvre are shown in green.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

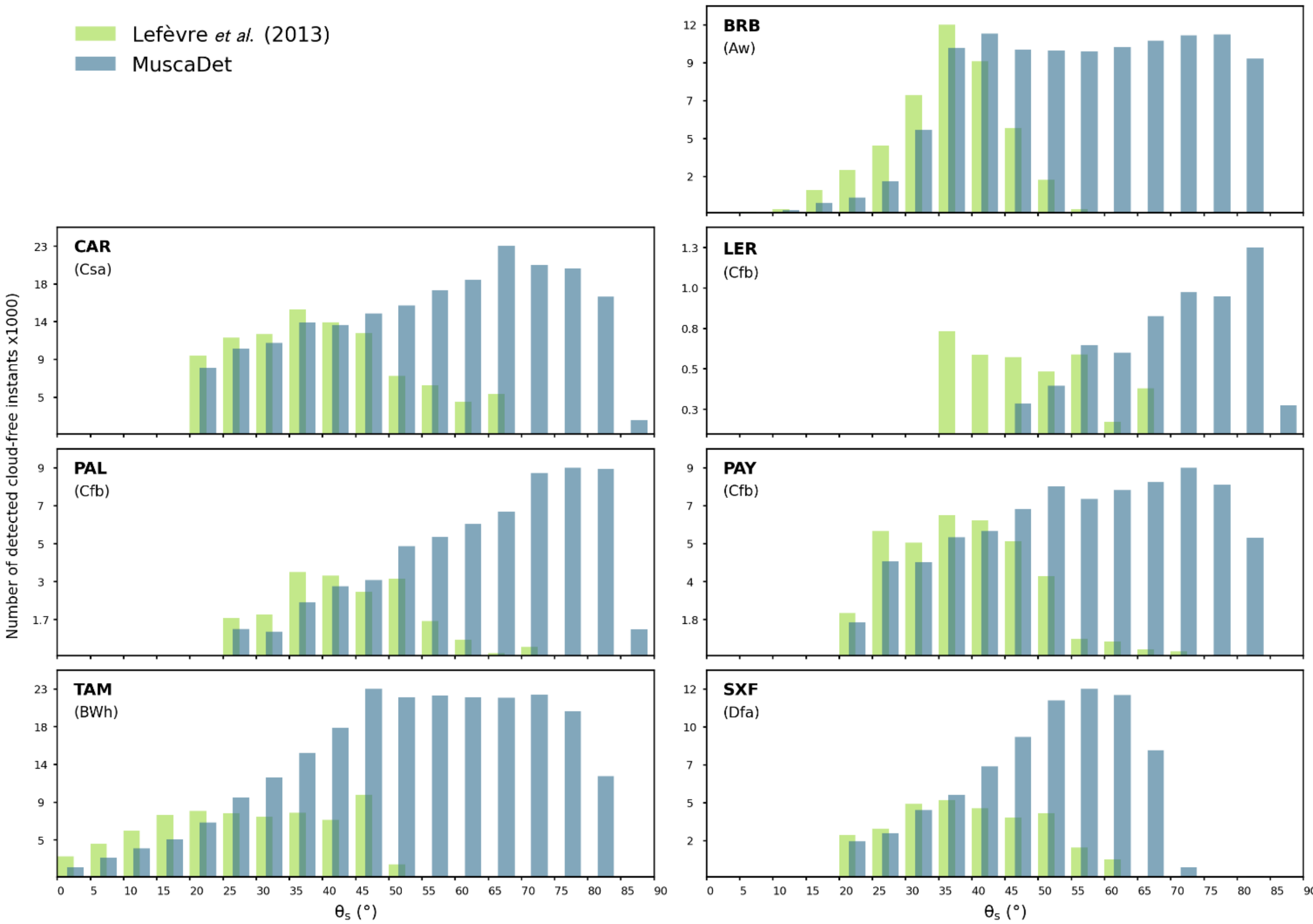

Figure 10 : Number of cloud-free instants detected by MuscaDet (blue bars) and by Lefèvre (green bars), shown as a function of θ_s at each station.

Fig. 10 highlights clear differences between the two detectors at every station. MuscaDet identifies more cloud-free instants than Lefèvre at high θ_s while it is the opposite at low θ_s. This pattern is observed across all seven BSRN stations.

A visual inspection of the measurements suggests that the low number of detections obtained by Lefèvre at high θ_s may be attributed to false negatives—i.e., missed cloud-free instants— that results from its use of a modified clearness index K_T' that does not fully remove the dependence on θ_s, particularly at high θ_s (see Fig. 1). The residual dependence of K_T' on θ_s may be interpreted by Lefèvre as a cloud-induced variability in irradiance, leading to systematic false negatives in this range. The inspection reveals also that the upper threshold of 0.3 of the diffuse fraction K in Lefèvre leads to a substantial number of false negatives. In cloud-free conditions, K can vary significantly and can reach high values at high θ_s, and a fixed threshold of 0.3 becomes overly restrictive. Batlles et al. (2000) proposed an upper threshold of K depending on θ_s. This dependence is illustrated in Fig. 11 using measurements at Tamanrasset over several consecutive days. The upper panel shows the diurnal evolution of E_{meas} and D_{meas} together with $E_{McClear}$ and $D_{McClear}$. The lower panel exhibits the temporal evolution of K, along

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

with the diffuse fraction computed with $E_{McClear}$ and $D_{McClear}$ for illustration and the thresholds of 0.3 and 0.9 used by respectively Lefèvre and MuscaDet. K exhibits pronounced diurnal variability, even under cloud-free conditions, and may approach unity at high θ_s. Fig. 11 illustrates that applying a fixed threshold of 0.3 leads to the rejection of valid cloud-free instants, particularly during morning and evening periods.

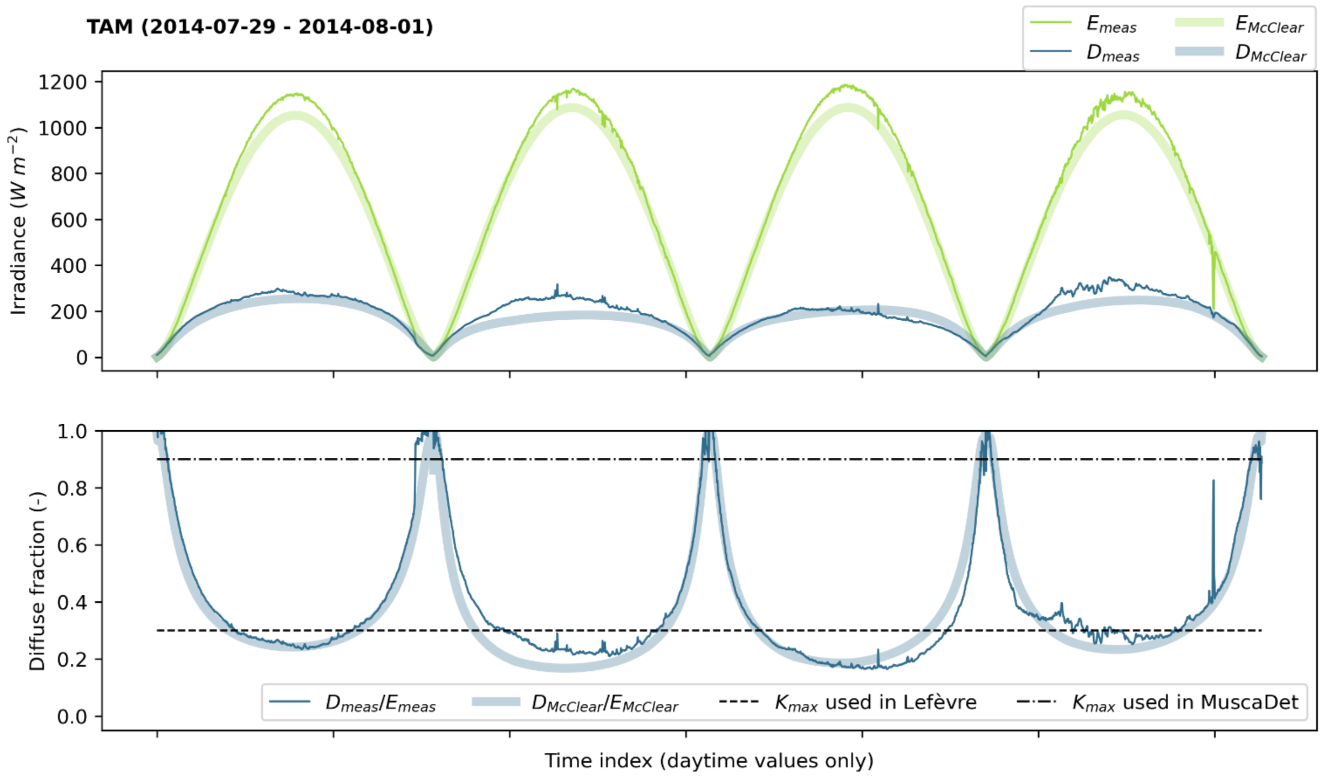

Figure 11: Daily profiles of E_{meas} and D_{meas} and the resulting diffuse fraction K at Tamanrasset over several days. Upper panel: green and light green lines represent E_{meas} and $E_{McClear}$, respectively, while blue and light blue lines D_{meas} and $D_{McClear}$. Lower panel: the diffuse fraction computed from measurements is shown in blue, and that computed from McClear in light blue. The upper threshold used by Lefèvre ($K_{\max} = 0.3$) is drawn in black dashed line, and that used in MuscaDet ($K_{\max} = 0.9$) is drawn in black dash-dotted line.

Another notable difference between the two detectors is the number of cloud-free instants detected at low θ_s. As an illustration, Fig. 12 exhibits four consecutive daily profiles of E_{meas} at Tamanrasset. Also drawn are the outputs of both detectors marked with various colors. Instants detected as cloud-free by both detectors are in yellow, those detected by MuscaDet only are in blue, and those detected by Lefèvre only are in green. The blue markers (MuscaDet only) corroborate the previous analysis and highlights the ability of MuscaDet to identify cloud-free instants at high θ_s, which are not detected by Lefèvre. Conversely, instants detected as cloud-free by Lefèvre only (green markers) are predominantly associated with large E_{meas}, corresponding to low θ_s, in agreement with the tendency noted in Fig. 11 of Lefèvre to identify a greater number of cloud-free instants than MuscaDet at low θ_s.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

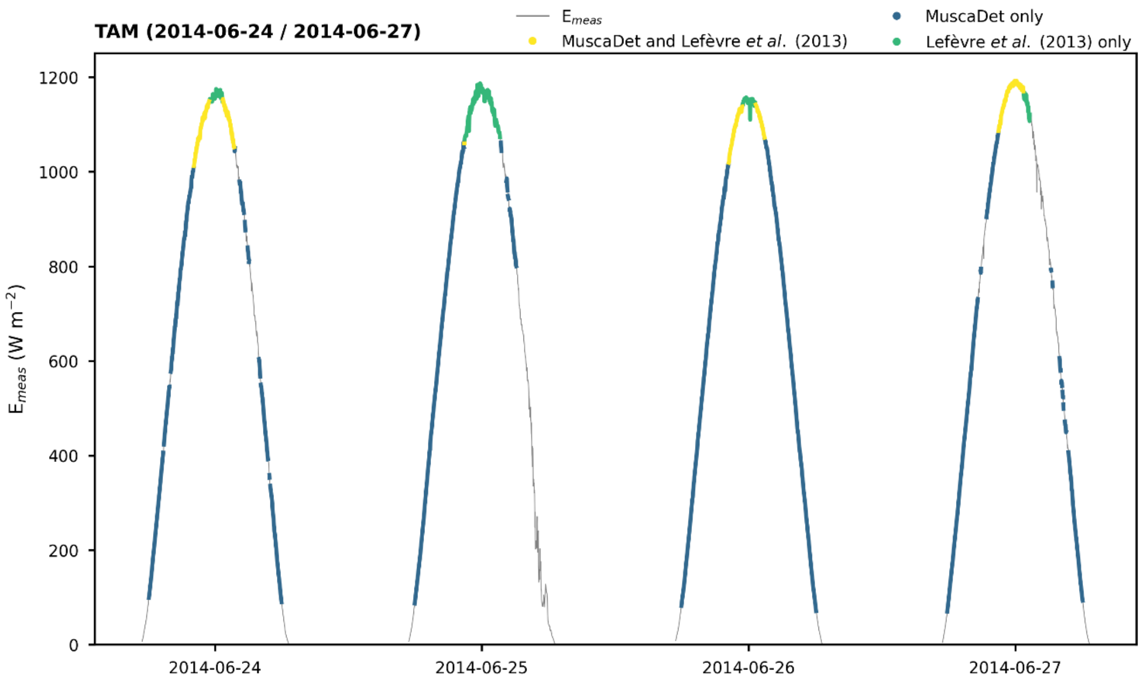

Figure 12 : Daily profiles of E_{meas} during four consecutive days at Tamanrasset. Instants detected as cloud-free by both detectors are in yellow, those detected by MuscaDet are in green, and those detected by Lefèvre only are in blue.

The green markers in Fig. 12 corroborate the previous analysis by highlighting the ability of MuscaDet to identify cloud-free instants at high solar zenithal angles, where Lefèvre exhibits systematic false negatives. Conversely, instants classified as cloud-free by Lefèvre, but not by MuscaDet (blue markers) are predominantly associated with higher values of E_{meas}, corresponding to lower solar zenithal angles. These observations are consistent with the tendency noted in Fig. 11 Lefèvre to identify a greater number of cloud-free instants than MuscaDet at low solar zenithal angles.

A closer examination of the blue markers in Fig. 12 shows that many of these instants correspond to two of the cases discussed in Section 5, namely "Low-amplitude fluctuations in irradiance in a globally cloud-free day" (fourth day in Fig. 12) and "Isolated dips in irradiance induced by passing clouds" (third day in Fig. 12). While these instants are "probably cloudy" according to MuscaDet, the test on the standard deviation of K_T', computed over a 180 min moving window in Lefèvre may reveal inaccurate as a brief dip in E_{meas} affects only a few minutes within this interval and the standard deviation over the window may remain below the prescribed threshold, yielding a detection of a cloud-free instant. This is also the case if a fluctuation in E_{meas} has limited amplitude and duration while the remainder of the window remains stable and close to cloud-free conditions, yielding a standard deviation less than the threshold. In the two cases, Lefèvre detects these instants as cloud-free, even though visual inspection suggests the likely presence of clouds.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

6.2 Implication for the verification of the McClear service

Lefèvre is currently used in the quarterly verification of the CAMS McClear service and may be replaced by MuscaDet with the aim at improving this verification. This sub-section aims at documenting the impact of the change in detector on the results of the verification. As an illustration, Fig. 13 exhibits 2-D histograms of E_{McClear} and E_{meas} for cloud-free instants only at Palaiseau. The left panel corresponds to instants identified as cloud-free by Lefèvre, whereas the right panel is for MuscaDet.

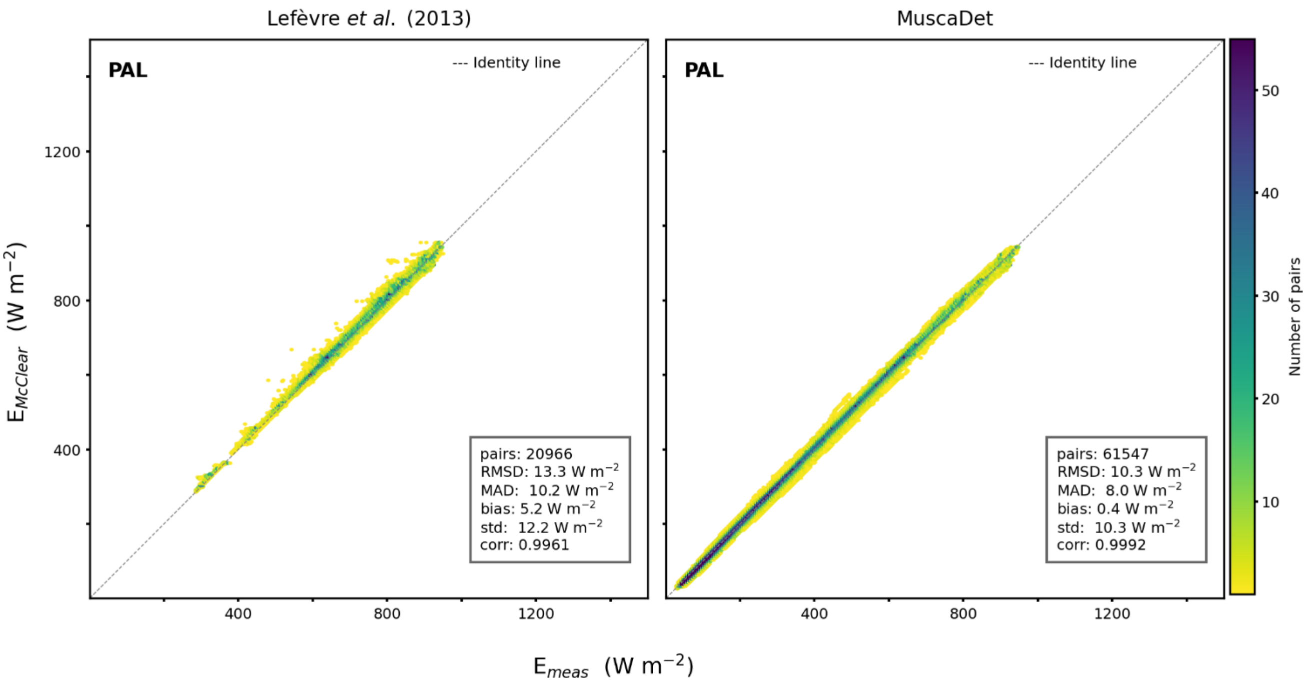

Figure 13 : 2D-histograms of E_{McClear} (y-axis) and E_{meas} (x-axis) for cloud-free instants at Palaiseau. Left panel: Lefèvre, right panel: MuscaDet. Each point represents a 1 min pair (E_{McClear}, E_{meas}) and the colour scale indicates the number of pairs. Several difference metrics are provided in the lower-right corner of each panel, where MAD stands for mean absolute difference, std for standard deviations of differences and corr for correlation coefficient.

The comparison of the 2D-histograms in Fig. 13 indicates that using MuscaDet instead of Lefèvre reduces the scattering of pairs around the identity line. In addition, MuscaDet yields more than double the number of paired samples, particularly due to its higher detection rate at large θ_s.

These differences are expected to produce two noticeable effects if MuscaDet is used in the verification of McClear. Firstly, instants affected by small-scale variability in E_{meas} and isolated irradiance dips are excluded, thereby reducing the variance of the differences between McClear and the measurements. Secondly, MuscaDet provides a more representative sampling of cloud-free instants at high solar zenithal angles while reducing the number of instants classified as cloud-free at low angles, yielding a more balanced statistical distribution of angles in the verification process.

The two detectors produce noticeably different statistics quantifying the differences between $\mathrm{E}_{McClear}$ and E_{meas} in Fig. 13. The correlation coefficient is 0.9960 with Lefèvre and 0.9990 with MuscaDet. The bias is respectively 5.2 and 0.4 W m^{-2}. The

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

standard deviation of differences and the RMSD are respectively 12.2 and 10.3 W m^{-2}, and 13.3 and 10.3 W m^{-2}. The improvement in these statistics between Lefèvre and MuscaDet is observed at most stations (Fig. 14).

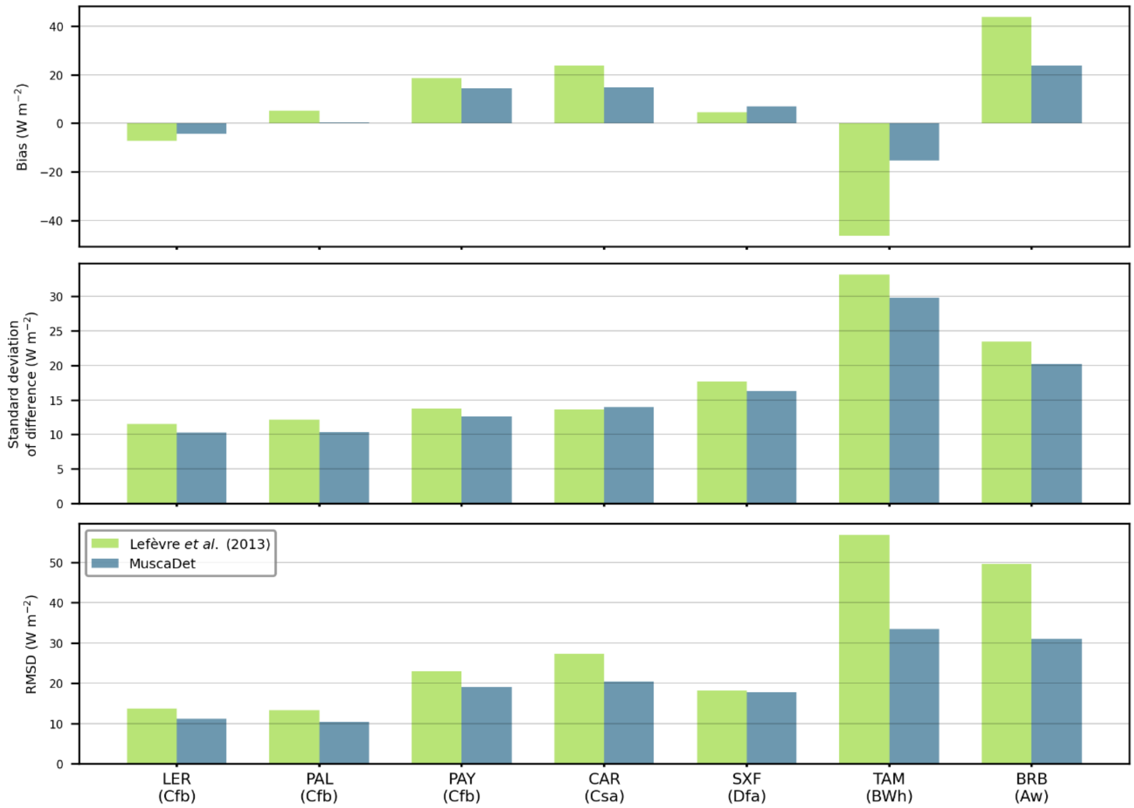

Figure 14 : Bias, standard deviation of differences and RMSD of $E_{McClear}$ vs E_{meas} at the 7 stations for each detector.

The observations above evidence that MuscaDet provides a fairly balanced statistical distribution of solar zenithal angles and irradiances. The replacement of Lefèvre by MuscaDet would therefore contribute to a more representative and consistent quarterly verification of the CAMS McClear service.

7 Conclusion

MuscaDet was designed as a robust alternative to cloud-free detection methods relying on external clear-sky models, with a particular focus on the operational verification of the McClear service within the Copernicus Atmospheric Monitoring Service. The proposed approach addresses limitations identified by Bright et al. (2020), as well as issues encountered with the detector of Lefèvre et al. (2013) currently used for the verification of the McClear service. Its primary objective is to ensure high confidence in detected cloud-free periods, while secondarily aiming to limit the number of false negatives—i.e., missed cloud-free instants.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

Building on the approach of Long and Ackerman (2000), MuscaDet relies on a local fit of the MLB function (Mueller et al., 2004), a simple parameterization describing the dependence of the irradiance on the solar zenithal angle under cloud-free conditions. Its main innovation lies in the combination of this parameterization with a multiscale analysis and an explicit consideration of uncertainties arising from both sensor measurements and atmospheric variability.

The verification shows that the method fulfills its primary objective of confidently identifying cloud-free conditions. Cloudy and cloud-free situations are consistently classified, while the use of multiple temporal scales and a stringent correlation criterion effectively prevent false positives (i.e. predicted cloud-free while cloudy), including those induced by short-term irradiance fluctuations. This conservative design, although leading to missed detections in marginal cases, reflects a deliberate trade-off between completeness and reliability, and is particularly well suited to applications such as the verification of clear-sky models.

Compared to the detector of Lefèvre et al. (2013), MuscaDet yields a more balanced distribution of detections with respect to the solar zenithal angle, detecting more cloud-free cases at high angles while reducing the occurrence of false positives at low angles. The resulting differences in the verification of the McClear service were analyzed in detail and could be attributed to methodological choices in the two detectors, in particular explaining the occurrence of spurious detections with the detector of Lefèvre et al. (2013). These results indicate that MuscaDet provides a more reliable basis for the verification of clear-sky model. These results suggest that MuscaDet could be considered as a potential alternative to the detector of Lefèvre et al. (2013) for the operational quarterly verification of the McClear service.

Despite these encouraging results, several limitations and open questions remain. In particular, low-amplitude variability in irradiance in globally cloud-free conditions was frequently observed. The origin of these fluctuations—whether related to thin cirrus, natural variability in atmospheric constituents, or measurement uncertainty—remains unclear. As a result, such situations are conservatively not classified as cloud-free, although their frequency at some sites suggests that improved characterization is needed. Future work could combine sky-imager observations with AERONET data to better understand this low-amplitude variability in irradiance and assess whether adaptations of the detection criteria are warranted under such conditions.

More generally, the definition of the tolerance parameter used in the detection criterion remains an important aspect of the method, as it governs the relaxation of the model to account for natural atmospheric variability and measurement-related deviations from the expected cloud-free behaviour. Although its current formulation accounts for the dependence on the solar zenithal angle, it remains heuristic. Its physical basis could be strengthened by incorporating models of pyranometer uncertainty and by accounting for the temporal variability of aerosols and water vapor, thereby improving the adaptability of the method across instruments and environments.

In summary, MuscaDet provides a robust, conservative, and physically-grounded framework for the detection of cloud-free instants, particularly well-suited to applications requiring high-confidence identification of cloud-free conditions. Further developments focusing on uncertainty characterization and improved observational constraints are expected to enhance its performance and extend its applicability.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

Appendix A: Derivation of the decision criterion

As explained in Sects. 4.1 and 4.2, MuscaDet relies on the agreement between E_{meas} and E_{MLB}. Quantifying this agreement requires the selection of an appropriate statistical metric. Several candidates are commonly used in similar contexts, including:

- the root-mean-square difference (RMSD) and the mean absolute different (MAD), which are expressed in W m^{-2} and therefore require thresholds that depend explicitly on irradiance level;
- their normalized counterparts, the normalized RMSD (nRMSD) and normalized MAD (nMAD), which alleviate the dependence on irradiance but may become unstable or poorly defined at low irradiance;
- the Pearson correlation coefficient (ρ) and the coefficient of determination (R^2), which are both dimensionless and bounded.

The normalized nature of ρ and R^2 makes them particularly attractive for cloud-free detectors, as it allows the definition of a decision threshold that is independent of levels in irradiance. Despite this apparent similarity, an important conceptual distinction exists between the two metrics. The coefficient of determination R^2 evaluates agreement based on the relative magnitudes of the overall variances of the two signals over the considered time window. It does not explicitly assess the temporal alignment of fluctuations at the scale of individual measurements. As a result, R^2 may attain high values in situations where the model and the measurements exhibit comparable levels of variability, even if their short-term fluctuations are not temporally synchronized. Such cases are undesirable for cloud-free detector, as they may lead to the erroneous classification of cloud-free conditions. In contrast, the Pearson correlation coefficient explicitly quantifies the degree of linear co-variation between the two signals over time, making it more sensitive to temporal mismatches and therefore better suited to the requirements of the present application.

Based on these considerations, a satisfactory level of agreement between E_{meas} and $f_{\mathrm{MLB}}(\theta_s)$ is achieved when the correlation coefficient exceeds a prescribed threshold ρ_{min}:

$$\rho\left(E_{meas}, f_{\mathrm{MLB}}(\theta_s)\right) > \rho_{min} \tag{A-1}$$

Preliminary tests using Eq. (A-1) with an illustrative threshold value ($\rho_{\mathrm{min}} = 0.99$) revealed the occurrence of false negatives under conditions where the dependence of E_{meas} on θ_s is weak. In such cases, the variance of $f_{\mathrm{MLB}}(\theta_s)$ over the considered time window is small, and a large fraction of the observed temporal variability in E_{meas} likely arises from measurement noise or from small-scale variations in atmospheric composition.

In practice, it is often assumed that E_{meas} can be expressed as the sum of the actual irradiance E_{true} and a measurement noise term $\varepsilon_{\mathrm{meas}}$ (Vuilleumier et al., 2014). Likewise, due to temporal variations in atmospheric constituents (e.g., aerosols or water vapour), a residual difference $\varepsilon_{\mathrm{MLB}}$ exists between $f_{\mathrm{MLB}}(\theta_s)$ and the irradiance in cloud-free conditions E_{c}:

$$E_{meas} = E_{true} + \varepsilon_{meas} \tag{A-2}$$

$$f_{\mathrm{MLB}}(\theta_s) = E_c + \varepsilon_{MLB} \tag{A-3}$$

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

The use of additive noise to represent variability in atmospheric conditions is a simplification, as such variability primarily affects atmospheric transmittance and would be more appropriately described by a multiplicative term. However, the dominant dependence of irradiance on solar zenithal angle and slowly varying atmospheric conditions is captured by the MLB fit. The remaining variability can therefore be interpreted as a residual fluctuation around the fitted signal, which justifies the use of additive noise.

If E_{true} and E_{c} were directly available, applying Eq. (A-1) would provide an effective test. In practice, however, the correlation is computed using E_{meas} and $f_{\mathrm{MLB}}(\theta_s)$, which differs from $\rho(E_{true}, E_c)$ due to measurement noise and residual atmospheric variability. The uncertainty terms $\varepsilon_{\mathrm{meas}}$ and $\varepsilon_{\mathrm{MLB}}$ must therefore be explicitly accounted for in the correlation-based test.

For the sake of simplicity and to obtain a tractable formulation, the two uncertainty contributions introduced above are combined into a single error term ε, which accounts for both measurement noise and atmospheric variability. This combined error term is associated with the measured irradiance E_{meas} and is assumed to be centered and normally distributed. Under this assumption, E_{meas} can be written as the sum of E_{true} and a stochastic term following a normal distribution with zero mean and standard deviation σ_{tol} representing the model tolerance used to account for the uncertainty terms $\varepsilon_{\mathrm{meas}}$ and $\varepsilon_{\mathrm{MLB}}$, and E_c is assumed equal to $f_{\mathrm{MLB}}(\theta_s)$:

$$E_{meas} = E_{true} + \varepsilon\,, \quad \varepsilon \sim \mathcal{N}(0, \sigma_{tol}) \tag{A-4}$$

$$f_{\mathrm{MLB}}(\theta_s) \approx E_c \tag{A-5}$$

Assuming that the correlation-based test (Eq. A-1) is satisfied for E_{true} and E_{c}, i.e. $\rho(E_{True}, E_c) > \rho_{min}$, the objective is to reformulate this test in terms of E_{meas} and $f_{\mathrm{MLB}}(\theta_s)$, while explicitly accounting for σ_{tol}. To this end, the Pearson correlation coefficient is expressed in its standard form as a function of the covariance and variances of the two variables:

$$\rho(E_{True}, E_{clear}) = \frac{Cov(E_{True}, E_c)}{\sqrt{Var(E_{True})\, Var(E_c)}} > \rho_{min} \tag{A-6}$$

Assuming that ε is centered (zero mean) and statistically independent of E_{c}, and using the approximation $E_{\mathrm{MLB}} \approx E_{\mathrm{c}}$, the covariance between the measured irradiance and the fitted MLB irradiance can be expressed as

$$\begin{aligned} Cov\left(E_{meas}, f_{\mathrm{MLB}}(\theta_s)\right) &= Cov(E_{true} + \varepsilon, E_c) \\ &= Cov(E_{true}, E_c) + Cov(\varepsilon, E_c) \\ &= Cov(E_{true}, E_c) \end{aligned} \tag{A-7}$$

where the last equality follows from the zero mean of ε and its independence with E_{c}, which imply $Cov(\varepsilon, E_{\mathrm{c}}) = 0$.

In contrast, the variance of E_{meas} is affected by ε. Since E_{true} and ε are independent, the variance of E_{meas} satisfies

$$Var(E_{meas}) = Var(E_{true} + \varepsilon) = Var(E_{true}) + Var(\varepsilon) = Var(E_{true}) + \sigma_{tol}^2. \tag{A-8}$$

Finally, the correlation-based test can be expressed in terms of E_{meas}, $f_{\mathrm{MLB}}(\theta_s)$, and σ_{tol} as follows:

$$\frac{Cov(E_{meas}, E_{MLB})}{\sqrt{max(0, Var(E_{meas}) - \sigma_{tol}^2)\, Var(f_{\mathrm{MLB}}(\theta_s))}} > \rho_{min} \tag{A-9a}$$

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

Or equivalently,

$$\rho\left(E_{meas}, f_{\mathrm{MLB}}(\theta_s)\right) > \rho_{min}\sqrt{\frac{max(0, Var(E_{meas})-\sigma_{tol}^2)}{Var(E_{meas})}} \qquad \text{(A-9b)}$$

The maximum functions in Eqs. (A-9a) and (A-9b) ensure numerical stability when $Var(E_{meas}) < \sigma_{tol}^2$. In such cases, the left-hand side of Eq. (A-9a) tends to infinity while the right-hand side of Eq. (A-9b) tends to zero, causing the test to classify the corresponding time step as cloud-free, except in some cases of overcast skies. This behaviour is consistent with the underlying assumption that measurements with very low variability are unlikely to reflect cloud-induced fluctuations.

Code availability. The algorithm is available on GitHub (https://github.com/YvesMSaintDrenan/MuscaDet-cloud-free-detection, last access: 2026-05-22)

Author contribution: YMSD and LW designed the algorithm, performed the analysis, and prepared the manuscript. CV reviewed the development of the statistical metric implemented in MuscaDet. ML carried out the comparison with the detector currently used in CAMS. ML, CV and YE contributed to the writing and revision of the manuscript.

Competing interests: The contact author has declared that none of the authors has any competing interests.

Supplement. The supplement related to this article is available online at (…).

Acknowledgements: This work was supported by the Copernicus Atmosphere Monitoring Service (CAMS), operated by the European Centre for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Union, and by the SciDoSol sponsorship chair at Mines Paris – PSL, funded by TotalEnergies, RTE, Somfy, TSE, and CalibSun. The authors thank all operators of the ground stations of the Baseline Surface Radiation Network (BSRN) for their valuable measurements, and the Alfred Wegener Institute for hosting the BSRN database. During the redaction of this paper, tools based on artificial intelligence (AI) were used. Specifically, "ChatGPT" by assisted in improving the wording and formulation of this manuscript. All ideas and concepts are developed without AI assistance exclusively by the authors.

References

Alia-Martinez, M., Antonanzas, J., Urraca, R., Martinez-de-Pison, F. J., and Antonanzas-Torres, F.: Benchmark of algorithms for solar clear-sky detection, Journal of Renewable and Sustainable Energy, 8, 033703, https://doi.org/10.1063/1.4950948, 2016.

Batlles, F. J., Olmo, F. J., Tovar, J., and Alados-Arboledas, L.: Comparison of cloudless sky parameterizations of solar irradiance at various Spanish midlatitude locations, Theoretical and Applied Climatology, 66, 81–93, https://doi.org/10.1007/s007040070034, 2000.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen-Geiger climate classification maps at 1-km resolution, Sci Data, 5, 180214, https://doi.org/10.1038/sdata.2018.214, 2018.

Blanc, P., Espinar, B., Geuder, N., Gueymard, C., Meyer, R., Pitz-Paal, R., Reinhardt, B., Renné, D., Sengupta, M., Wald, L., and Wilbert, S.: Direct normal irradiance related definitions and applications: The circumsolar issue, Solar Energy, 110, 561–577, https://doi.org/10.1016/j.solener.2014.10.001, 2014.

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

Bright, J. M., Sun, X., Gueymard, C. A., Acord, B., Wang, P., and Engerer, N. A.: Bright-Sun: A globally applicable 1-min irradiance clear-sky detection model, Renewable and Sustainable Energy Reviews, 121, 109706, https://doi.org/10.1016/j.rser.2020.109706, 2020.

Ellis, B. H., Deceglie, M., and Jain, A.: Automatic detection of clear-sky periods using ground and satellite based solar resource data, 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC), 2293–2298, 2018.

Gschwind, B., Wald, L., Blanc, P., Lefèvre, M., Schroedter-Homscheidt, M., and Arola, A.: Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions – McClear-v3, metz, 28, 147–163, https://doi.org/10.1127/metz/2019/0946, 2019.

Gueymard, C. A.: REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation – Validation with a benchmark dataset, Solar Energy, 82, 272–285, https://doi.org/10.1016/j.solener.2007.04.008, 2008.

Gueymard, C. A., Bright, J. M., Lingfors, D., Habte, A., and Sengupta, M.: A posteriori clear-sky identification methods in solar irradiance time series: Review and preliminary validation using sky imagers, Renewable and Sustainable Energy Reviews, 109, 412–427, https://doi.org/10.1016/j.rser.2019.04.027, 2019.

Ineichen, P., Barroso, C. S., Geiger, B., Hollmann, R., Marsouin, A., and Mueller, R.: Satellite Application Facilities irradiance products: hourly time step comparison and validation over Europe, International Journal of Remote Sensing, 30, 5549–5571, https://doi.org/10.1080/01431160802680560, 2009.

Inman, R. H., Edson, J. G., and Coimbra, C. F. M.: Impact of local broadband turbidity estimation on forecasting of clear sky direct normal irradiance, Solar Energy, 117, 125–138, https://doi.org/10.1016/j.solener.2015.04.032, 2015.

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., and Morcrette, J.-J.: McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions, Atmos. Meas. Tech., 6, 2403–2418, https://doi.org/10.5194/amt-6-2403-2013, 2013.

Li, H., Zhang, M., Wang, L., Su, X., and Lu, Y.: Effects of different types of aerosols on diffuse radiation based on global AERONET, JGR Atmospheres, 130, e2024JD042701, https://doi.org/10.1029/2024JD042701, 2025.

Long, C. N. and Ackerman, T. P.: Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects, J. Geophys. Res., 105, 15609–15626, https://doi.org/10.1029/2000JD900077, 2000.

https://doi.org/10.5194/egusphere-2026-3204
Preprint. Discussion started: 20 August 2026

Mamajek, E. E., Prsa, A., Torres, G., Harmanec, P., Asplund, M., Bennett, P. D., Capitaine, N., Christensen-Dalsgaard, J., Depagne, E., Folkner, W. M., Haberreiter, M., Hekker, S., Hilton, J. L., Kostov, V., Kurtz, D. W., Laskar, J., Mason, B. D., Milone, E. F., Montgomery, M. M., Richards, M. T., Schou, J., and Stewart, S. G.: IAU 2015 Resolution B3 on Recommended Nominal Conversion Constants for Selected Solar and Planetary Properties, https://doi.org/10.48550/arXiv.1510.07674, 26 October 2015.

Mueller, R., Dagestad, K. F., Ineichen, P., Schroedter, M., Cros, S., Dumortier, D., Kuhlemann, R., Olseth, J. A., Piernavieja, G., Reise, C., Wald, L., and Heinnemann, D.: Rethinking satellite-based solar irradiance modelling: The SOLIS clear-sky module, Remote Sensing of Environment, 91, 160–174, https://doi.org/10.1016/j.rse.2004.02.009, 2004.

Ohmura, A., Gilgen, H., Hegner, H., Müller, G., Wild, M., Dutton, E. G., Forgan, B., Fröhlich, C., Philipona, R., Heimo, A., König-Langlo, G., McArthur, B., Pinker, R., Whitlock, C. H., and Dehne, K.: Baseline surface radiation network (BSRN/WCRP): New precision radiometry for climate research, Bull. Amer. Meteor. Soc., 79, 2115–2136, https://doi.org/10.1175/1520-0477(1998)079%253C2115:BSRNBW%253E2.0.CO;2, 1998.

Perez, R., Ineichen, P., Seals, R., and Zelenka, A.: Making full use of the clearness index for parameterizing hourly insolation conditions, Solar Energy, 45, 111–114, https://doi.org/10.1016/0038-092X(90)90036-C, 1990a.

Perez, R., Ineichen, P., Seals, R., Michalsky, J., and Stewart, R.: Modeling daylight availability and irradiance components from direct and global irradiance, Solar Energy, 44, 271–289, https://doi.org/10.1016/0038-092X(90)90055-H, 1990b.

Polo, J., Zarzalejo, L. F., Martín, L., Navarro, A. A., and Marchante, R.: Estimation of daily Linke turbidity factor by using global irradiance measurements at solar noon, Solar Energy, 83, 1177–1185, https://doi.org/10.1016/j.solener.2009.01.018, 2009.

Quesada-Ruiz, S., Linares-Rodríguez, A., Ruiz-Arias, J. A., Pozo-Vázquez, D., and Tovar-Pescador, J.: An advanced ANN-based method to estimate hourly solar radiation from multi-spectral MSG imagery, Solar Energy, 115, 494–504, https://doi.org/10.1016/j.solener.2015.03.014, 2015.

Reno, M. J. and Hansen, C. W.: Identification of periods of clear sky irradiance in time series of GHI measurements, Renewable Energy, 90, 520–531, https://doi.org/10.1016/j.renene.2015.12.031, 2016.

Saint-Drenan, Y.-M. and Wald, L.: On the assessment of hourly means of solar irradiance at ground level in clear-sky conditions by the ERA5, JRA-3Q, and MERRA-2 reanalyses, Atmosphere, 16, 949, https://doi.org/10.3390/atmos16080949, 2025.

Vuilleumier, L., Hauser, M., Félix, C., Vignola, F., Blanc, P., Kazantzidis, A., and Calpini, B.: Accuracy of ground surface broadband shortwave radiation monitoring, JGR Atmospheres, 119, https://doi.org/10.1002/2014JD022335, 2014.

Wandji Nyamsi, W., Saint-Drenan, Y.-M., Arola, A., and Wald, L.: Further validation of the estimates of the downwelling solar radiation at ground level in cloud-free conditions provided by the McClear service: the case of Sub-Saharan Africa and the Maldives Archipelago, Atmos. Meas. Tech., 16, 2001–2036, https://doi.org/10.5194/amt-16-2001-2023, 2023.

Xie, Y. and Liu, Y.: A new approach for simultaneously retrieving cloud albedo and cloud fraction from surface-based shortwave radiation measurements, Environ. Res. Lett., 8, 044023, https://doi.org/10.1088/1748-9326/8/4/044023, 2013.

Zhang, W., Kleiber, W., Florita, A. R., Hodge, B.-M., and Mather, B.: A stochastic downscaling approach for generating high-frequency solar irradiance scenarios, Solar Energy, 176, 370–379, https://doi.org/10.1016/j.solener.2018.10.019, 2018.