

A solar resource classification algorithm for global horizontal irradiance time series based on frequency domain analysis

F SCI

Cite as: J. Renewable Sustainable Energy 13, 036102 (2021); <https://doi.org/10.1063/5.0045032>
Submitted: 22 January 2021 . Accepted: 18 April 2021 . Published Online: 28 June 2021

 C. Lewis, J. M. Strauss, and A. J. Rix

COLLECTIONS

 This paper was selected as Featured

 This paper was selected as Scilight



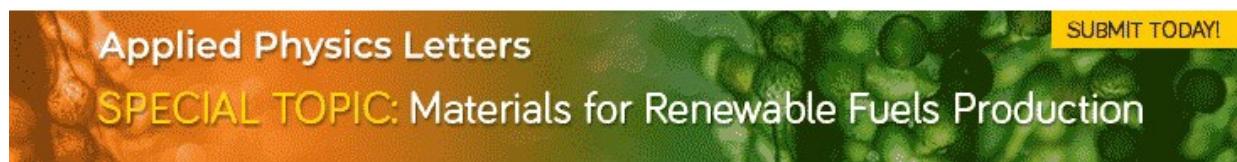
[View Online](#)



[Export Citation](#)



[CrossMark](#)



Applied Physics Letters

SUBMIT TODAY!

SPECIAL TOPIC: Materials for Renewable Fuels Production

A solar resource classification algorithm for global horizontal irradiance time series based on frequency domain analysis



Cite as: J. Renewable Sustainable Energy 13, 036102 (2021); doi: 10.1063/5.0045032

Submitted: 22 January 2021 · Accepted: 18 April 2021 ·

Published Online: 28 June 2021



View Online



Export Citation



CrossMark

C. Lewis,^{a)} J. M. Strauss, and A. J. Rix

AFFILIATIONS

Department of Electrical and Electronic Engineering, Stellenbosch University, Stellenbosch, South Africa

^{a)}Author to whom correspondence should be addressed: carmenl@sun.ac.za

ABSTRACT

A novel daily and 15-min period solar resource classification algorithm, based only on the Fourier transform of global horizontal irradiance (GHI) and calculable parameters, is presented. These methods can identify GHI series that are not attenuated by clouds, whether the clouds may or may not be visible. Additional sky condition classes are also adopted in which the representative conditions of the remaining days or periods are classified. Although the proposed day classification is too conservative to be widely applicable, the 15-min period classification is generally in good agreement with the Perez sky clearness index while providing greater sensitivity to turbid conditions and algorithm flexibility. It is shown, however, that the adoption of the clearness index in the proposed 15-min period method less successfully distinguished between clear and constant overcast conditions under winter continental conditions due to relatively large solar zenith angles. It is also shown that the proposed 15-min period and Reno and Hansen statistical methods are generally in good agreement on the detection of clear conditions under temperate conditions. The classification results of the proposed method are further shown in comparison to high-resolution sky images to elucidate the character of each class and the classification algorithm flexibility.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0045032>

I. INTRODUCTION

A. Photovoltaic (PV) power and the variability of solar energy

A significant source of uncertainty and variability in photovoltaic (PV) power originates from the capricious nature of the atmosphere and the extinction event within, which affects solar radiation transmission. It is not only climate heterogeneity that introduces variability in energy sources but also weather variability, spatial characteristics (such as latitude, altitude, level of urbanization, and geographical features), and temporal characteristics (such as seasonality and anomaly events).

Understanding atmospheric variability, the physical properties of transitional cloud cover and interdependence of solar radiation are therefore critical in solar radiation modeling and forecasting.^{1,2} However, due to the inherent complexity of solar radiation transmittance and often a lack of collocated measurements to describe the intrinsic variability, contributions to radiation extinction are still only predicted moderately well at higher spatial and temporal resolutions.³

Prior research suggests that the intrinsic variability in solar radiation and therefore PV power production introduce several power system challenges, which include adequate accuracy in estimating PV

system yield, energy storage dimensioning, power quality divergence, and generation and load balancing.^{4,5} Technical challenges in the distribution network also increase as PV grid penetration increases; these challenges include reverse power flow, network congestion, voltage unbalance, and harmonic distortion.⁶ Accurate high-resolution short-term radiation modeling and forecasting may therefore contribute to system operator decision-making to mitigate some of these challenges.

B. Sky condition classification in solar energy methodologies

Characterizing clear sky radiation is pivotal in many solar radiation evaluation, modeling, and forecasting methodologies.¹ As a result, identification of clear days and periods is required to assess these methodologies. Within the pursuit of sky conditions classification, clear sky conditions are inherently detected.

The definition of clear sky conditions varies in literature but is often based on visual interpretation, e.g., sky conditions are clear if no visible clouds are located within the sky dome,⁷ or clear sky radiation occurs under cloudless skies.⁸ However, even under these conditions, some atmospheric constituents indicate spatiotemporal variation and,

as a result, influence solar radiation transmission through the atmosphere varyingly. In some cases, clear sky irradiance is defined as the theoretical or idealized irradiance under ideal clear sky conditions, but the classification methodology is often more accurate if clear sky conditions are defined subjectively.⁹ More complex definitions of a clear sky are adopted if the objective is sky condition classification rather than the identification of clear sky conditions, despite the inherent subjectivity of classification. In this article, a clear sky refers to a certain period of time where clouds may or may not be visible within the sky dome, but it does not influence (attenuate) the global horizontal irradiance (GHI).

In literature, classification and detection methodologies are as abundant as forecasting methodologies, as observed in the reviews by Gueymard *et al.*¹⁰ and Bright *et al.*¹¹ Perhaps the most commonly used sky condition indicator is the *clearness index* (k_t), which is the ratio of surface measured global horizontal irradiance (GHI, F_G) and broadband irradiance at the top of the atmosphere,^{4,12}

$$k_t = \frac{F_G}{F_0 \cos \theta}, \quad (1)$$

where F_0 is the normal extraterrestrial irradiance and θ is the solar zenith angle. Equation (1) infers that a higher k_t value indicates a clearer day; although it is simple to calculate, a threshold or limit is required to detect clear periods, which is inherently dependent on the climatic character of the location under analysis.⁸ *Clearness index* thresholds for detecting clear days vary between 0.7 and 0.9 in literature;^{13,14} however, Gueymard *et al.*¹⁰ proposed that an average k_t threshold to detect an entire clear day is inaccurate since k_t is not only a function of cloudiness.

A more sophisticated version of the *clearness index* is the *Perez sky clearness index* (k_e), which is an indicator of sky conditions based on the ratio between the diffuse and direct irradiance,¹⁵

$$k_e = \frac{\frac{F_D + F_I}{F_D} + k\theta^3}{1 + k\theta^3}, \quad (2)$$

where $k = 1.041$ if the zenith angle θ is in radians or $k = 5.53 \times 10^{-6}$ if θ is in degrees. The index in (2) is used to classify sky conditions into eight discrete categories according to the thresholds in Table I.

Although Gueymard *et al.*¹⁰ proposed that the *Perez sky clearness index* has not been validated for higher temporal resolutions, Suárez-

TABLE I. The discrete *Perez sky clearness index* (k_e)¹⁵ categories for (2), with a color index as used in Sec. II C.

| Color | Category | Lower bound | Upper bound |
|-------|----------|-------------|-------------|
| 1 | Overcast | 1 | 1.065 |
| 2 | | 1.065 | 1.230 |
| 3 | | 1.230 | 1.500 |
| 4 | | 1.500 | 1.950 |
| 5 | | 1.950 | 2.800 |
| 6 | | 2.800 | 4.500 |
| 7 | | 4.500 | 6.200 |
| 8 | Clear | 6.200 | ∞ |

García *et al.*¹⁶ successfully applied this index to irradiance series with resolutions up to half-hourly.

A popular clear sky detection method is a statistical criteria algorithm by Reno and Hansen.⁷ The algorithm consists of five statistical criteria that have to be fulfilled by a measured GHI time series. In order to identify clear periods, the measured values are compared to theoretical clear sky model derived values in a sliding window. This method is often used in comparative detection and classification studies.^{10,11} However, it is found that this statistical criteria method performs less satisfactorily under more extreme conditions, such as high turbidity and low radiation.

C. Validation of sky condition classifications

It is proposed in literature¹¹ that qualitative validation of classification and detection methods is more appropriate, in part due to the subjectivity of classifications but also due to a lack of appropriate baseline references. As a result, quantitative validation should be limited and applied subjectively.

Several baseline references for classification and forecasting are presumed in literature, such as pre-classified CIE (Commission Internationale de l'Eclairage) skies,¹⁶ cloud data from a meteorological observer,¹⁷ ground-based aerosol optical depth retrievals,⁹ or validation with total sky imagers (TSIs), which is most prevalent in literature.^{7,18} However, Gueymard *et al.*¹⁰ proposed that TSIs do not provide an ideal ground truth due to several systematic and operational limitations, including the scarcity of collocated TSIs at the correct temporal resolution.

Due to the capricious nature of transitional cloud cover and relatively short lifetime of some cloud forms, higher temporal sky image resolutions are required than provided by most satellite products.¹⁹ However, despite the higher resolution of ground-based imagery, it often has limited spatial coverage.

D. The need for simplified classification methods

Sky condition classification and clear condition identification have been overcomplicated in literature, often requiring several atmospheric features and additional radiation components, such as the methodologies by Bright *et al.*¹¹ and Suárez-García *et al.*¹⁶

Although Bright *et al.*¹¹ proposed that more than one irradiance component is necessary to discern true sky conditions, it inherently limits the applicability of a classification method to data-scarce locations. Even more limiting is the utilization of theoretical clear sky irradiance in classification methodologies, which is often assumed available and accurate in literature.¹⁰ This is not always true, specifically in developing countries where radiation and atmospheric measurements are scarce and often not at the required spatiotemporal resolution.

In developing countries, a scarcity of long-term ground-based meteorological data is noted since 1980s.²⁰ This data scarcity is also increasing since new station development is limited and maintenance capacity deficits results in fewer permanent measurement locations.¹³ The causality of this data scarcity is a significant reduction in research output, which results in a scarcity of local and remote measurement validation, and an absence of literature prerequisite models, such as theoretical clear sky models. Considering this, many methodologies in

literature become significantly less applicable to these locations with a data scarcity.

It is therefore proposed that the representative sky conditions of a day or period can be classified solely from broadband global horizontal irradiance (GHI) measurements and the accessible *clearness index*. The proposed methods are, therefore, independent of solar and atmospheric models, and only require the most widely measured radiation component and calculable parameters. These methods have dual purposes: (1) to identify days or periods where the measured GHI is indistinguishable from GHI under a sky where clouds may or may not be visible but do not attenuate the GHI and (2) to classify not-clear days and periods into other descriptive classes.

II. METHODS

A. Fourier transforms and the GHI approximation

Considering the simple empirical clear sky formulation²¹

$$F_{G0} = 910 \cos \theta - 30, \quad (3)$$

where F_{G0} is the broadband clear sky GHI and θ is the solar zenith angle, it is inferred that a certain cosine or its complementary sinusoid function can reasonably approximate a day GHI time series. It is noted in Fig. 1 that half of the regular period of a sine wave may approximate the measured GHI between $00:00 \leq t \leq 23:59$. A Gaussian approximation is also considered, which is often considered a reasonable representative function for stationary time series due to its predictable Fourier transform.²²

Frequency domain analyses provide information regarding the underlying physical process of the time series, in this case, the character of solar radiation within a day or period. Assuming that the measured GHI series is periodic and reasonably continuous, it may be analyzed as in the spectral representation,²³

$$F(t) = \sum_n A_n \sin(2\pi f_n t + \varphi_n), \quad -\infty < t < \infty, \quad (4)$$

where t is time, A_n represents amplitude, f_n is the frequency, and φ_n is the phase of each harmonic function. The power within each harmonic is then equal to $|A_n|^2$. Equation (4) may be decomposed with Fourier analyses into a series of weighted sines and cosines that

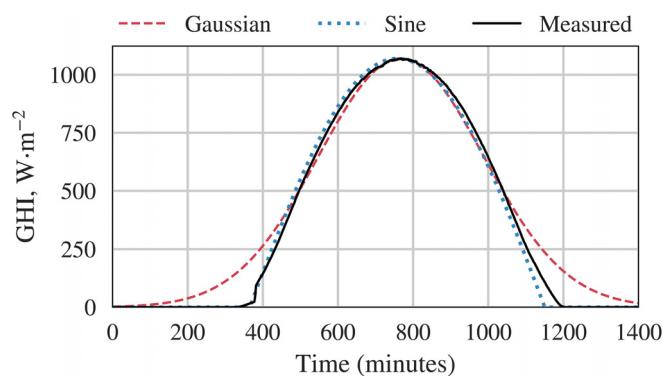


FIG. 1. The time domain for a measured *clear* day GHI, and Gaussian and sinusoid approximations.

contain a fundamental frequency f_0 and real integer multiples of the frequency (i.e., $2f_0$, $3f_0$, $4f_0$, etc.).²⁴

In practice, every point of the Fourier transformed (from time to frequency domain) function requires a discrete integral, and the function is limited to finite sums. This requires discretization of the continuous series into a set of samples that is converted back into a continuous series, which is only applicable if the sampling frequency ($f_s = 1/\Delta t$) is greater or equal to the Nyquist frequency ($f_c = 1/2\Delta t$), where Δt is the sampling frequency.

The predictable features of the Gaussian approximation anticipate that (1) the GHI series is decidedly localized in the time domain, (2) the maximum power in the frequency domain is at 0 Hz, and (3) the frequency series decays exponentially. As a result, a time series function is first delocalized through a first derivative of the specific function, which removes the unchanging component of the series. The resulting power spectral density (PSD) plots of the delocalized functions are shown in Fig. 2.

A first difference filter rather than a true derivative is applied since derivation favors narrow time variability, resulting in a noisy real-time signal and poorer signal-to-noise ratio than the original series.

B. The classification of an entire day of GHI into classes clear, intermittent, and cloudy

The frequency theory suggests that a frequency domain function is characterized by (1) the overall function shape and (2) the variability within the function. Considering the frequency response of the approximation functions in Fig. 2, it is assumed that the amplitude at the fundamental frequency suggests the shape of the function, while the rate of decay indicates the smoothness. Two theorems are therefore proposed to classify sky conditions into classes *clear*, *intermittent*, and *cloudy* as follows:

Theorem 1: If $\max(A_n) = |F(f_0)|^2$ or $\max(A_n) = |F(f_1)|^2$ for $f(t), 00:00 \leq t \leq 23:59$, the day is *clear* or *intermittent*.

Theorem 2: Let k_c be the number of elements of $\frac{A_n}{\max(A_n)} \times 100\%$ that are greater or equal to 1.00%, with A_n the amplitude at each frequency within $|F(f)|^2$ and k_c the decay number; if the day is not *cloudy* and $k_c \leq 10$, the day is *clear*.

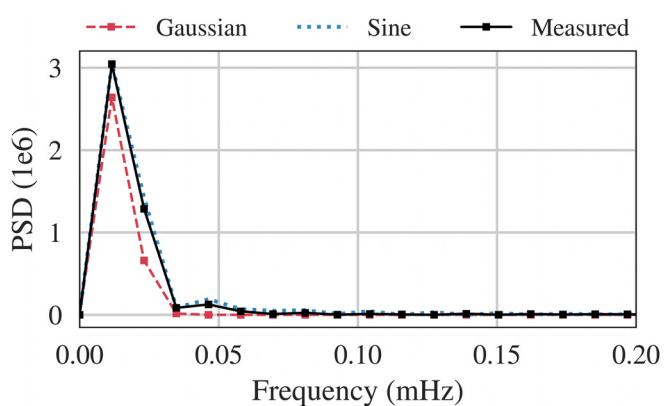


FIG. 2. The resulting delocalized measured GHI, and Gaussian and sine approximations in the frequency domain.

The first theorem is resultant from the approximation function analyses, where the maximum power amplitude ($\max(|F(f)|^2)$) at f_0 suggests a dominant bell-curve shape in $f(t)$ (the time series function). Application of this theorem to measured GHI time series indicates that the maximum amplitude coincides at f_0 in summer months and some spring and autumn days. The maximum amplitude coincides at f_1 for *clear* and *intermittent* predominantly winter days, while considerable power is still located at f_0 . This power shift is attributed to the decreased amplitude and narrowed bell-curve in the time domain, which translates into a compressed delocalized difference curve with a decreased maximum and average power in the frequency domain.

Despite the seemingly exponential decay of the approximation functions in the frequency domain, the measured GHI series response includes several peaks outside of the exponential decay function; the decay is also not continuous. As a result, a new decay number is formulated in **Theorem 2** to describe the variability observed in the PSD curve.

The proposed algorithm to classify the representative sky conditions of an entire day from 1-min resolution GHI measurements is presented in [Appendix A, Fig. 24](#).

Classification of sky conditions is subjective, no matter the classification method. The advantage of the decay number k_c is that it provides a clearness scale: a smaller k_c indicates a clearer day. As a result, varying k_c allows variation of conservatism as required. The *intermittent* class further provides a fail-safe to this conservatism by grouping days that may be varyingly classified by other methods.

C. The classification of a 15-min period of GHI into five classes

Although the proposed method has the capability of classifying the representative sky conditions of an entire day, it is often too conservative to be applicable for all locations. The proposed method is therefore protracted to rather classify 15-min periods of broadband GHI, without significantly increasing the complexity or dependence on measured parameters.

The presumption of characterization by shape and smoothness in the day classification holds for a period classification, however, the frequency function shape cannot inform the period classification without assuming the functions of other periods within the same day. In analogy, however, **Theorem 1** is a *measure of amplitude*, while **Theorem 2** is a *measure of variability*.

Since the *clearness index* (k_t) in (1) is a measure of the expected shape of GHI under clear sky conditions, it may provide the *measure of amplitude* of a period independently while only requiring the widely calculated extraterrestrial irradiance and solar zenith angle. However, due to the inherent location dependence of k_t , it is only used to identify *overcast* periods. Another limitation of incorporating k_t is the limited applicability to periods where $\theta \geq 85^\circ$. As observed in [Fig. 3](#), the first and last (one to two) periods can therefore not be classified without assimilating the uncertainty of k_t .

Although the smoothness or variability of a curve may be described by various methods such as a non-linear *curvature* function,²⁵ the analogues second derivative and standard deviation (SD) thereof are used to simplify the *measure of variability*. However, this is not applied to previously classified *overcast* periods due to the often (constant) smoothness of GHI under these conditions, which is similar to the variability under clear conditions.

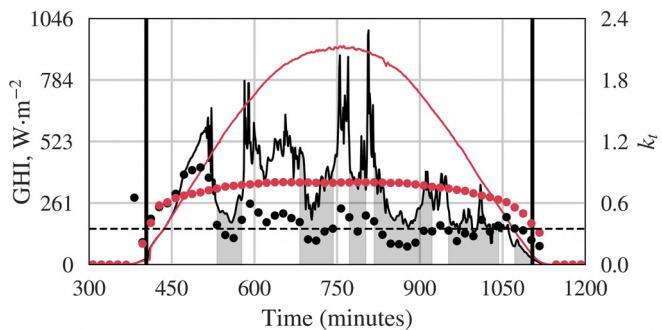


FIG. 3. Measure of amplitude: Measured GHI time series of two specific days with the calculated 15-min period clearness index (k_t) designated by scatter plots, the overcast threshold indicated by a horizontal dash line, and the identified overcast periods shaded; the classification is only applicable between the two vertical lines as a result of the k_t zenith angle limitation.

The second derivative in the period classification is

$$y'' = \frac{d}{dx} \left(\frac{dy}{dx} \right). \quad (5)$$

In other words, the second derivative is the first difference of dy/dx , divided by the first difference of x . The SD is the *sample* standard deviation

$$SD = \sqrt{\frac{\sum_{i=1}^N (y''_i - \bar{y}'')^2}{N-1}}, \quad (6)$$

where \bar{y}'' is the mean value of the second derivative and N is the number of sample observations. Due to the large SD values calculated, the SD thresholds in [Fig. 25](#) are scaled by 10^{-6} .

The *measure of variability* is illustrated in [Fig. 4](#), and it is observed that a relationship between SD and cloudiness exists: a lower SD indicates a smoother GHI line. The class thresholds indicated by dashed lines in [Fig. 4](#) are, however, subjective and may be varied with required conservatism. If the aim of the classification is clear sky detection, the lowest SDs may be retrieved to identify the “clearest” periods.

Although the objective of period classifications is often near real-time temporal resolutions, variability can only be characterized by

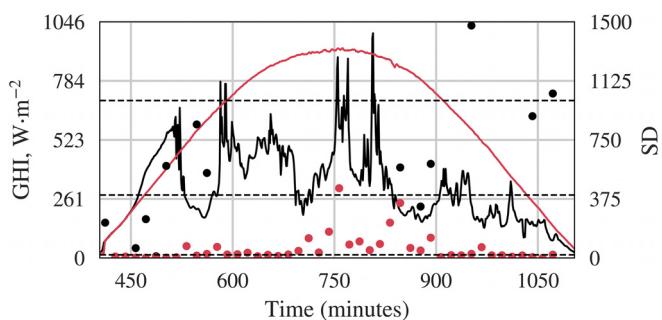


FIG. 4. Measure of variability: Measured GHI time series of two specific days with the calculated 15-min period standard deviation (SD) of the second derivative designated by scatter plots for $\theta \leq 85^\circ$; horizontal dash lines indicate class thresholds.

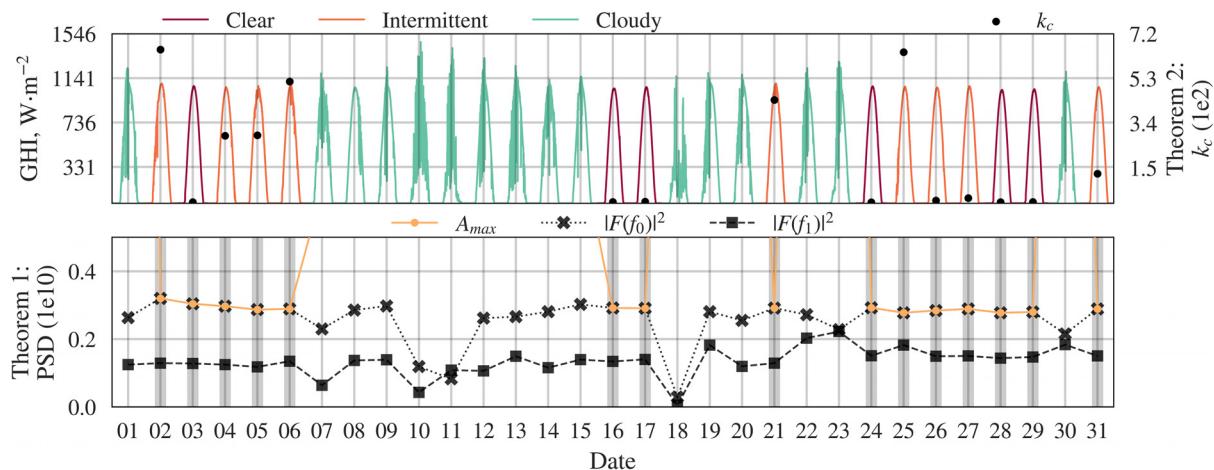


FIG. 5. The measured GHI for January 2019 classified per day into classes *clear*, *intermittent*, and *cloudy*, according to the PSD as in **Theorem 1** and k_c in **Theorem 2**. The GHI is measured at Stellenbosch, South Africa by SAURAN.²⁶ k_c is only calculated (and therefore displayed) for classes *clear* and *intermittent*.

considering an adequate number of sample periods (measurements); a 15-min period was found adequate to describe the variability of a 1-min resolution GHI time series. The proposed classification may therefore be considered a *representative* 15-min period classification since the sky conditions at each minute within the period is collectively considered within the period classification.

It is observed in the abridged flow chart in [Appendix A Fig. 25](#) that the addition of the *clearness index* does not introduce significant complexity to the *measure of amplitude*, while the *measure of variability* is still based on the fundamentals of the day classification. Although five classes, namely, *clear*, *partially clear*, *partially cloudy*, *cloudy*, and *overcast* are introduced, the classification of the first four classes only depend on the SD of the second derivative. As a result, the number of classes and the SD thresholds characterizing each class may be adapted to the required conservatism.

III. RESULTS

A. Entire day classification results

The entire day classification results for January are shown in [Fig. 5](#)—since this is a summer month, the *clear* and *intermittent* classified days have predominantly $\max(A_n) = |F(f_0)|^2$ in **Theorem 1**. k_c is only calculated for these two classes, with $k_c > 10$ classified as *intermittent* in **Theorem 2**. The value of $\max(A_n)$ is of no consequence to the classification algorithm; however, it is observed in [Fig. 5](#) that days with increased variability have higher $\max(A_n)$ values.

It is observed in [Fig. 6](#) that although September is a spring month, some days within the month rather satisfies $\max(A_n) = |F(f_1)|^2$ in **Theorem 1**; however, both theorems still hold for this and other months. Furthermore, it is observed that $|F(f_0)|^2$ and $|F(f_1)|^2$ are closer in value in September than January.

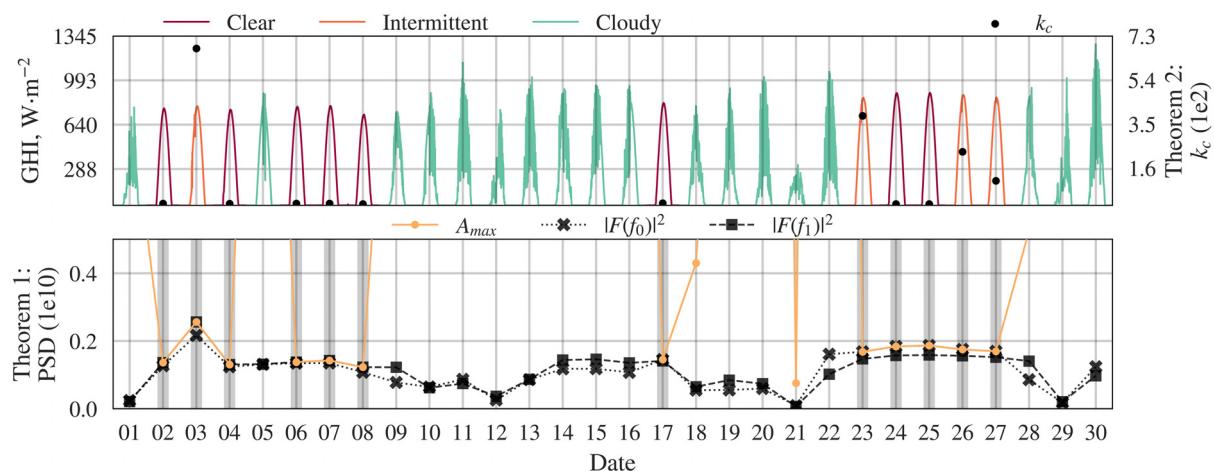


FIG. 6. The measured GHI for September 2019 classified per day into classes *clear*, *intermittent*, and *cloudy*, according to the PSD as in **Theorem 1** and k_c in **Theorem 2**. The GHI is measured at Stellenbosch, South Africa by SAURAN.²⁶

TABLE II. The purposively selected SAURAN (Southern African Universities Radiometric Network)²⁶ and BSRN (Baseline Surface Radiation Network)²⁷ stations for validation of the period classification, with latitude and longitude in degrees, and altitude in meters.

| Climate | Name | Country | Latitude | Longitude | Altitude |
|-------------|---------------------------|-------------------|----------|-----------|----------|
| Continental | Sioux Falls | South Dakota, USA | 43.73 | -96.62 | 473 |
| Tropical | Brasilia | Brazil | -15.60 | -47.71 | 1023 |
| Arid | Tamanrasset | Algeria | 22.79 | 5.53 | 1385 |
| Polar | Concordia Station | Antarctica | -75.10 | 123.38 | 3233 |
| Temperate | Stellenbosch ^a | South Africa | -33.93 | 18.87 | 119 |

^aA SAURAN station, while all other are BSRN stations.

B. Preliminary period classification results

The proposed Fourier transform period classification method is preliminarily validated for locations with varying climates, as listed in **Table II**. Due to the limitations of classification validations as discussed in Sec. IC, a qualitative validation is predominantly pursued through the comparison of the proposed method and *Perez sky clearness index* k_e ¹⁵ classification results.

Although the day classification may also be applied to any of these locations, the conservatism of such an entire day classification may be too limiting for locations that observe few *clear* days.

The locations in **Table II** are chosen based upon climatic variability as to include both temperate and extreme conditions in the validation process. The classification results for specific days exhibited in this section are purposively selected to represent all conditions and seasons.

While the thresholds in the proposed method and the *Perez index* may be adjusted based on location, these thresholds are kept constant across all locations and seasons within the validation in this section. This will elucidate the sensitivity and flexibility, and therefore indicate the possible location and seasonal independence of the method. All subsequent classification results of the proposed method are based on the algorithm and thresholds in Fig. 25, Appendix A.

1. Continental: Sioux Falls

Sioux Falls has a continental climate, characterized by significant variation in annual temperatures but little seasonal variation in

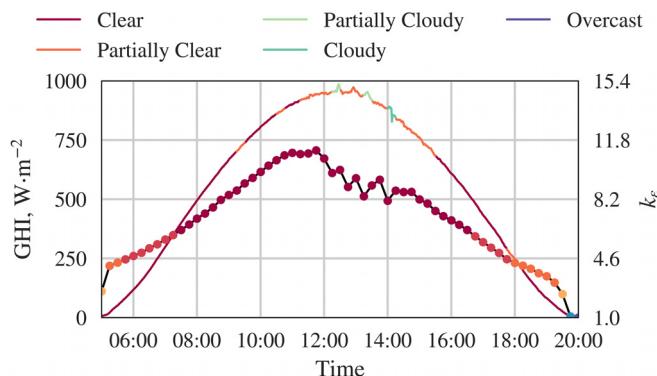


FIG. 7. 26 May 2018, Sioux Falls, South Dakota, USA,^{27,28} the proposed 15-min period classification results with legend as in the figure, and the *Perez sky clearness index* k_e designated by the scatterplot, with legend as in **Table I**.

precipitation. Due to the relatively smaller zenith angles toward the summer season, almost all periods between sunrise and sunset are classified despite the limit on applicable zenith angles, as observed in Fig. 7. The proposed period method classifies periods of turbidity around solar noon as *partially clear*, while more significant peaks are classified as *partially cloudy*. A single trough is classified as *cloudy*. Although these variations are observed in the *Perez index* scatterplot in Fig. 7, the thresholds as is classifies all of these periods as *clear* and therefore cannot distinguish between periods of relatively small variation. Since a higher k_e value indicates clearer conditions, it is slightly more complex to change the class conservatism in the *Perez index* classification.

It is also observed in Fig. 7 that the *Perez index* is more partial to the *measure of amplitude*, where periods closer to solar noon are classified toward the *clear* end of the classification spectrum.

The GHI series in Fig. 8 is purposively selected to illustrate the limitations of the *clearness index* (k_e) adoption and the proposed method. It is observed that the proposed method misidentifies two extended periods of *clear*, which start at 08:30 and 13:00. Considering that both periods are relative smooth, it is assumed that k_b as proposed in Fig. 25, Appendix A, cannot distinguish between *clear* and *overcast* under these conditions, which is its only function.

Considering the conditions in Fig. 8 and the formulation of k_e in (1), it is noted that although the mean GHI is relatively low on this day, the solar zenith angle is relatively large. In fact, the minimum solar zenith angle on this day is $\sim 66^\circ$, which is approximately 10° larger than at more temperate locations (such as Stellenbosch), under

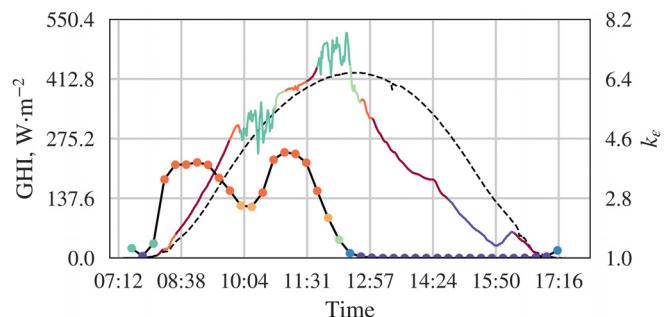


FIG. 8. 2 January 2018, Sioux Falls, South Dakota, USA,^{27,29} the proposed 15-min period classification and *Perez sky clearness index* k_e results with legends as described in Fig. 7; a subsequent relatively *clear* day GHI is indicated by a dash line.

the same conditions. In other words, irradiance under *overcast* conditions still produces a relatively large k_t due to the large zenith angle at this location. As a result, the adoption of k_t to distinguish between *clear* and constant *overcast* periods under winter continental climates is often less successful than in summer months and more temperate climates.

However, due to the relative simplicity of the proposed method, several adaptations are possible to overcome this limitation. A condition of $0 < k_t \leq 0.7$, similar to what is observed in literature^{13,14} in Sec. IB, may be more appropriate to ensure the *measure of amplitude* can identify *overcast* periods. Perhaps more appropriately, the k_t threshold may be adapted to be season and/or solar zenith angle specific. Furthermore, based on availability, an appropriate clear sky model or additional irradiance components may be applied to distinguish between these two conditions.

Another limitation of the proposed method is the representative nature thereof. In Fig. 8, it is observed that the 15-min period at 11:30 is seemingly misclassified as *clear*. The period indicates little oscillation, and the first few minutes of the period seem *clear*; however, due to the representative 15-min period considered, the method cannot distinguish within the period between specific classes. However, increasing the proposed method resolution may resolve this limitation of sub-period classification, as discussed in Sec. IV B.

2. Tropical: Brasilia

The Brasilia station is located within a tropical savanna climate, which is characterized by an extensive dry winter season and wetter humid summer months. In tropical moist climates such as Brasilia, greater surface heating and higher humidity are expected, resulting in fewer *clear* days.

Conversely to the *clearness index* adaption required at Sioux Falls, the proposed threshold of $0 < k_t \leq 0.35$ is appropriate for identifying *overcast* winter conditions at Brasilia, as observed in Fig. 9. It is observed that the proposed method and *Perez index* are in good agreement on the *overcast* periods; however, the zenith bias observed in the *Perez index* in Fig. 7 is also observed in Fig. 9, where the index increases with decreasing zenith angle despite some variability within the GHI curve.

At Brasilia, it is observed that most days through the year is characterized by relatively *clear* mornings, while *cloudy* to *overcast* periods increase toward afternoon, as observed in Figs. 9 and 10.

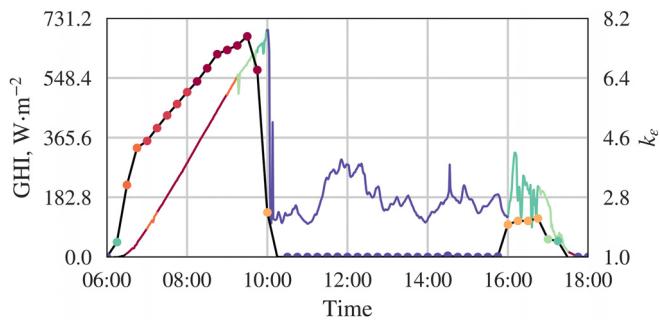


FIG. 9. 17 May 2018, Brasilia, Brazil.^{27,30} the proposed 15-min period classification and Perez sky clearness index k_e results with legends as described in Fig. 7.

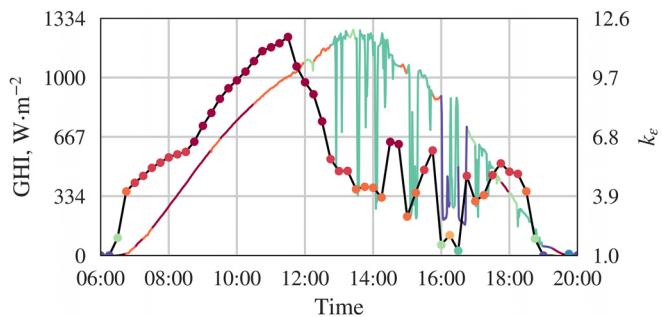


FIG. 10. 21 December 2018, Brasilia, Brazil.^{27,31} the proposed 15-min period classification and Perez sky clearness index k_e results with legends as described in Fig. 7.

It is observed in Fig. 10 that all five proposed classes are identified on this day; considering this figure, the prevalent defining GHI character of each sky condition class can be summarized as follows:

- *Clear*: Relatively smooth and primarily defined by the shape, which is similar to the sine approximation curve in Fig. 1, or the theoretical *clear sky* GHI.
- *Partially clear*: Defines turbid or hazy periods, which do not necessarily indicate cloud formation but do include significant atmospheric “noise.” The shape of theoretical *clear sky* GHI is still observed.
- *Partially cloudy*: Increased variability within the GHI curve, with distinct peaks and troughs. However, the shape of theoretical *clear sky* GHI is still observed.
- *Cloudy*: Distinct peaks and troughs that occur frequently enough to obscure the shape of the GHI curve.
- *Overcast*: Often characterized by a smooth curve, not unlike *clear* conditions; however, the periods have distinctly lower GHI than expected.

3. Arid: Tamanrasset

Tamanrasset is arid, with dry conditions persisting throughout the winter season and very high temperatures in the summer season. A considerable increase in water vapor is marked in summer seasons, resulting in increased humidity advection closer to the surface. During these periods of moisture advection, greater cloud coverage percentages are noted. Furthermore, due to Tamanrasset’s location within the Saharan Desert, it is characterized by considerable dust transport.³² Bright *et al.*¹¹ disclosed that many clear sky detection methods perform unsatisfactorily under these conditions at Tamanrasset. It has, as a result, become a notable location for solar characterizing in literature.^{33,34}

It is observed in Fig. 11 that the proposed method is significantly more sensitive to the small variations expected under considerable dust events.

This sensitivity of the proposed method to turbid or hazy periods is also observed in the winter seasons, as in Fig. 12. The proposed method also identifies relatively short periods of *partially cloudy* to *overcast* conditions under low GHI conditions, such as late afternoon,

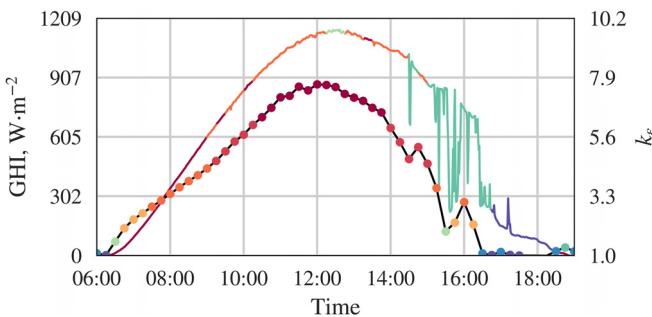


FIG. 11. 28 August 2018, Tamanrasset, Algeria.^{27,35} the proposed 15-min period classification and Perez sky clearness index k_e results with legends as described in Fig. 7.

while the *Perez index* predominantly classifies the periods in Fig. 12 as *clear*.

4. Polar: Concordia Station

The polar climate of Antarctica is predominantly characterized by winter-type conditions throughout the year. Furthermore, days within the winter seasons are relatively short with measured GHI often of $\sim 0 \text{ W m}^{-2}$ for the entire day, whereas summer days are long and measured GHI often of $\neq 0 \text{ W m}^{-2}$ even throughout evening hours. Due to the exclusion of GHI periods with $\theta \geq 85^\circ$, an increasing number of GHI periods are not classified by the proposed method due to predominantly larger zenith angles toward the winter season. It is observed in Fig. 13 that the proposed method and *Perez index* predominantly agree on the classifications, with the *Perez index* slightly more biased toward more *clear* periods.

Due to the relatively smaller zenith angles ($\theta \leq 85^\circ$) throughout the summer season, the proposed method classifies all periods in $0:00 \leq t \leq 23:59$ and the global irradiance component does not decrease to 0 W m^{-2} during the evening, as observed in Fig. 14. The *Perez index* only classifies between 02:00 and 19:45 in Fig. 14 due to missing direct measurements, rather than irradiance that decrease to 0 W m^{-2} .

It is again observed in Fig. 14 that the *Perez index* is more partial to the *measure of amplitude*, where periods closer to solar noon are classified toward the *clear* end of the classification spectrum, despite similar variability to other periods within the day.

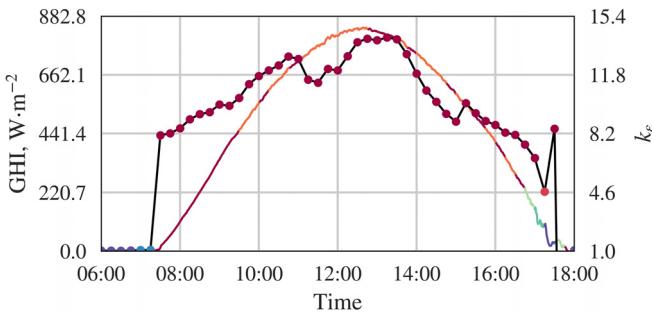


FIG. 12. 31 December 2018, Tamanrasset, Algeria.^{27,36} the proposed 15-min period classification and Perez sky clearness index k_e results with legends as described in Fig. 7.

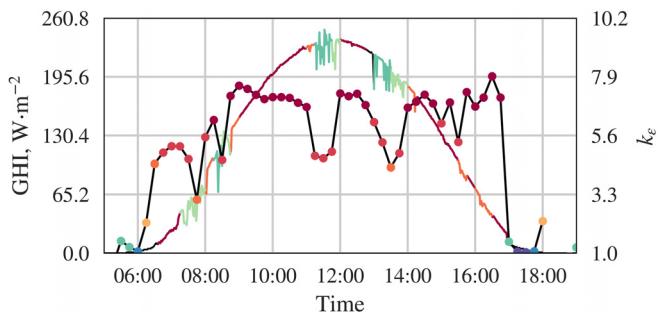


FIG. 13. 28 March 2018, Concordia Station, Antarctica.^{27,37} the proposed 15-min period classification and Perez sky clearness index k_e results with legends as described in Fig. 7.

5. Temperate: Stellenbosch

In Mediterranean temperate climates, such as Stellenbosch, dry conditions persist through the warm summer season and rainfall in the mild winter season.

It is observed in Fig. 15 that despite relatively low GHI under winter conditions, the proposed method identifies the seemingly *overcast* periods at 08:30 and 17:00. Although both the proposed method and the *Perez index* identify several *clear* periods in the morning, the *Perez index* identifies a period of distinct variability as *clear* at 12:45, indicating the relative insensitivity of this index.

Due to the temperate climate of this location, the classification results are similar to the representative findings of the previous locations.

C. Supplementary examples of temperate period detection and classification results

1. Preliminary clear detection results

If the proposed resource classification algorithm is applied with the intent of identifying *clear* periods, it is of interest how it performs in comparison to the *Reno and Hansen* method,⁷ as introduced in Sec. IB. Due to the limitations of the *Reno and Hansen* method,¹¹ it is comparatively applied only under temperate conditions.

For simplicity, the `detect_clearsky` and `get_clearsky` PVLIB functions³⁹ are used with the *Ineichen* and *Perez* model,⁴⁰

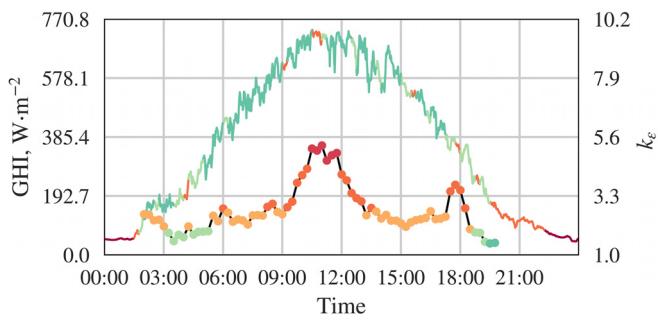


FIG. 14. 17 November 2018, Concordia Station, Antarctica.^{27,38} the proposed 15-min period classification and Perez sky clearness index k_e results with legends as described in Fig. 7.

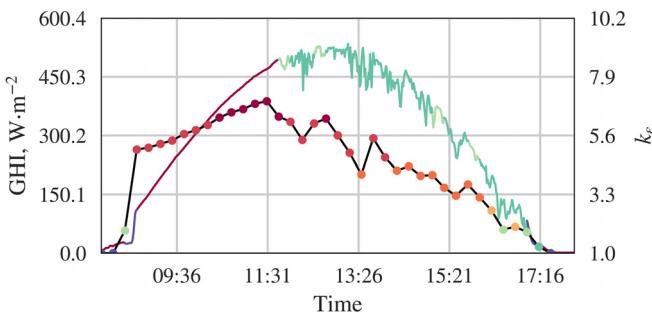


FIG. 15. 18 June 2019, Stellenbosch, South Africa:²⁶ the proposed 15-min period classification and Perez sky clearness index k_e results with legends as described in Fig. 7.

which is reported to perform well with limited input data.⁷ The *clear* sky detection results of the proposed 15-min period and *Reno and Hansen* method⁷ for a limited number of days in April and December in Stellenbosch are presented in Figs. 16 and 17.

The *Reno and Hansen* method⁷ identifies 1-min periods of measured GHI, while the proposed method identifies the representative conditions over a 15-min period. As result, the proposed method identifies fewer clear periods. However, considering for example frames 3 and 8 in Fig. 16 and frame 6 in Fig. 17, it is surmised that the proposed method is also more conservative, which results in the detection of fewer *clear* periods.

This type of conservatism is often welcomed in “cloudless” detection methodologies since it provides a relative guarantee that the GHI is measured under *clear* conditions,¹⁰ as often required in the process of clear sky model conditioning and validation.

The detection of *clear* periods may be considered a binary classification. A confusion matrix is often adopted to assess the binary classification of *clear* conditions,^{10,16,33} as in Table III. However, due to a baseline deficit, a confusion matrix is rather used here to assess the similarity between the proposed and *Reno and Hansen* methods. In other words, the proposed *clear* detection results are considered as the “predicted” series, while the *Reno and Hansen* results are considered the “observed.” As a result, TP indicates that both the proposed and *Reno and Hansen* methods classify a period as *clear*, whereas FP indicates that the proposed method identifies a period as *not clear* that the *Reno and Hansen* method identifies as *clear*, and so forth.

The proposed method results are filled forward to meet the 1-min resolution of the *Reno and Hansen* method for the entire 2019 series at Stellenbosch, South Africa. The confusion results are limited

to periods where $\theta \leq 85^\circ$ to meet both the criteria of the proposed method and limit the number of TNs.

Due to the interpretation of the confusion matrix in this context of comparison, TP and TN indicate that the proposed and *Reno and Hansen* methods similarly classify periods as either *clear* or *not clear*, while FP and FN indicate dissimilarity. In other words, the confusion matrix and the results thereof do not necessarily indicate the accuracy of either method, but rather they indicate how similar or dissimilar the two methods classify periods as *clear* or *not clear*.

As a result, *accuracy*, as in (B1) provided in Appendix B, is a measure of similarity in the detection of both *clear* and *not clear* periods, while *precision* in (B2) is a measure of *clear* detection similarity of the two methods.

The *accuracy* in Table IV indicates that the two methods similarly classify ~70%–95% of the periods, despite the resolution disparity. The *precision* results indicate that more than 93% of the periods are similarly identified as *clear*, while less than 7% of the periods are identified as *clear* by the proposed method but as *not clear* by the *Reno and Hansen* method in terms of the positives (*clear*). The *recall* (B3), and by inference the *F-score* (B4), is lower than the other statistics due to consideration of FNs. In other words, the proposed method classifies a considerable number of periods as *not clear* that is classified as *clear* by the *Reno and Hansen* method. As a result, these two statistics are indicative of conservatism in the proposed method.

It is also observed in Table IV that the two methods seem to be more similar in winter (June to August) than summer months, which may be ascribed to an increased number of TNs (both methods identify a period as *not clear*) in winter.

The percentage of TP, TN, FP, and FN in each month is illustrated in Fig. 18—it is observed that the highest percentage of TNs is found throughout the year, while the FPs are the lowest. This indicates that the two methods agree on a high percentage of *not clear* periods, while very few periods are classified by the proposed method as *clear* that is classified as *not clear* by the *Reno and Hansen* method. It is noted that while higher FPs occur in winter months, higher FNs occur in summer months.

The relatively large percentage of FNs (vs FPs) in Fig. 18 is indicative of greater conservatism of the proposed method, as previously mentioned. However, considering the lower percentage of FNs in June to August, it can also be said that the two methods are similarly conservative in winter months.

2. Period classification results with sky images

An independent radiometric station is in development in close proximity to the SAURAN (Southern African Universities

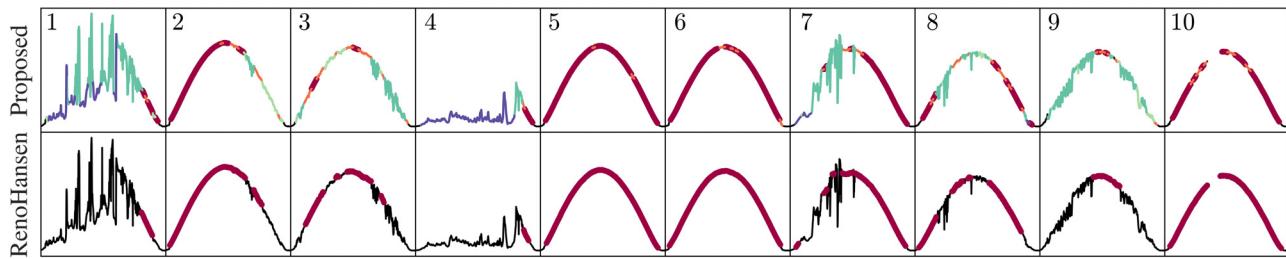


FIG. 16. April 2019, Stellenbosch, South Africa:²⁶ the proposed 15-min period and *Reno and Hansen*⁷ clear detection results, with the legend as described in Fig. 7.

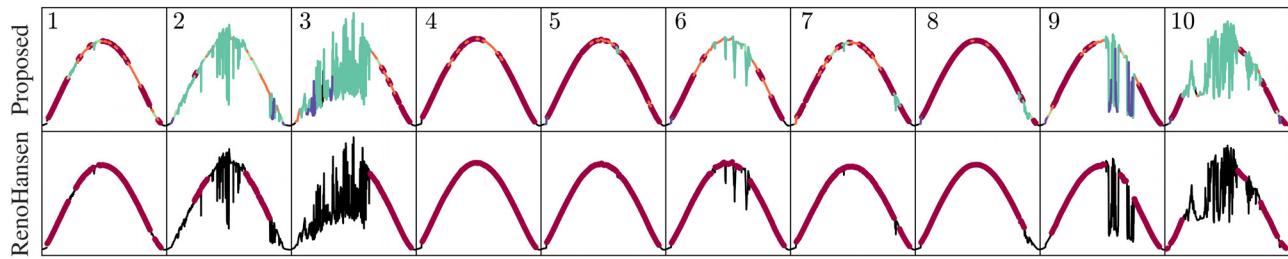


FIG. 17. December 2019, Stellenbosch, South Africa:²⁶ the proposed 15-min period and *Reno and Hansen*⁷ clear detection results, with the legend as described in Fig. 7.

Radiometric Network) radiometric station at Stellenbosch, South Africa. This additional radiometric station includes high-resolution sky photographs at 10-min intervals, which provides an opportunity to preliminarily verify the proposed period classification results with observed conditions at this temperate location. The supplementary examples further provide clarification on the proposed classification classes and the expected observed conditions within the specific classifications, beyond only the identification of *clear* conditions.

The first example in Fig. 19 illustrates the subjectivity of the *partially clear* class. The difference in curve smoothness between periods classified as *clear* or *partially clear* in this figure is almost indiscernible. Furthermore, the sky images appear to indicate predominantly *clear* conditions; however, as discussed in Sec. III B 2, the *partially clear* class does not necessarily indicate cloud formation but rather indicates the presence of atmospheric “noise” which is not always discernible in sky images.

Although the sky images in Fig. 19 illustrates a relatively narrow field-of-view, the camera is mounted on a double axis tracker, allowing several sky image angles. Currently, a single horizontal, direct, direct-tilted, and north-tilted view is captured within ~ 2 min; these four views are deemed representative of the sky dome and therefore considered collectively here. However, only a single view is shown in Figs. 19–23 due to space limitations.

It is then considered that the *partially clear* class pertains more to conservative classifications, as applicable to clear sky detection methods. However, as discussed in Sec. II C, the *measure of variability* is subjective and the SD thresholds are flexible to allow application required conservatism.

The *partially clear* class may also serve as a grouping of periods that represent a state between *clear* and *cloudy*, similar to the *intermittent* class in the proposed day classification algorithm, where it was found that days classified within this class is often varying classified by other methods.

TABLE III. Confusion matrix, with TP true positive, FP false positive, FN false negative, and TN true negative.

| | Predicted | |
|------------------|--------------|------------------|
| Observed | <i>Clear</i> | <i>Not clear</i> |
| <i>Clear</i> | TP | FN |
| <i>Not clear</i> | FP | TN |

Conversely, in Fig. 20, it is observed that the transition from *clear* to *partially clear* at 10:30 is due to the advent of minor cloud formation over the 15-min period (10:30 to 10:45) classified as *partially clear*. The transition at 10:45 to *cloudy* is due to the development of cloud formation closer to the solar disk in the last three 15-min periods in Fig. 20.

The last period in Fig. 20 transitions to *clear* in Fig. 21, according to the proposed method, although a minor cloud formation is observed in the sky image at 11:25. This again illustrates the subjectivity of a classification, specifically when periods before and after a specific moment are not considered within the classification algorithm. The classified *clear* period does, however, indicate that the cloud formation within this period is insignificant to the degree that it does not change the amount of irradiance observed.

Although the cloud formation at 11:25 and 11:35 seems similar in Fig. 21, the classification transitions from *clear* to *partially cloudy*. This is due to cloud formation increasing within *partially cloudy* classified period (toward 11:45). Due to the residual cloud formations near the solar disk, the succeeding periods are classified as *cloudy*. The *partially cloudy* class seems to indicate transitional representative conditions, i.e., either an increase or decrease in cloud formation within the 15-min period.

TABLE IV. The similarity confusion matrix class statistics per month for 2019, Stellenbosch, South Africa; the year average for each statistic is also supplied.

| Month | Accuracy | Precision | Specificity | F-score | Recall |
|-----------|----------|-----------|-------------|---------|--------|
| January | 0.7145 | 0.9568 | 0.9708 | 0.6669 | 0.5118 |
| February | 0.8166 | 0.9772 | 0.9753 | 0.8225 | 0.7101 |
| March | 0.8602 | 0.9667 | 0.9872 | 0.7712 | 0.6414 |
| April | 0.8537 | 0.9727 | 0.9776 | 0.8397 | 0.7387 |
| May | 0.8998 | 0.9328 | 0.9759 | 0.8207 | 0.7327 |
| June | 0.9333 | 0.9827 | 0.9909 | 0.9056 | 0.8398 |
| July | 0.9433 | 0.9512 | 0.9776 | 0.9112 | 0.8745 |
| August | 0.9105 | 0.9717 | 0.9819 | 0.8874 | 0.8166 |
| September | 0.9001 | 0.9756 | 0.9847 | 0.8736 | 0.7909 |
| October | 0.9054 | 0.9750 | 0.9948 | 0.7630 | 0.6268 |
| November | 0.7710 | 0.9825 | 0.9883 | 0.7294 | 0.5800 |
| December | 0.8102 | 0.9930 | 0.9951 | 0.7787 | 0.6404 |
| Year | 0.8599 | 0.9698 | 0.9833 | 0.8142 | 0.7086 |

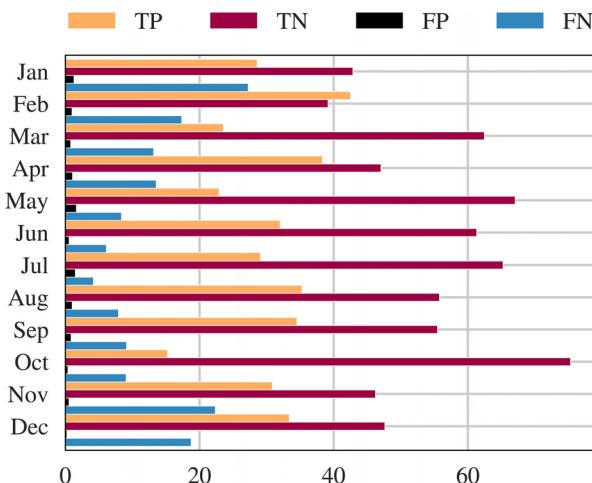


FIG. 18. 2019, Stellenbosch, South Africa:²⁶ the confusion matrix clear and not clear classification results (in percentage) for the proposed vs the Reno and Hansen method.

The transition from *partially cloudy* in Fig. 21 to *partially clear* in Fig. 22 is representative of the decrease in cloud formation within the proximity of the solar disk. Similar to the transition at 10:45 in Fig. 20, the transition from *partially clear* to *cloudy* indicates formation of clouds in close proximity of the solar disk.

It is further observed in Fig. 22 that the *overcast* class also represents conditions where the entire sky dome is not necessarily cloud covered; the class rather represents conditions where the solar disk is predominantly obscured. The transition from *overcast* to *cloudy* at 13:15 is due to reduced cloud formation around the solar disk.

It is observed in Fig. 23 that the class *cloudy* encapsulates conditions of varying cloud amount and position in relation to the solar disk. However, it is imperative to consider that while the classification is a representative characterization of conditions over a 15-min period, the sky images indicate instantaneous sky conditions. Notwithstanding, it may be noted from Fig. 23 that a relation between the instantaneous sky image and the SD of a 15-min period does exist, as theorized in Sec. II C.

Although a higher SD indicates greater variability within the period, it also alludes to an increase in cloud formation within the sky dome, and subsequently, greater probability of cloud formations within proximity and even obstruction of the solar disk.

IV. CONCLUSIONS AND FUTURE WORK

A. Conclusions

It is proposed that representative sky conditions can be classified primarily through the Fourier transform of broadband global horizontal irradiance time series. It is discussed in Sec. II C that characterization of the solar resource depends on two measures of the GHI time series curve, namely, the *measure of amplitude* and *measure of variability*, which applies to the classification of both an entire day series and an appropriate period within the series.

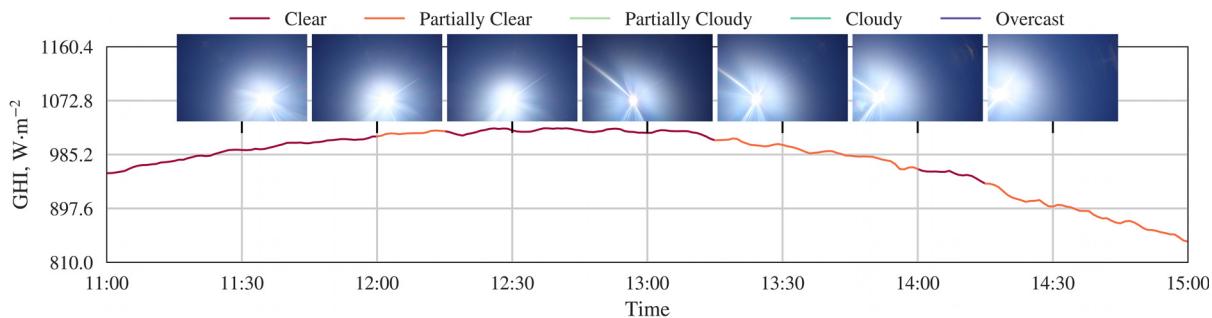


FIG. 19. 20 November 2020, Stellenbosch, South Africa:²⁶ the proposed 15-min period classification of a relatively clear day with horizontal sky photographs per 30 min between 11:00 and 15:00.

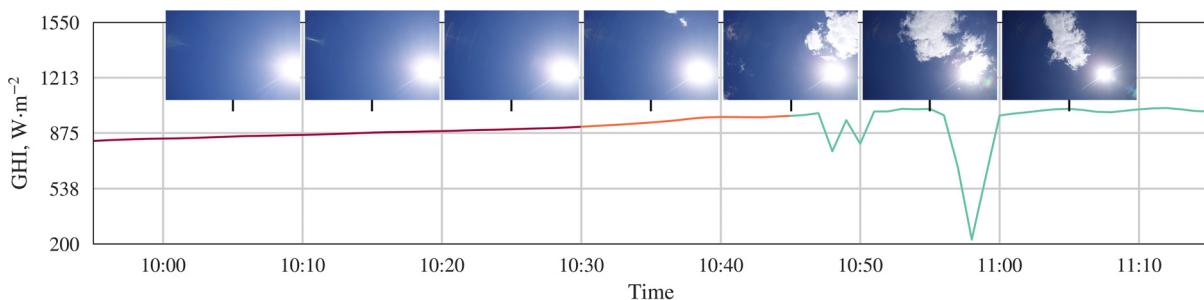


FIG. 20. 28 November 2020, Stellenbosch, South Africa:²⁶ the proposed 15-min period classification with horizontal sky photographs per 30 min between ~10:00 and 11:00, with a classification color index as in Fig. 19.

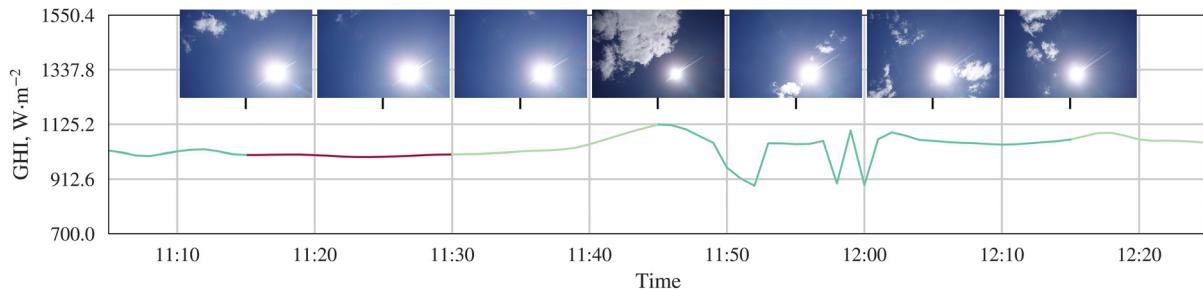


FIG. 21. 28 November 2020, Stellenbosch, South Africa.²⁶ the proposed 15-min period classification with horizontal sky photographs per 30 min between ~11:00 and 12:00, with a classification color index as in Fig. 19.

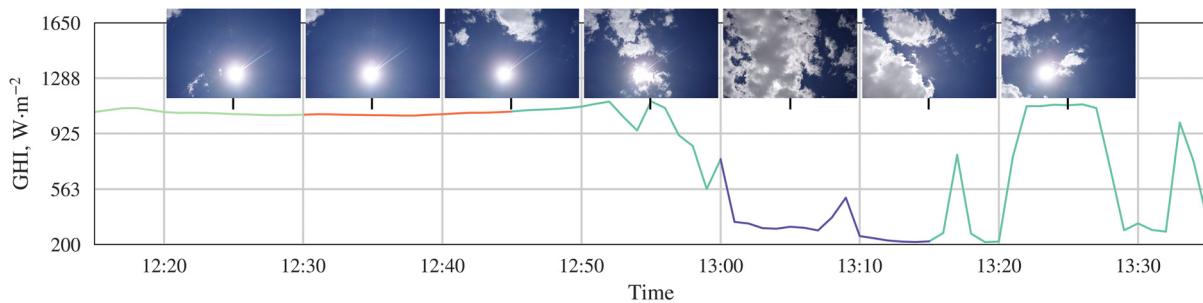


FIG. 22. 28 November 2020, Stellenbosch, South Africa.²⁶ the proposed 15-min period classification with horizontal sky photographs per 30 min between ~12:00 and 13:00, with a classification color index as in Fig. 19.

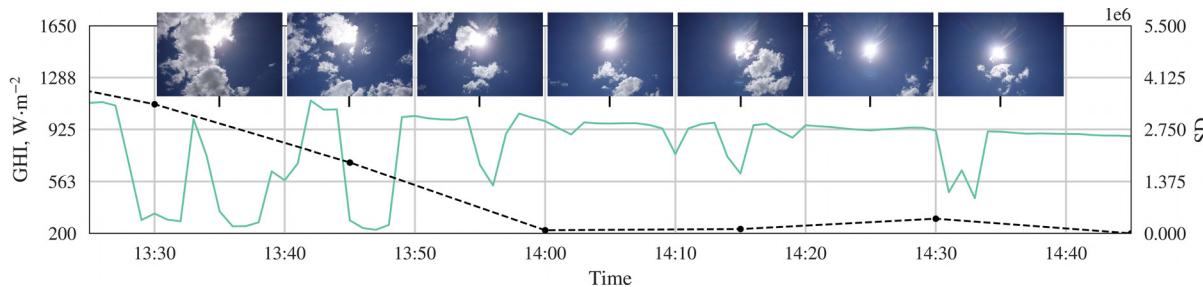


FIG. 23. 28 November 2020, Stellenbosch, South Africa.²⁶ the proposed 15-min period classification with direct-tilted sky photographs per 30 min between ~13:00 and 14:00, and the period standard deviation (SD) used for the measure of variability, with a classification color index as in Fig. 19.

The inductive classification method, as described in Sec. II B, supplies a rigorous classification of entire day conditions that fits for preliminary conditioning of clear sky models, specifically in more temperate locations. Despite the classification discrimination introduced by classifying entire days, the Fourier transform-based algorithm is a convenient approach to reduce the interdependence of the classification on location, modeled conditions, and other atmospheric parameters.

The day classification is as a result, augmented to rather classify 15-min period GHI time series into five classes, namely, *clear*, *partially clear*, *partially cloudy*, *cloudy*, and *overcast*, as described in Sec. II C. The resulting algorithm is, however, still concise and requires minimal inputs, as observed in Fig. 25.

It was shown in Sec. III B that the proposed period classification is relatively impartial to climatic conditions and that alteration of the thresholds would rather promote flexibility in classification conservatism. However, it was observed that the adoption of the *clearness index* results in limitations, specifically under winter continental conditions, i.e., low irradiance and high solar zenith angles. The model is adequately conservative to detect clear periods for the conditioning of clear sky models while granting flexibility for wider classification of the solar resource. It was shown in Sec. III C 1 that the proposed method is generally in good agreement with the *Reno and Hansen* method on the detection of clear conditions under temperate conditions; however, the proposed method is more conservative in the classification of *clear*.

The advantages of the proposed method are as follows: (1) it is clear sky model-independent, i.e., pre-modeled clear sky irradiance is not required, (2) additional atmospheric parameters and/or irradiance components are not required, (3) additional transitional classes, namely, *partially clear* and *partially cloudy* are included to identify periods of intermittent states, and (4) the proposed algorithm thresholds are applicable to varying climatic conditions, which may be conveniently altered to set the level of required classification conservatism.

Solar resource or sky condition classification is subjective; however, the proposed method is not more subjective than other classification methods found in literature. Although the consolidation of a classification method with sky images may reduce uncertainty, it does not necessarily decrease subjectivity due to the abstraction of the classes within the photovoltaic research space.

Subjectivity of the classification is further promoted by aggregation of various atmospheric states within the classified period, specifically so under rapidly changing sky conditions. While considering a more near real-time temporal resolution would reduce classification subjectivity, variability can only be characterized by considering an adequate number of applicable sample periods, limiting the currently proposed classification method.

B. Future work

Near real-time solar resource classification may be beneficial as modeling and forecasting resolutions increase. The resolution of the proposed classification method can be increased by considering a *moving measure of variability*, similar to a traditional statistical moving average, whereby samples around a specific GHI time-series point influences the classification of the specific point; the surrounding samples do not automatically assume the same classification as the specific point but are rather in turn influenced by their surrounding samples. This *moving measure of variability* can increase classification resolution to the resolution of the GHI time series.

Although the primary use of a GHI time series is advantageous for data-scarce locations, it is shown in literature that classification accuracy may increase with additional irradiance component consolidation. It may therefore be beneficial to similarly apply the *measure of amplitude* and *measure of variability* methods to a direct normal irradiance (DNI) time series and assimilate the results to the GHI classification to provide a less subjective classification. Analysis of DNI has the additional advantage of elucidating solar disk obscuration, without inferring the conditions of the entire sky dome.

ACKNOWLEDGMENTS

This work was supported by Eskom through EPPEI (Eskom Power Plant Engineering Institute) and TESP (Tertiary Education Support Programme).

APPENDIX A: CLASSIFICATION ALGORITHMS

The flow chart for the fundamental theory of the entire day classification is shown in Fig. 24.

The flow chart for the 15-min period classification of a 1-min resolution broadband GHI time series is shown in Fig. 25.

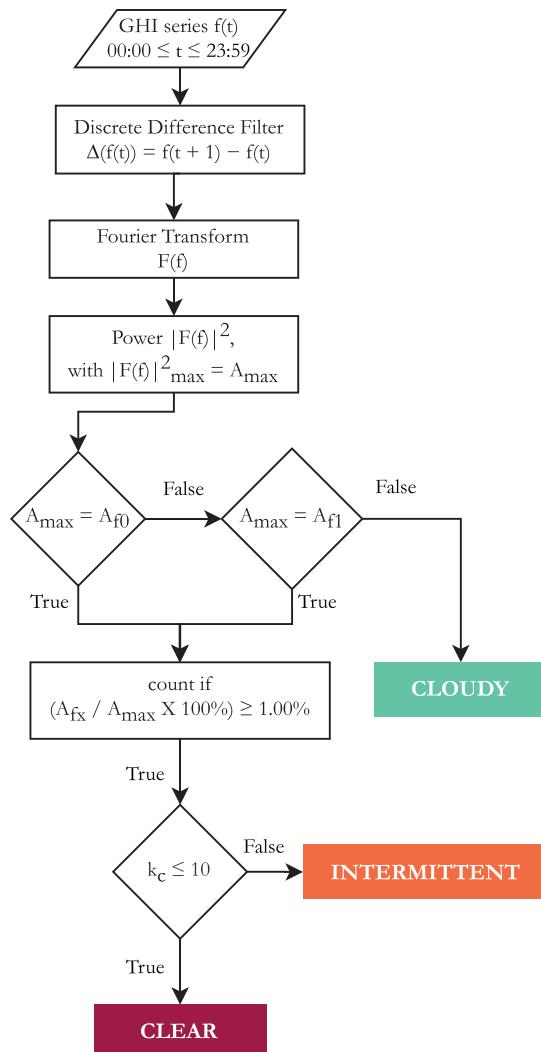


FIG. 24. Diagrammatic representation of the proposed algorithm classifying a single GHI day into classes *clear*, *intermittent*, and *cloudy*.

APPENDIX B: CONFUSION MATRIX CLASS STATISTICS

The *accuracy* or classification rate is

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (\text{B1})$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative, according to the confusion matrix in Table III. *Precision* is

$$\text{precision} = \frac{TP}{TP + FP}. \quad (\text{B2})$$

Recall,

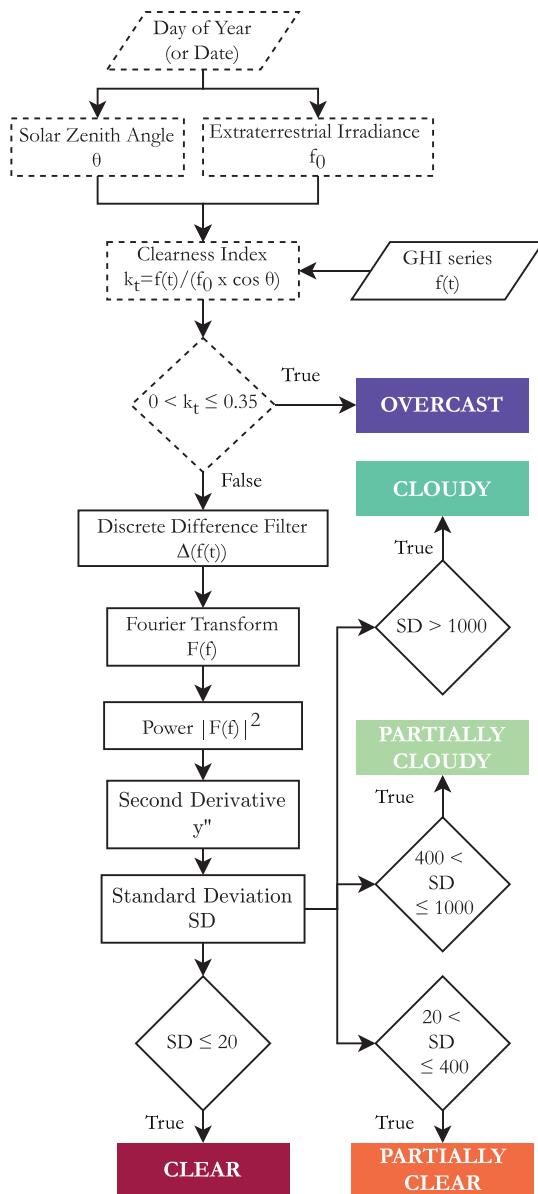


FIG. 25. Diagrammatic representation of the proposed algorithm classifying a single 15-min period of GHI into classes *clear*, *partially clear*, *partially cloudy*, *cloudy*, and *overcast*; the measure of amplitude is indicated by dashed squares.

$$recall = \frac{TP}{TP + FN} \quad (B3)$$

indicates how well a classification is recognized. *Recall* and *precision* may be combined in the *F-score* as follows:

$$F\text{-score} = \frac{2 \times recall \times precision}{recall + precision}. \quad (B4)$$

Specificity,

$$specificity = \frac{TN}{FP + TN} \quad (B5)$$

is a measure of the assignment to the positive class, which is the classification of *clear* in detection methods.

DATA AVAILABILITY

The data that support the findings of this study are openly available in SURAN at <https://sauran.ac.za/>, Ref. 26 and in BSRN at <https://dataportals.pangaea.de/bsrn/>, Refs. 27–31, and 35–38.

REFERENCES

- ¹P. Ineichen, “A broadband simplified version of the Solis clear sky model,” *Sol. Energy* **82**, 758–762 (2008).
- ²T. Muneer, S. Younes, and S. Munawwar, “Discourses on solar radiation modeling,” *Renewable Sustainable Energy Rev.* **11**, 551–602 (2007).
- ³J. Polo, C. Fernández-Peruchena, V. Salamatikis, L. Mazorra-Aguiar, M. Turpin, L. Martín-Pomares, A. Kazantzidis, P. Blanc, and J. Remund, “Benchmarking on improvement and site-adaptation techniques for modeled solar radiation datasets,” *Sol. Energy* **201**, 469–479 (2020).
- ⁴G. M. Lohmann, “Irradiance variability quantification and small-scale averaging in space and time: A short review,” *Atmosphere* **9**, 264 (2018).
- ⁵E. d. F. M. Júnior and R. Rüther, “The influence of the solar radiation database and the photovoltaic simulator on the sizing and economics of photovoltaic-diesel generators,” *Energy Convers. Manage.* **210**, 112737 (2020).
- ⁶A. Kharrazi, V. Sreram, and Y. Mishra, “Assessment techniques of the impact of grid-tied rooftop photovoltaic generation on the power quality of low voltage distribution network – A review,” *Renewable Sustainable Energy Rev.* **120**, 109643 (2020).
- ⁷M. J. Reno and C. W. Hansen, “Identification of periods of clear sky irradiance in time series of GHI measurements,” *Renewable Energy* **90**, 520–531 (2016).
- ⁸S. Younes and T. Muneer, “Clear-sky classification procedures and models using a world-wide data-base,” *Appl. Energy* **84**, 623–645 (2007).
- ⁹J. A. Ruiz-Arias, C. A. Gueymard, and T. Cebecauer, “Direct normal irradiance modeling: Evaluating the impact on accuracy of worldwide gridded aerosol databases,” *AIP Conf. Proc.* **2126**, 190013 (2019).
- ¹⁰C. A. Gueymard, J. M. Bright, D. Lingfors, A. Habte, and M. Sengupta, “A posteriori clear-sky identification methods in solar irradiance time series: Review and preliminary validation using sky imagers,” *Renewable Sustainable Energy Rev.* **109**, 412–427 (2019).
- ¹¹J. M. Bright, X. Sun, C. A. Gueymard, B. Acord, P. Wang, and N. A. Engerer, “Bright-Sun: A globally applicable 1-min irradiance clear-sky detection model,” *Renewable Sustainable Energy Rev.* **121**, 109706 (2020).
- ¹²G. Reikard and C. Hansen, “Forecasting solar irradiance at short horizons: Frequency and time domain models,” *Renewable Energy* **135**, 1270–1290 (2019).
- ¹³T. Hove and E. Manyumbu, “Estimates of the Linke turbidity factor over Zimbabwe using ground-measured clear-sky global solar radiation and sunshine records based on a modified ESRA clear-sky model approach,” *Renewable Energy* **52**, 190–196 (2013).
- ¹⁴A. Kuye and S. Jagtap, “Analysis of solar radiation data for Port Harcourt, Nigeria,” *Sol. Energy* **49**, 139–145 (1992).
- ¹⁵R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart, “Modeling daylight availability and irradiance components from direct and global irradiance,” *Sol. Energy* **44**, 271–289 (1990).
- ¹⁶A. Suárez-García, M. Díez-Mediavilla, D. Granados-López, D. González-Peña, and C. Alonso-Tristán, “Benchmarking of meteorological indices for sky cloudiness classification,” *Sol. Energy* **195**, 499–513 (2020).
- ¹⁷F. Battles, F. Olmo, J. Tovar, and L. Alados-Arboledas, “Comparison of cloudless sky parameterizations of solar irradiance at various Spanish mid latitude locations,” *Theor. Appl. Climatol.* **66**, 81–93 (2000).
- ¹⁸R. H. Inman, H. T. Pedro, and C. F. Coimbra, “Solar forecasting methods for renewable energy integration,” *Prog. Energy Combust. Sci.* **39**, 535–576 (2013).

- ¹⁹J. Manning and R. Baldick, "A novel clear-sky index for color sky imagery used in short-term irradiance forecasting," *J. Renewable Sustainable Energy* **12**, 013701 (2020).
- ²⁰M. M. Khan and M. J. Ahmad, "Estimation of global solar radiation using clear sky radiation in Yemen," *J. Eng. Sci. Technol. Rev.* **5**, 12 (2012).
- ²¹F. Antonanzas-Torres, R. Urraca, J. Polo, O. Perpiñán-Lamigueiro, and R. Escobar, "Clear sky solar irradiance models: A review of seventy models," *Renewable Sustainable Energy Rev.* **107**, 374–387 (2019).
- ²²W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory* (Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1986).
- ²³L. H. Koopmans, *The Spectral Analysis of Time Series*, Probability and Mathematical Statistics, 2nd ed. (Academic Press, 1995).
- ²⁴J. F. James, *A Student's Guide to Fourier Transforms*, Student's Guides, 3rd ed. (Cambridge University Press, 2011).
- ²⁵S. Pal and S. C. Bhunia, *Engineering Mathematics* (Oxford University Press, 2015).
- ²⁶M. J. Brooks, S. Du Clou, W. L. Van Niekerk, P. Gauché, C. Leonard, M. J. Mouzouris, R. Meyer, N. Van der Westhuizen, E. E. Van Dyk, and F. J. Vorster, "SAURAN: A new resource for solar radiometric data in Southern Africa," *J. Energy South. Afr.* **26**, 2–10 (2015).
- ²⁷A. Driemel, J. Augustine, K. Behrens, S. Colle, C. Cox, E. Cuevas-Agulló, F. M. Denn, T. Duprat, M. Fukuda, H. Grobe *et al.*, "Baseline Surface Radiation Network (BSRN): Structure and data description (1992–2017)," *Earth Syst. Sci. Data* **10**, 1491–1501 (2018).
- ²⁸J. Augustine, "Basic measurements of radiation at station Sioux Falls (2018–05)," 2019.
- ²⁹J. Augustine, "Basic measurements of radiation at station Sioux Falls (2018–01)," 2019.
- ³⁰E. B. Pereira, "Basic measurements of radiation at station Brasilia (2018–05)," 2019.
- ³¹E. B. Pereira, "Basic measurements of radiation at station Brasilia (2018–12)," 2019.
- ³²J. Cuesta, D. Edouart, M. Mimouni, P. H. Flamant, C. Loth, F. Gibert, F. Marnas, A. Bouklila, M. Kharef, B. Ouchène *et al.*, "Multiplatform observations of the seasonal evolution of the Saharan atmospheric boundary layer in Tamanrasset, Algeria, in the framework of the African Monsoon multidisciplinary analysis field campaign conducted in 2006," *J. Geophys. Res.: Atmos.* **113**, D00C07, <https://doi.org/10.1029/2007JD009417> (2008).
- ³³M. Alia-Martinez, J. Antonanzas, R. Urraca, F. Martinez-de Pison, and F. Antonanzas-Torres, "Benchmark of algorithms for solar clear-sky detection," *J. Renewable Sustainable Energy* **8**, 033703 (2016).
- ³⁴X. Sun, J. M. Bright, C. A. Gueymard, B. Acord, P. Wang, and N. A. Engerer, "Worldwide performance assessment of 75 global clear-sky irradiance models using principal component analysis," *Renewable Sustainable Energy Rev.* **111**, 550–570 (2019).
- ³⁵S. Baika and S. Abdessadak, "Basic measurements of radiation at station Tamanrasset (2018–08)," 2018.
- ³⁶S. Baika and S. Abdessadak, "Basic measurements of radiation at station Tamanrasset (2018–12)," 2019.
- ³⁷A. Lupi, "Basic and other measurements of radiation at Concordia Station (2018–03)," 2019.
- ³⁸A. Lupi, "Basic and other measurements of radiation at Concordia Station (2018–11)," 2018.
- ³⁹W. F. Holmgren, C. W. Hansen, and M. A. Mikofski, "PVLIB Python: A Python package for modeling solar energy systems," *J. Open Source Software* **3**, 884 (2018).
- ⁴⁰P. Ineichen and R. Perez, "A new airmass independent formulation for the Linke turbidity coefficient," *Sol. Energy* **73**, 151–157 (2002).