Energy 284 (2023) 129046

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Modelling the spectral influence on photovoltaic device performance using the average photon energy and the depth of a water absorption band for improved forecasting

Rajiv Daxini *, Robin Wilson, Yupeng Wu

Faculty of Engineering, University of Nottingham, Nottingham, NG7 2RD, UK

ARTICLE INFO

Keywords:
Photovoltaic performance
Average photon energy
Spectrum
Water absorption band
Spectral correction
Forecasting

ABSTRACT

Accurate forecasts of solar panel performance can improve grid penetration, enable cost evaluation prior to project implementation, and improve fault detection during operation. Spectral correction functions (SCFs) used to model the influence of the solar spectrum in such forecasting models are typically based on either proxy representations of the spectrum, using parameters such as air mass, or parameters derived directly from the spectrum, such as the average photon energy (APE). Although the latter is more accurate, the APE is argued in some studies not to be a unique characteristic of the spectrum and to suffer from increased uncertainty when analysing spectra at longer wavelengths. This study first derives APE spectral correction function coefficients for three PV technologies — multicrystalline (mSi), triple junction amorphous silicon (aSi-T), and cadmium telluride (CdTe). Based on an analysis of uncertainty in the SCF for each of the three devices, this study proposes a new spectral correction function based on the average photon energy, φ, and the depth of a water absorption band, ε. The additional index enables spectra to be characterised by unique combinations of φ and ε. Several water absorption bands are tested and the 650–670 nm band is found to yield the most accurate SCF for all three PV devices. An optimal parameterisation of the SCF for each PV device, as well as a cost-accuracy-balanced parameterisation, is presented. Improvements in the prediction accuracy of up to 60% for both the mSi and aSi-T modules, and around 20% for the CdTe module, are achieved by the proposed model with respect to a comparable two-variable proxy SCF, namely the air mass and precipitable water function. Compared with the single-variable APE SCF, $f(\varphi)$, the proposed model improves the prediction accuracy by around 10% for the aSi-T and mSi modules, and by around 2% for the CdTe module. No new data are required for the proposed model compared with $f(\varphi)$ as the same spectra used to calculate the APE are used to calculate ε. The proposed spectral correction function can easily be integrated into wider photovoltaic performance models for improved forecasting.

1. Introduction

Solar irradiance is one of the main factors affecting PV performance [1,2]. However, due to the non-linear spectral response of photovoltaic (PV) modules, measurements of broadband irradiance alone are insufficient for accurately modelling the energy yield of PV systems under realistic operating conditions.

The impact of spectral irradiance on PV performance has been studied extensively [3–7] and shown to be especially significant for wide band gap semiconductor technologies such as amorphous silicon and cadmium telluride. Its omission from performance models can lead to annual forecasting errors of up to 20% for such devices [8–10]. For concentrating photovoltaic (CPV) and multi-junction (MJ) systems, spectral errors of up to 30% have been reported in locations with extreme atmospheric conditions [11]

The performance of any particular PV device is usually known for a set of reference test conditions (RTC), which include a reference spectrum. To correct for the difference in performance due to the spectrum under realistic operating conditions (ROC), a spectral correction function (SCF) is used. SCFs are either based on proxy variables used to represent the spectrum, or values derived directly from the solar spectrum.

Proxy variables are environmental parameters that are considered to have a dominant effect on the solar spectrum, and therefore whose

* Corresponding author.
E-mail addresses: Rajiv.Daxini@nottingham.ac.uk (R. Daxini), Yupeng.Wu@nottingham.ac.uk (Y. Wu).

https://doi.org/10.1016/j.energy.2023.129046
Received 21 February 2023; Received in revised form 24 July 2023; Accepted 9 September 2023
Available online 13 September 2023

value can serve as an indirect indicator of the shape of the prevailing spectrum. Examples of such variables include air mass (AM) [12], clearness index (K_t) [13], and atmospheric precipitable water content W [14]. The proxy variables are chosen based on the significance of their impact on the solar spectral distribution and the specific spectral response characteristics of the devices under investigation. For example, AM values between 1.0 and 5.0 have been found to result in performance fluctuations of up to 5% in c-Si modules and over 5% for CdTe modules [15]. Furthermore, CdTe devices have been found to be particularly susceptible to changes in W, with variation between 0.5–5.0 cm resulting in performance fluctuations of −4% to + 5% [15].

In terms of values derived directly from the spectrum, the average photon energy (APE) [16] is frequently used [17–20]. The APE is the mean energy of the photons contained within a predetermined waveband of spectral irradiance. A higher average photon energy indicates a blue-shifted spectrum, while a lower energy indicates a red-shifted spectrum. Since the parameter selection in proxy-variable SCFs offer only a limited range of information about the spectrum, the APE has been used as an alternative parameter to characterise the solar spectrum and investigate its influence on PV performance. Several studies have shown that spectra with a higher average photon energy can lead to increases in the operation efficiency of aSi (single– and double-junction), CdTe, and perovskite PV devices, while decreases are observed for technologies such as mSi, mono-Si, and Copper-Indium-Selenium PV devices [21–23]. These studies analyse in detail the relationship between PV performance and the spectrum, as characterised using the APE, but do not go on to develop an APE-based SCF for PV performance forecasts. Where regression analysis has been used to quantify the correlation between the APE and a PV performance indicator [24], validation of the proposed model(s) in terms of predictive accuracy is lacking. Daxini et al. [25] proposed a new SCF based on the APE, whose predictive accuracy is validated using field data and compared to that of two proxy-variable SCFs, namely the air mass function, $f(AM_a)$ [12], and the combined air mass and clearness index function, $f(AM_a, K_t)$ [8]. The APE SCF is shown to reduce the mean absolute prediction error by almost 50% with respect to $f(AM_a)$ and almost 40% with respect to $f(AM_a, K_t)$. However, the proposed model is only validated for a single amorphous silicon PV device and therefore it is necessary for future work to validate the model with different types of PV technology.

Another limitation of the aforementioned APE-PV studies centres on the uniqueness of the APE parameter in terms of its ability to represent solar spectra. Dips in spectral irradiance in one particular waveband may be countered by increases in irradiance in another waveband, leading to two differently shaped spectra but the same average photon energy. By adopting the International Electrotechnical Commission's methodology for rating the spectral matching of a solar simulator [26], Minemoto et al. [27] show that an APE value yields a spectral irradiance distribution with a relatively small standard deviation . These results are supported by Tsuji et al. [28] who conclude the APE is a unique characteristic of the solar spectrum at three test sites across Japan. On the other hand, more recent research has cautioned against using the APE parameter for spectral analysis [7,29]. For example, Nofuentes et al. (2017) analyse the coefficient of variation of the APE parameter, rather than the standard deviation as used by Minemoto et al. and find that APE values used to represent the spectrum have an uncertainty of over 3% between 450–900 nm, and 5%–11% outside of this range [30]. They conclude that the APE is not a bijective index and cannot be used reliably for spectral analysis. Given that the only study to have derived and validated an SCF based on the APE parameter demonstrates the proposed methodology solely for aSi PV devices, which have a spectral response range below 900 nm, the bijectivity issue highlighted in [30] may not have been a problem. Two research gaps in the existing literature are clear:

1. The use of the APE parameter to derive an SCF for PV devices with wider spectral responses.
2. Uncertainty in the APE parameter when analysing spectral effects on PV devices, in particular those with wider spectral responses.

Ishii et al. [31] show that the primary driver for changes in the shape of spectra that maintain the same APE are negatively correlated depths of water absorption bands (ε_w) and atmospheric windows (ε_a), hereinafter collectively referred to as "spectral bands", or ε. It is suggested that a solar spectral distribution may therefore be characterised uniquely for all wavelengths with both the APE and an additional index, namely the depth of one such spectral band.

In this work, an SCF based on the APE parameter is derived and validated for a range of different PV technologies, including those with a spectral response beyond 900 nm. Based on an analysis of the uncertainty in the single-variable APE SCF, an advanced model based on both the APE and the depth of a spectral band is proposed to address the uncertainty in $f(\varphi)$ caused by the uncertainty in the φ. The new model is developed and validated using empirical data for three PV technologies — multicrystalline, cadmium telluride, and triple-junction amorphous silicon. The combined φ-ε model addresses the bijectivity issue of the APE parameter whilst still retaining the higher accuracy achieved through the use of direct representation of the spectrum rather than the traditional proxy representation. As part of the model validation, the predictive accuracy of the proposed model is compared to that of the APE spectral correction and a proxy-variable function. New model coefficients are derived for the latter, which represent an additional contribution of this work to the literature.

2. Methodology

In this section, first, the overall research framework is presented. The following subsections then explain methods and parameters used for data collection, processing, and analysis.

2.1. Research framework

Fig. 1 is a flow chart illustrating the research framework and overall processes undertaken to achieve the main aims of this study, which are to validate the $f(\varphi)$ SCF for different PV technologies, and develop and validate a new SCF based on two spectral indices — the average photon energy (φ) and the depth of a water absorption band (ε).

The first stage of the work is to derive the $f(\varphi)$ model coefficients for three PV technologies not previously examined in the literature. Based on an analysis of the uncertainty in the $f(\varphi)$ correlation for these technologies, and published spectral irradiance analysis, four spectral bands are determined as potential candidates for ε in the proposed $f(\varphi, \varepsilon)$ spectral correction model. An iterative algorithm is used to fit different surface functions to a correlation of $I_{scn} = f(\varphi, \varepsilon)$, for each ε band and for each device under test (DUT). In this stage, the optimal fitting function is determined for each ε band, for each DUT. The best fitting function for each ε band is then used to make predictions of I_{scn}, which are compared with measurement-derived values of I_{scn}, to determine the optimal ε band for each DUT. This leads to a final model of $f(\varphi, \varepsilon)$ with the optimal ε band and functional form. The final $f(\varphi, \varepsilon)$ SCF for each DUT is then validated through a comparison of its fitting and I_{scn} prediction accuracy with those of existing SCFs, namely $f(\varphi)$, and $f(AM_a, W)$.

2.2. Spectral shift in PV performance

As spectral changes affect the measured short-circuit current of a PV device, I_{sc}, a normalised form of I_{sc}, is used to characterise PV performance under the prevailing spectrum. I_{sc} is first translated to a reference temperature (T_r = 25 °C) and reference irradiance (G_0 = 1000 W m^{-2}) as follows [32]:

$$I'_{sc} = \frac{I_{sc}}{1 + \hat{\alpha}_{I_{sc}}(T_c - T_r)} \left[\frac{G_0}{G_{poa}}\right]. \quad (1)$$

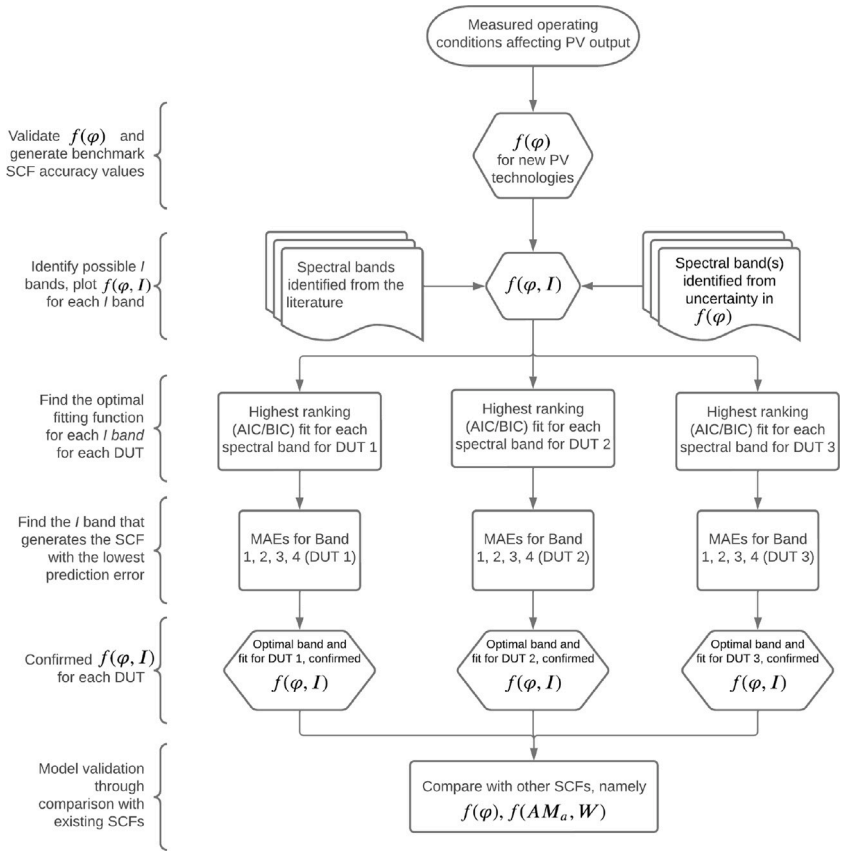

Fig. 1. Research framework.

Here, $\hat{\alpha}_{I_{sc}}$ [°C^{-1}] is the short-circuit current temperature coefficient, T_c [°C] is the cell temperature, and G_{poa} is the measured plane of array irradiance. T_c is estimated using a one-dimensional thermal conduction model [32] based on the module temperature, T_m, which is measured:

$$T_c = T_m + \frac{G_{poa}}{G_0}\Delta T. \tag{2}$$

ΔT is the temperature difference between the module and cell and is determined based on the structure of the DUT. ΔT is set as 3 °C in this study based on the structure of the three modules tested.

The purely spectral influence is expressed by the ratio of I'_{sc} to the reference current, I_{sc0}, which is the current measured at T_r, G_0, and AM1.5 (reference spectral irradiance):

$$I_{scn} = \frac{I'_{sc}}{I_{sc0}}. \tag{3}$$

2.3. Characterisation of the solar spectrum

The two parameters – APE and ε – used to characterise the measured solar spectral distribution are introduced in this section.

2.3.1. Average photon energy

The average photon energy, APE or φ, is an indicator of the overall shape of the solar spectrum. It is calculated by dividing the total energy in the spectrum by the number of photons it contains, as follows:

$$\varphi[\text{eV}] = \frac{1}{q}\left(\frac{\int_a^b E_\lambda d\lambda}{\int_a^b \Phi_\lambda d_\lambda}\right). \tag{4}$$

$E(\lambda)$ [W m^{-2} nm^{-1}] is the spectral irradiance, $\Phi(\lambda)$ [m^{-2} nm^{-1}] is the spectral photon flux density, q [C] is the electron charge, and a [nm] and b [nm] are the upper and lower wavelength limits, respectively, of the considered waveband.

2.3.2. Depth of a spectral band

The depth of any spectral band, ε, is calculated by integrating the spectral irradiance between the lower (c) and upper (d) wavelength limits of the specified band,

$$\varepsilon = \int_c^d E_\lambda d\lambda. \tag{5}$$

Fig. 2 shows the intercorrelation between the atmospheric window and water absorption band in the solar spectrum, as first observed by Ishii et al. (2012) [31]. One pair of bands is highlighted as an example. Band A (1000–1050 nm) is an atmospheric window and Band B (1110–1160 nm) is a water absorption band. The red and blue arrows indicate a higher and lower spectral irradiance in Band A and B, respectively. Therefore, as the spectral irradiance (SI) in Band A increases, that of Band B decreases, and vice versa.

In this study, the recommendation of Ishii et al. (2012) to use a combination of (φ, I) to characterise the solar spectrum is adopted to develop a new spectral correction function. The band selection is discussed in Section 6.

2.4. Devices tested

The devices tested in this study are multicrystalline silicon (mSi), cadmium telluride (CdTe), and amorphous silicon triple junction (aSi-T). A summary of the technical specifications of each device tested is presented in Table 1.

These three devices are a representative sample of PV technologies to demonstrate the proposed model from a technical perspective as well as the perspective of meaningful practical application. Firstly, amorphous and CdTe devices are especially important for building-integrated photovoltaics applications [36–40]. The application aSi-based devices in particular has also been explored for PV-thermoelectric systems [41–43]. In terms of mSi, their positive balance between

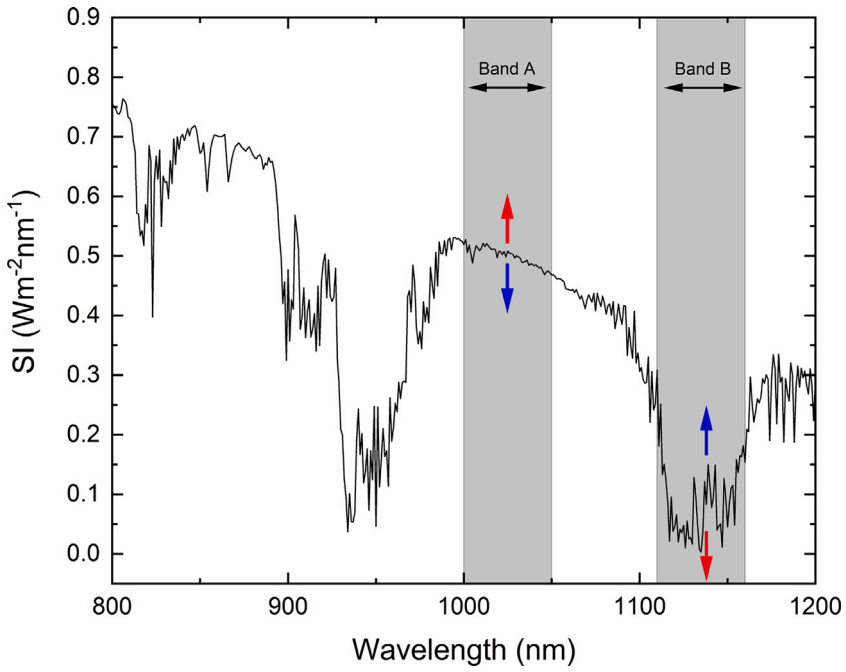

Fig. 2. Atmospheric window (Band A, 1000–1050 nm) and water absorption band (Band B, 1110–1160 nm) as highlighted in [31]. The red and blue arrows indicate the negative correlation between the spectral irradiance (SI) in bands A and B. As the SI of band A increases (red upwards), that of band B decreases (red downwards), and vice versa with the blue arrows. The SI data were measured at the University of Nottingham test site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Technical specification and database identifier of the three DUTs as detailed in the NREL database [33–35].

Parameter	DUT		
	aSi-T	CdTe	mSi
Database identifier	aSiTriple28325	CdTe75669	mSi0251
I_{sc} (A)	4.39	1.14	2.67
V_{oc} (V)	23.22	85.15	21.90
P_{mpp} (W)	58.67	65.77	43.88
$\alpha_{I_{sc}}$ (%°C^{-1})	0.000981	0.00051	0.00057

Table 2
Equipment at the OTF and their respective measurements.

Parameter	Instrument
Wind speed/direction, precipitation, temperature, relative humidity, pressure	Vaisala WCT529 sensor
Plan-of-array irradiance	Kipp & Zonen CMP22 pyranometer
Datalogger	Campbell Scientific CR1000
Data Logger Communications	RAVEN XE-EVDO
PV module IV curve	Daystar MT5 multi-tracer
Module back surface temperature	Omega CO1-T Style I thermocouple

Table 3
Additional data processing measures undertaken to filter the NREL data measured in Golden, Colorado, USA. The $I_{scn} = 0.17$ example point occurred on 2013-01-02 at 1045h.

Retention criteria	Explanation
Irradiance > 200 W m^{-2}	Reduce noise but retain overcast conditions
Time 0800 h–1600 h	Retain daylight hours only
Manual I_{scn} filter	Example: $I_{scn,CdTe} = 0.17$, all other $I_{scn} > 0.8$

cost and efficiency means that they have dominated the global PV market [44] for rooftop applications, in particular on residential properties [45]. The devices selected cover a range of PV types, namely thin film, multijunction, and crystalline, and have different spectral response characteristics, thus providing a comprehensive sample of technology types and constructions against which the proposed methodology can be validated.

2.5. Data acquisition and analysis

The PV performance and meteorological data used in this study are sourced from a public database published by the National Renewable Energy Laboratory (NREL) [33–35] and used widely in a range of PV and meteorological studies [46–49]. The data used in this study were measured at NREL's Outdoor Test Facility (OTF) located in Golden, Colorado. The parameters measured and the devices used to measure them at the OTF are summarised in Table 2. The database's associated publications [33–35] explain the measurement campaign, data processing, and data reliability and uncertainty in more detail. The additional data processing actions undertaken in this work are summarised in Table 3.

The PV performance and meteorological data are only measured for a total duration of one year. However, two datasets are required in this analysis — one to develop the new SCF and a second to validate it. It is essential to maintain sufficient data in both sets to ensure they are representative of real operating conditions throughout the entire year [50–52]. Since multiple years of data are not available, the available one-year dataset was sorted chronologically and split by extracting data at a ratio of 2:1, which is a commonly used ratio in both PV and non-PV research [53–55]. This means that, moving through the dataset chronologically, every third measurement of the required parameters (SI, G_{poa}, I_{sc}, etc.) was extracted to form the model validation dataset while the remaining two were retained for the model development. This separation enables the application of out-of-sample testing in the validation stage where data not previously used in the model development are used to validate the model. It also means that data from the start to the end of the year are included in both the model development and validation datasets.

The Levenberg–Marquardt (LM) algorithm is used in the fitting process to determine the regression coefficients for correlations in this study. For each spectral band used in the new model, the surface

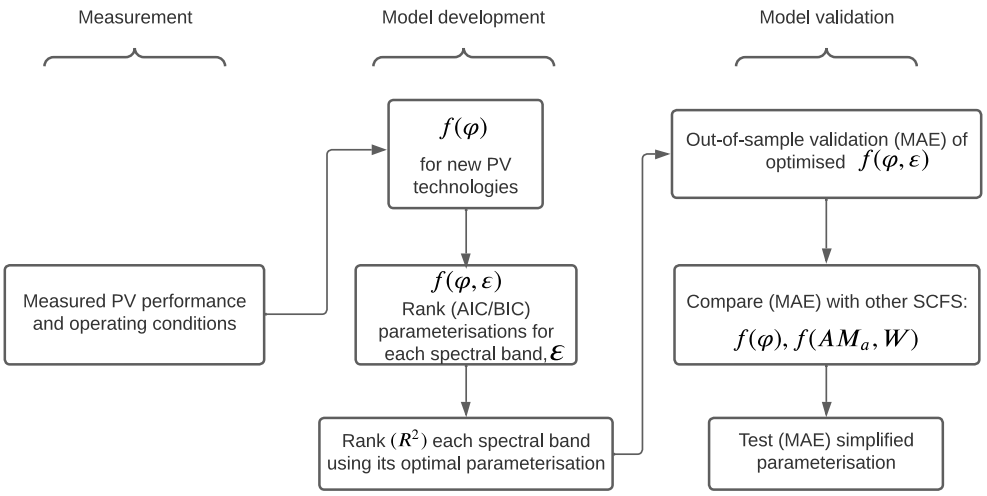

Fig. 3. Structure of the analysis in this study.

functions fitted to the data are ranked according to their Bayesion and Akaike Information Criterion (BIC, AIC) scores to determine the best fit function for each spectral band. The BIC and AIC yield a rating that balances accuracy and complexity of the model [56], although the BIC involves a greater penalty for increased complexity due to the higher weighting of the term for the number of model parameters. The coefficient of determination, R^2, is calculated for the best fit surface function for each spectral band to enable a comparison of fitting accuracy between the models. A higher R^2 indicates that the surface function explains more of the variation in the dependent variable. However, a higher R^2 value does not necessarily mean a better prediction model as there is always a risk of overfitting to the data used to develop the model [57,58]. Therefore, in the validation stage, the optimal parameterisation, according to the BIC/AIC scores, for all four spectral bands is used to predict values of I_{scn}.

The prediction accuracies of the models based on different spectral bands are then compared to determine the optimal ε band. In addition, comparing this ε-band optimisation result to the ε band ranking according to the R^2 values offers an indication of whether any overfitting exists in the model parameterisation.

The prediction accuracy of $f(\varphi,\varepsilon)$, using the optimal ε band and functional form, is then compared to the accuracy of existing SCFs, namely $f(\varphi)$ and $f(AM_a,W)$ to test whether there is any benefit to using the new model. Finally, the sensitivity of the prediction accuracy to the functional form of $f(\varphi,\varepsilon)$ is then analysed to determine the flexibility of the model for different use cases that could impose varying levels of computational resources or requirements.

2.6. Summary

The structure of the remaining analysis in this study is summarised in Fig. 3. The flow chart describes each step of the analysis and, where applicable, the statistical metrics adopted at each stage are included in brackets. The measurement stage is based on the data collection and processing described in Section 2.5. The measured meteorological and PV performance data are presented in Sections 3 and 4. The model development stage is presented across Sections 5 and 6, where new coefficients of $f(\varphi)$ are presented and the new $f(\varphi,\varepsilon)$ model is developed and optimised. The model validation stage is presented across Section 7, where the optimised form of $f(\varphi,\varepsilon)$ is validated, and the I_{scn} prediction accuracies of $f(\varphi,\varepsilon)$, $f(\varphi)$ (new coefficients), and $f(AM_a,W)$ are compared. In the final step, a simplified parameterisation of $f(\varphi,\varepsilon)$ is validated in Section 7.4.

3. Climatic conditions at the test site

The Golden test site is at an elevation above sea level of approximately 1800 m. The monthly variation in irradiance and ambient temperature at the site are shown in Fig. 4. The climate is a typical mild-to-cold winter (mean temperature 10 °C) and warm–hot summer (mean temperature 30 °C). The monthly mean irradiance on the plane of the PV test array varies between approximately 600 W m^{-2} and 700 W m^{-2} throughout the year.

Fig. 5 presents the variation in the monthly mean values of several meteorological parameters that directly affect the solar spectrum, namely air mass, atmospheric precipitable water content, and aerosol optical depth at 500 nm (AOD). AOD is measured at the OTF, while W is calculated using measurements of relative humidity and ambient air temperature [59,60]. AM_a is calculated using the Kasten and Young model [61] combined with a linear pressure correction based on the site elevation [62] to account for the decreased density of air molecules at different elevations [63]. The air mass follows the expected seasonal sinusoidal trend that is dictated by the change in solar elevation over the course of one year. Atmospheric water vapour follows a similar but inverted sinusoidal pattern, which is similar to that of the annual ambient air temperature variation. Increased temperatures in the summer lead to increased surface water evaporation and the greater capacity of warmer air to hold moisture [64,65] The range of these parameters is between around 1.0–2.5 and 0.5–2.0 cm, which are typical of this region. The AOD levels break the sinusoidal trend with an atypical spike in April. Historical weather data from a variety of meteorological stations and airports, compiled in Ref. [66], report unstable conditions during April of 2013. Most days during this period featured one or a combination of rainfall, haze, and dust. The latter two, in particular, may have contributed to the increase in aerosol optical depth during the month of April 2013.

Fig. 6 shows the annual variation in the daily mean and monthly mean values of the average photon energy at the OTF. The variation in APE also exhibits a sinusoidal pattern with a decrease in the winter months (October–February) relative to the summer months March–September. The dashed line shows the AM1.5 reference spectrum, which, between 350–1050 nm, is equal to 1.88 eV [67]. Relative to the AM1.5 reference spectrum, the prevailing spectral irradiance conditions during the summer months are similar to AM1.5, with a slight blue shift in July. On the other hand, the winter months experience significantly red-shifted spectra relative to the AM1.5 spectrum. This red shift in winter results from shorter-wavelength radiation being more strongly affected by Raleigh scattering when the solar elevation is lower.This

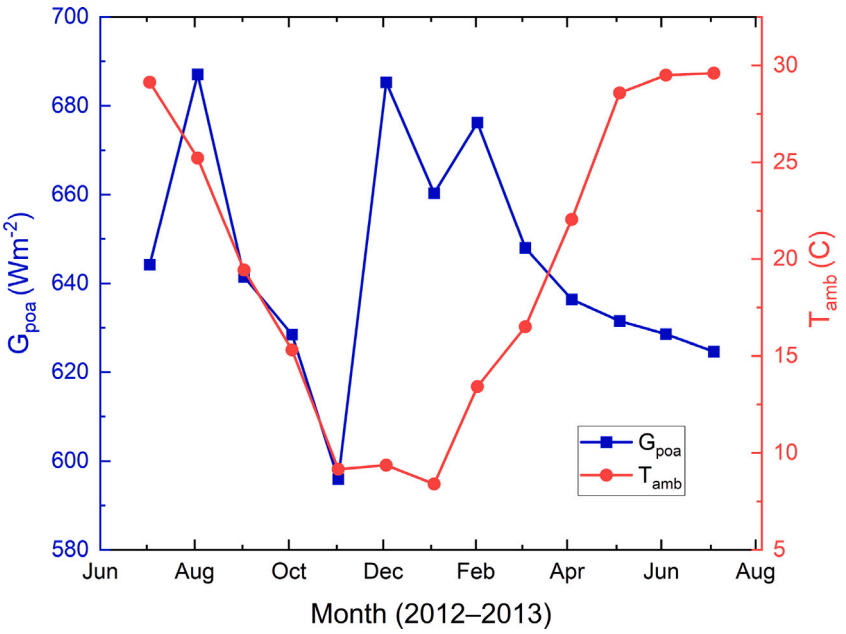

Fig. 4. Mean monthly global plane-of-array irradiance (G_{poa}) and ambient temperature (T_{amb}) measurements at the OTF.

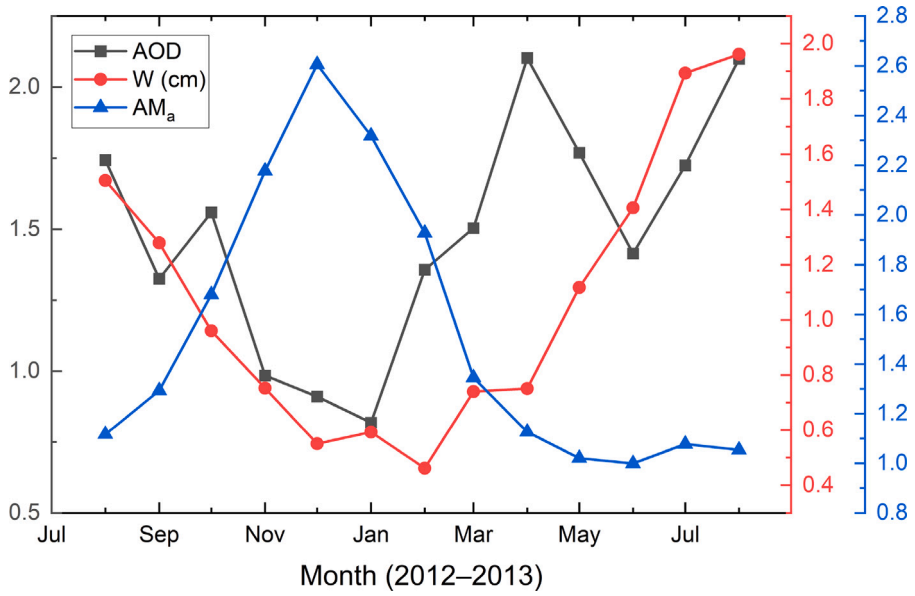

Fig. 5. Annual variation in several key environmental parameters that influence the solar spectrum in Golden, Colorado. Data points represent the monthly mean values for each parameter.

pattern is in line with the trends observed in Fig. 5 where atmospheric variables that result in greater scattering of longer-wavelength light, resulting in a blue-shifted spectrum, are more prevalent in summer months, during which time the solar elevation (air mass) is also higher (lower). As a result of the winter red shift, the PV devices, in particular the CdTe- and aSi-based technologies, may operate less efficiently ($I_{scn} < 1$) in the winter compared with the summer. The annual variation in I_{scn} is discussed in greater detail in Section 4.

4. Spectral dependence of PV performance

Before presenting the results of the SCF analysis, the spectral response characteristics of the DUTs are presented in this section. In addition, an analysis of the temporal variation in I_{scn} over the course of one year is presented for each DUT to illustrate their respective dependence on the solar spectrum.

Fig. 7 shows each device's normalised spectral response, behind which the AM1.5 reference spectrum is plotted. CdTe has the narrowest spectral response and exhibits sharp cut off wavelength at around 900 nm, which corresponds to its band gap of around 1.5 eV [70]. The wide band gap (narrow spectral response range) results in a relatively strong spectral dependence of CdTe performance as CdTe PV devices can only utilise a relatively small proportion of the available spectral irradiance. The other two devices – aSi-T and mSi – all have much broader spectral response ranges. For the mSi device, this results in a reduced spectral dependence. However, this is not the case with aSi-T, which is a triple-junction device. In a multijunction device, the spectral responses of each junction are combined to provide a wider overall

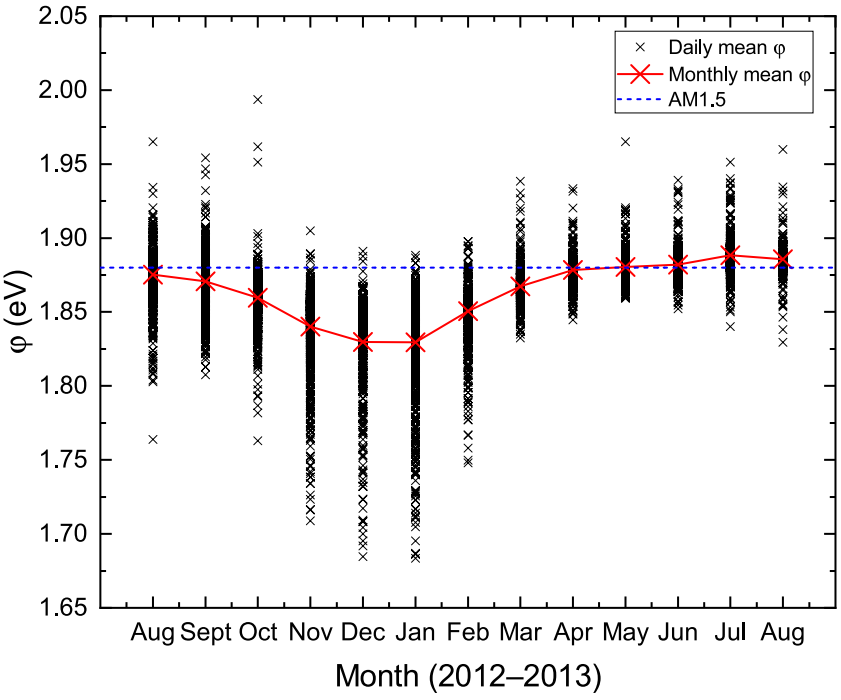

Fig. 6. Annual variation in the average photon energy at the OTF in Golden, Colorado.

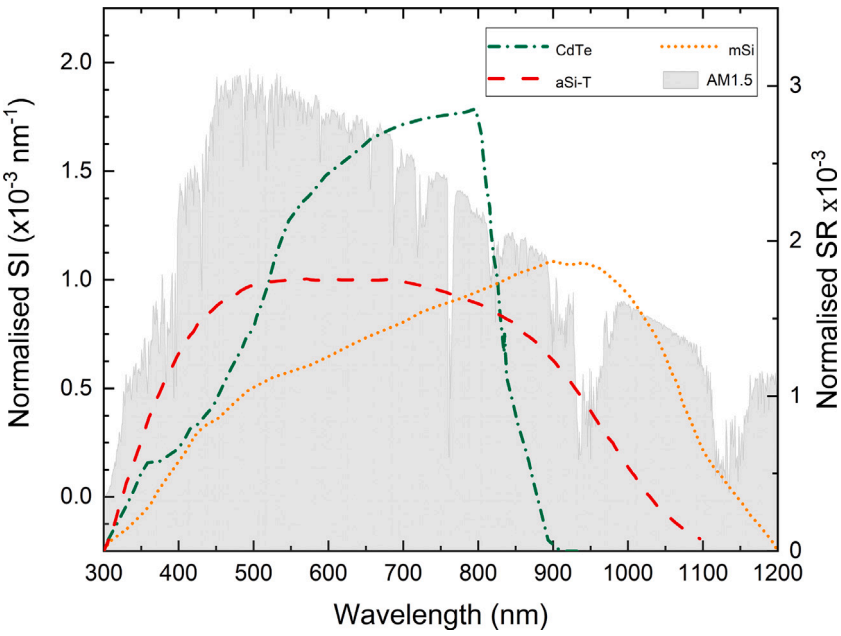

Fig. 7. Normalised spectral response (SR) of the three PV technologies investigated in this study. Behind the spectral response curves is the spectral irradiance (SI) for the AM1.5 reference spectrum, normalised between 280 and 1200 nm. SR data are sourced from Ref. [68] (mSi and CdTe) and Ref. [69] (aSi-T).

spectral response range for the device. However, the construction of multijunction devices is such that the junctions may be considered as cells connected in series. Therefore, the current flowing through each junction must be equal, which limits the overall device current to that of the least productive junction. Therefore, multijunction devices exhibit a strong spectral dependence [71,72] due to the particular spectral response characteristics of the individual subcells (junctions).

To understand the spectral response behaviour of the DUTs quantitatively in the field, Fig. 8a shows the annual variation of I_{scn} for the devices. I_{scn} is an indicator of the purely spectral influence on the short-circuit current of a PV module, as defined in Section 2.2. Values of $I_{scn} > 1$ indicates higher performance under the prevailing spectrum with respect to the performance under the reference conditions, $I_{scn} < 1$ indicates decreased performance, while $I_{scn} = 1$ indicates there is no difference between the performance under the prevailing spectrum and under reference conditions.

As expected, the performances of the wider band gap (CdTe) and multijunction (aSi-T) devices have a stronger spectral dependence than the narrower band gap single-junction device (mSi). With respect to RTC (where $I_{scn} = 1$), the monthly mean variations plotted in Fig. 8a show variation of up to −9%, +6% for aSi-T, −5%, +1% for CdTe, and −2%, +0% for mSi. The maximum observed variations on the 15-minute-averaged timescale of the original dataset are −40% to +20% for aSi-T, −20% to +10% for CdTe, and around ±5% for mSi. It is important

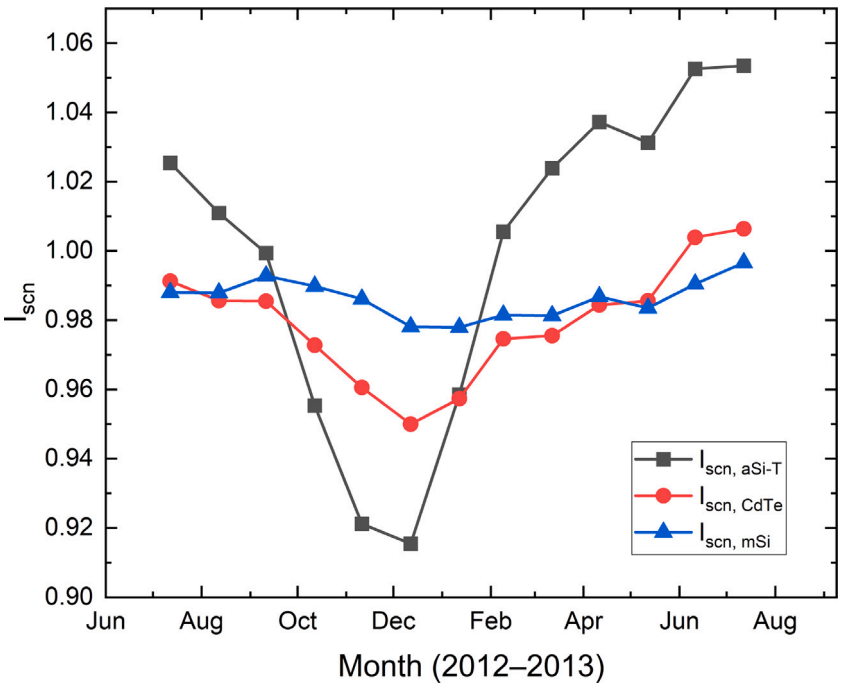

Fig. 8a. Annual variation in monthly mean I_{scn} for the three DUTs — aSi-T, CdTe, and mSi.

to note that I_{scn} for the aSi-T device may be influenced by non-spectral effects that are not accounted for in the normalisation process. In the summer in particular, thermal annealing of the device may lead to an increase in efficiency [73,74], which could in part be responsible for the significant rise in I_{scn}, in particular when compared with the mSi and CdTe devices.

Variation in I_{scn} for all three devices follow a seasonal sinusoidal pattern over the year. The reduction in I_{scn} in the winter months may be attributed to an average drop in the solar elevation during this time of the year. The resultant increase in solar air mass shifts the solar spectrum to longer wavelengths on average, which results in a reduction in efficiency, especially for the CdTe and aSi-T devices. This change in the solar spectrum is also indicated by a drop in the average photon energy at the same time, which is shown in Fig. 6.

5. APE spectral correction

In this section, the coefficients of $f(\varphi)$ are derived for each of the three DUTs, which expands on Ref. [25]. The coefficients for each of the modules are summarised in Table 5. An analysis of the uncertainty in $f(\varphi)$, in addition to a review of existing studies, informs the selection of the spectral band used for ε in the main analysis of this study where a new SCF based on the APE and ε is presented.

5.1. APE spectral correction

Based on the discussion in Section 4, the correlation between I_{scn} and φ for the mSi device would be expected to be weak compared to the same correlations for the aSi-T and CdTe devices. This estimation is borne out in the results shown in Fig. 8d. The coefficients of determination, R^2, for each of the fits are listed in Table 4.

All three devices show an increase in efficiency with the average photon energy, although the rate of increase is greatest for the aSi-T and CdTe modules. For the CdTe module, which has a relatively wide band gap, the blue-shifted spectra (higher APE) contain a larger proportion of photons with sufficient energy to generate a photocurrent. For the aSi-T module, a similar principle applies but in this case it is the top junction (wide band gap) of the module that is engaged more effectively in the overall current generation by the blue-shifted spectra. Generally speaking for both the CdTe and aSi-T modules, the active region for these thin-film technologies is typically within the first micrometer of the cell (top-cell in the aSi-T case). This improves their spectral response to shorter-wavelength irradiation as irradiance in the blue-visible range is typically absorbed within 0.1–1 μm of cell depth [75]. Although the mSi data also show an increase in I_{scn} with APE, the correlation is weak and the fit is unreliable as evidenced by the extremely low R^2 (< 0.01). The increased uncertainty in the mSi fit may be attributed in part to the wider and flatter spectral response range of the module, which means that the device performance has a greater dependence on APE values that have a higher uncertainty.

Table 4
R^2 values for the fourth order polynomial fits of $I_{scn} = f(\varphi)$ for each DUT.

DUT	R^2
aSi-T	0.90
CdTe	0.54
mSi	0.10

The aSi-T data show the least variability of the three PV devices, with the tightest correlation and highest R^2. However, one notable feature across all devices is that in the range of approximately $1.85\text{eV} \leq \varphi \leq 1.90\text{eV}$ there is a larger spread in I_{scn} relative to the variation in φ, when compared to other values of φ. In this φ range, the same or similar values of φ yield different values of I_{scn}. This means that spectra resulting in different levels of PV performance possess APE values that are insufficiently different to represent the change in I_{scn} to which they lead. This observation substantiates previous work that argues the APE index is incapable of uniquely representing solar spectral distributions [29,30]. Improvement in the reliability and prediction accuracy of $f(\varphi)$ is likely to be achieved through attributing the different I_{scn} values to unique value(s) representing the spectrum. Previous research has identified variation in atmospheric water vapour to be the driver of uncertainty in the APE value [30,31]. Several studies have also found that atmospheric water vapour has a significant impact on the performance of solar panels, in part through its effect on the solar spectrum [76,77]. These studies find a high variation in PV output due

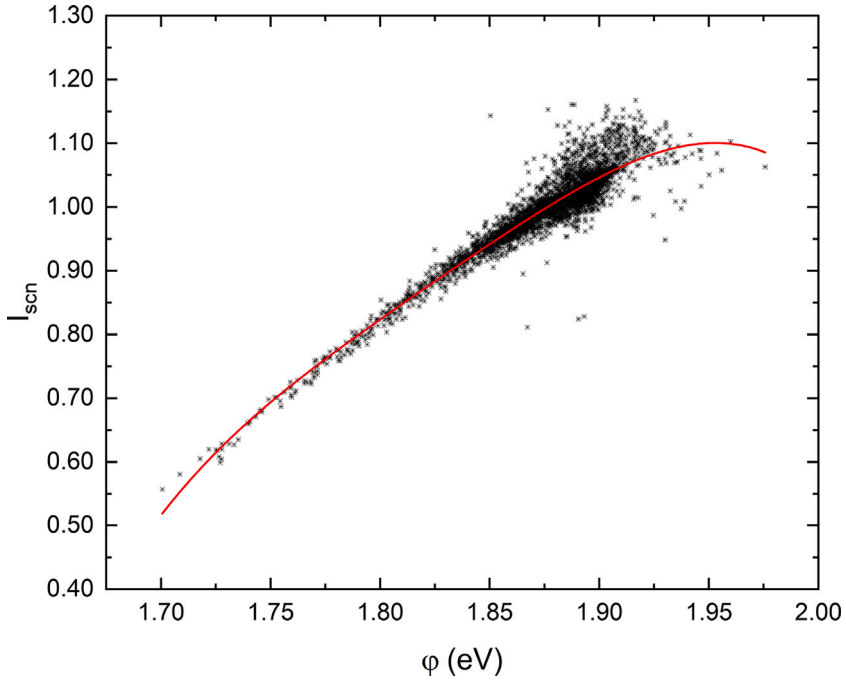

Fig. 8b. $I_{scn} = f(\varphi)$ SCF for the aSi-T device. The coefficient of determination, R^2 is high, at around 0.90.

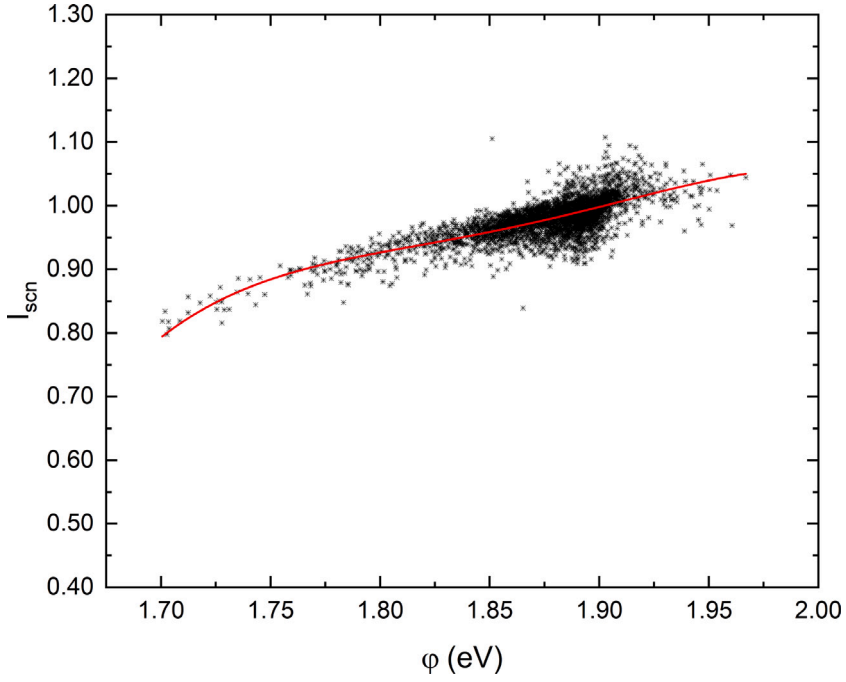

Fig. 8c. $I_{scn} = f(\varphi)$ SCF for the CdTe device. The coefficient of determination, R^2 is medium-high, at around 0.54.

to variable atmospheric precipitable water content levels, in particular for CdTe modules. Guechi et al. [78] find that the short-circuit current of a CdTe module can vary by over 3% as a result of variation in W from 0.5 to 4.0 cm, while a hydrogenated silicon device is found to vary by less than 2.5%. Passow et al. [15] report a more significant variation in CdTe performance of −4% to + 5% when W ranges between 0.5–5.0 cm.

Four 20 nm-wide wavebands centred about 660 nm, 720 nm, 815 nm, and 940 nm, which are identified in [78,79], are used as the positions of water absorption bands in the subsequent analysis. Through comparing the SCFs derived from multiple wavebands, it is possible to determine whether the uncertainty around $1.85\text{eV} \leq \varphi \leq 1.90\text{eV}$ can be resolved through the inclusion of a water absorption band in the $f(\varphi)$ SCF — $f(\varphi, \varepsilon)$.

In this section, model coefficients for $f(\varphi)$ have been derived for the three PV technologies investigated in this study — aSi-T, CdTe, and mSi. The model coefficients are presented in Table 5. $f(\varphi)$ coefficients for these three PV types have not previously been published in the

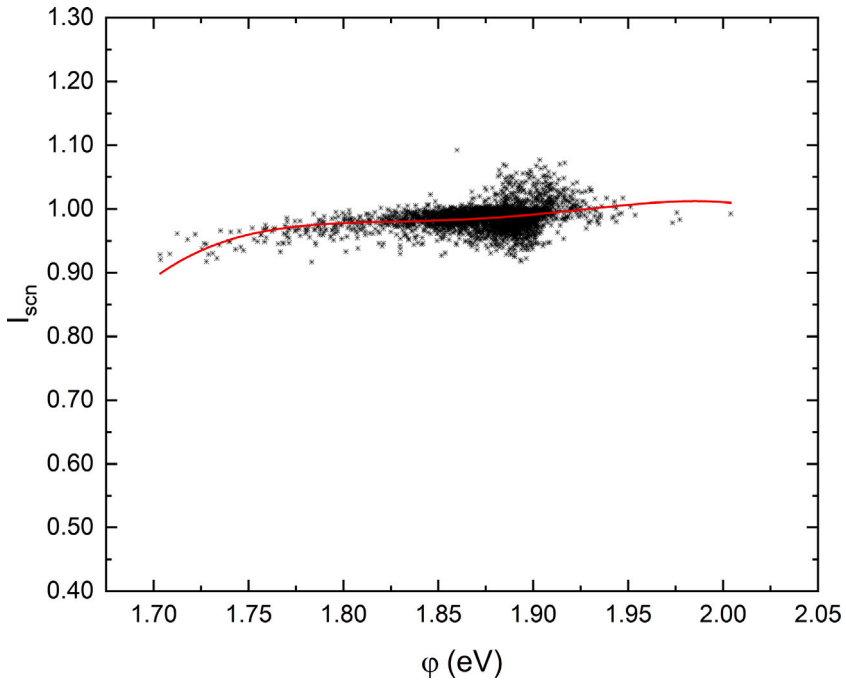

Fig. 8d. $I_{scn} = f(\varphi)$ SCF for the mSi device. The coefficient of determination, R^2 is low, at around 0.1.

Table 5
Fourth order polynomial coefficients for $f(\varphi)$ for each PV device. The fits to which these values refer are shown in Figs. 8b, 8c, and 8d for the aSi-T, CdTe, and mSi devices, respectively.

Device	Coefficient				
	a_0	a_1	a_2	a_3	a_4
aSi-T	−2681.4825	5873.6537	−4828.0300	1764.8774	−241.9810
CdTe	−1745.5747	3752.4391	−3022.5415	1081.5722	−145.0411
mSi	−1469.6501	3139.0754	−2511.0929	892.1727	−118.78122

literature. Furthermore, analysis of the uncertainty in $f(\varphi)$ for each of these three devices, in addition to a review of existing literature, has helped identify four spectral bands as candidates for ε in the $f(\varphi, \varepsilon)$ SCF. The following section presents the new model, $f(\varphi, \varepsilon)$, which addresses the uniqueness issue of the APE parameter that is responsible for the high degree of uncertainty in $f(\varphi)$.

6. APE-ε spectral correction

In this section, an additional index is added to the APE spectral correction, $f(\varphi)$. The index is the depth of a water absorption band, ε. The purpose of this additional index is to help distinguish observations that have similar APE values but different values of I_{scn}, in particular in the range of $1.85\text{eV} \leq \varphi \leq 1.90\text{eV}$ where this phenomenon is prevalent. Based on the analysis in Section 5.1, the following four wavebands are used: 650–670 nm, 710–730 nm, 810–830 nm, and 930–950 nm. For each combination φ-ε, the optimal surface fitting function is determined for further analysis.

6.1. Parameterisation of the I_{scn} -φ-ε correlation

For each waveband, the 22 default surface fitting functions available in OriginLab [80] were fit to the data. Typically, around eight fits converged for each waveband. The highest ranking fit for each waveband, according to the BIC and AIC values, was selected for further analysis. These fits for each waveband are summarised in Tables 6, 7, and 8 for the aSi-T, CdTe, and mSi modules, respectively. The extent to which each surface function explains the variability in I_{scn} is characterised by the coefficient of determination, R^2, the values of which are also included in the aforementioned tables.

Table 6
R^2 values for the highest ranking surface fit to the data for each of the four wavebands tested for the aSi-T device.

ε	R^2	Function
650–670	0.915	LogNormal2D
710–730	0.913	ExtremeCum
810–830	0.912	ExtremeCum
930–950	0.913	ExtremeCum

Table 7
R^2 values for the highest ranking surface fit to the data for each of the four wavebands tested for the CdTe device.

ε	R^2	Function
650–670	0.584	Poly2D
710–730	0.604	RationalTaylor
810–830	0.571	Parabola2D
930–950	0.596	DoseResp2D

Table 8
R^2 values for the highest ranking surface fit to the data for each of the four wavebands tested for the mSi device.

ε	R^2	Function
650–670	0.295	RationalTaylor
710–730	0.242	ExtremeCum
810–830	0.237	LogNormal2D
930–950	0.269	LogNormal2D

For all three devices, an improvement in fitting accuracy is achieved relative to $f(\varphi)$ model regardless of the waveband chosen for ε. The improvement in the SCF achieved through the inclusion of ε is not only apparent in the R^2 value, but also visually evident in the correlations plotted for each device in Figs. 9a, 9b, and 9c. In these three graphs, the correlations using the 650–670 nm spectral band are presented as an example, although the described effect is present for all spectral bands.

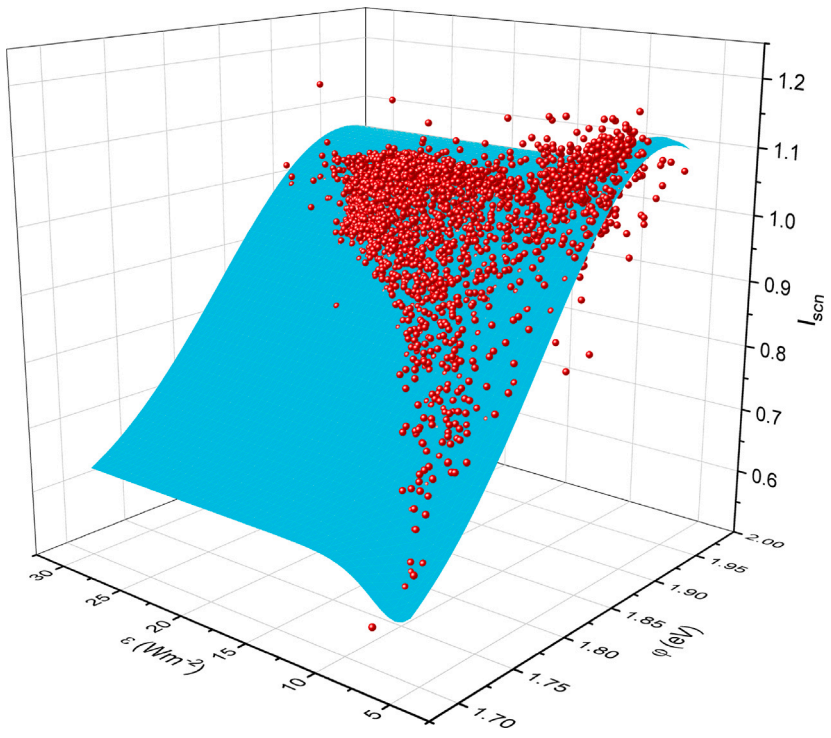

Fig. 9a. $I_{scn} = f(\varphi, \varepsilon)$ for the aSi-T device, where ε is set as the 650–670 nm band.

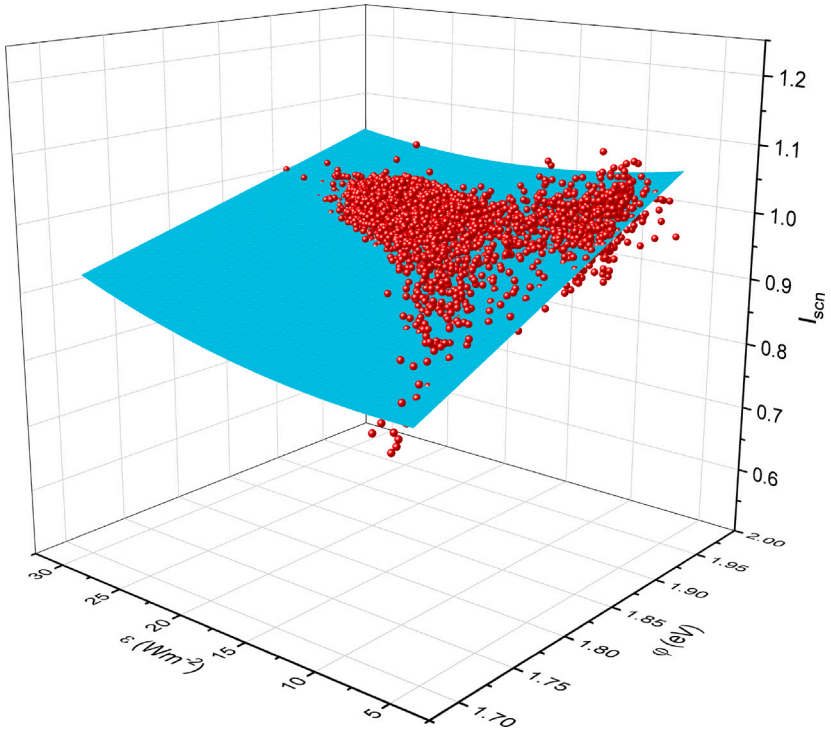

Fig. 9b. $I_{scn} = f(\varphi, \varepsilon)$ for the CdTe device, where ε is set as the 650–670 nm band.

The grouping of observations around the 1.85eV $\leq \varphi \leq$ 1.90eV range for all PV devices, which was discussed in Section 5.1, is spread out into the third dimension created by the additional index. As a result, different values of I_{scn} that were originally attributed to the same or similar values of APE are now attributed to unique combinations of APE and ε.

For the aSi-T device, the improvement in R^2 (Table 4 vs. 6) is relatively small (<0.02) and the fitting accuracy of all wavebands is similar. In contrast, the differences between the R^2 values for the different wavebands is more distinct for the CdTe device, as shown in Table 7. The range in R^2 is around 0.04 ($0.56 < R^2 < 0.60$), which is greater than the range for the aSi-T device. This may be due

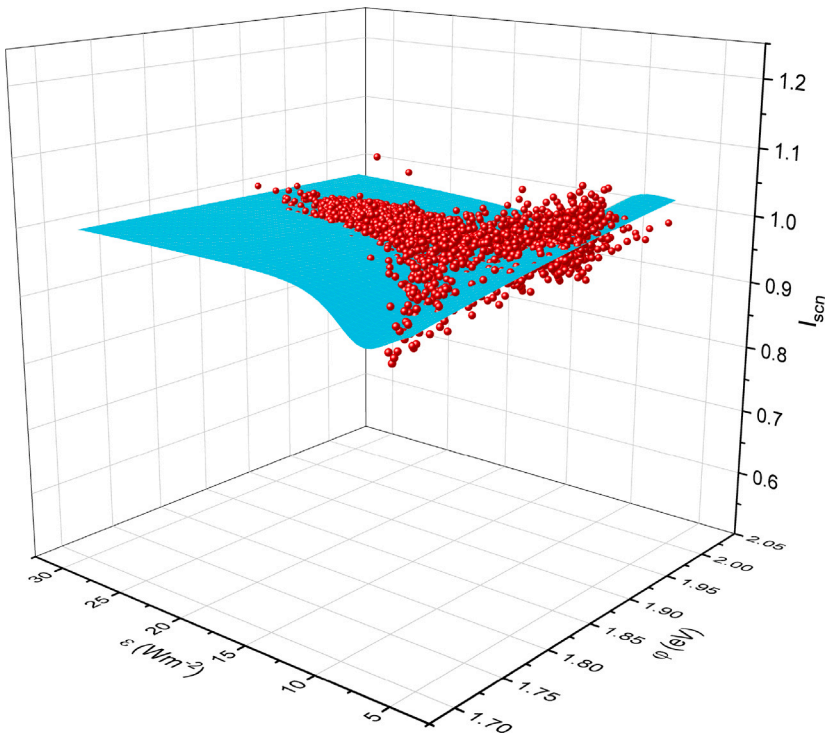

Fig. 9c. $I_{scn} = f(\varphi, \varepsilon)$ for the mSi device, where ε is set as the 650–670 nm band.

to the higher susceptibility of CdTe performance to changes in atmospheric water vapour content, hence $f(\varphi, \varepsilon)$ could be more sensitive to the water absorption band (ε) selection. Furthermore, the relative improvement of the spectral correction function through the inclusion of the additional index is higher for CdTe than for the aSi-T device. Comparing $f(\varphi)$ for both devices (Figs. 8b and 8c), the variation in I_{scn} for similar values of APE is significantly greater for the CdTe module than for the aSi-T device, in particular between $1.85\text{eV} \leq \varphi \leq 1.90\text{eV}$. This was to be expected given the wider spectral response of the CdTe module, compared with a single-junction aSi module (discussed in [25]), and the issues discussed earlier regarding the uncertainty in APE values for analysing spectra at longer wavelengths. Therefore, the inclusion of an additional index that helps to distinguish between different performance observations that are matched to the same APE would be expected to benefit the CdTe model more than the aSi-T model.

For the mSi module, the variation in I_{scn} about the 1.85–1.90 eV band reaches up to around ±10%, which is even greater than that of the CdTe module. Hence, it would be expected that the attribution of these different performance observations to uniquely characterised spectra, as opposed to similarly characterised spectra as in Fig. 8d, would increase the model accuracy significantly. The additional index increases the value of R^2 for the spectral correction function by around 300% (a factor of three). All R^2 values for the mSi module are significantly lower than those resulting from the fits for the other PV devices, but this is to be expected given the flatter spectral response of mSi device and its greater dependence on APE values that have a higher uncertainty, compared with the aSi-T and CdTe modules, as discussed earlier. What is more important than the absolute value of the R^2 coefficient is its relative increase through the inclusion of an additional index in the spectral correction model.

The variation in the highest ranking spectral band for each device, according to the R^2, means that no single spectral band can be declared optimal for all devices at this stage of the analysis. The 650–670 nm ranks highest for the aSi-T and mSi devices, although by only a small margin for the former. For the CdTe module, the 650–670 nm band ranks third highest. Moreover, R^2 is only an indicator of how well the surface function fits to the plotted data points, but provides no information on whether the resulting model is a good predictor of the dependent variable when supplied with an arbitrary set of data as an input. A high R^2 value could result from overfitting to the model development dataset and the derived model may not provide accurate predictions when supplied with new data. Therefore, in the following section, the predictive accuracies of the SCFs derived for all four ε bands are analysed to identify the optimal waveband, ε.

6.2. Optimisation of ε

The highest ranking parameterisations for each ε band are used to predict I_{scn} values for each DUT using the validation dataset described in Section 2.5. A simple calculation of the annual Mean Absolute Error (MAE) is used here to compare the prediction accuracies of each SCF for each PV device. The MAE is calculated as follows:

$$\text{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}, \tag{6}$$

where y_i is the predicted value, x_i is the true value, and n is the total number of data points. The MAE in this case is dimensionless as x and y are the true and predicted values of I_{scn}, respectively, which themselves are dimensionless. First, as an overall indication of accuracy, the annual mean MAE is examined. A more detailed analysis of the MAE on different timescales throughout the year is presented in Section 7, where the proposed SCF is validated.

Table 9 shows the MAE values calculated using the best fit function, as identified in Section 6.1, for the aSi-T, CdTe, and mSi devices. The results in Table 9 show that the wavebands that yielded SCFs with the highest R^2 value do not necessarily result in the lowest MAE. This may be a sign of overfitting in some of the models, as postulated in Sections 2.5 and 6.1. The 650–670 nm waveband consistently results in the lowest MAE for each of the DUTs. However, for aSi-T device, the 810–830 and 930–950 wavebands result in the same MAE as the 650–670 waveband. For the CdTe device, the 710–730 waveband results in the same MAE as the 650–670 waveband. Despite these similarities, given that the 650–670 waveband is the most consistent across all

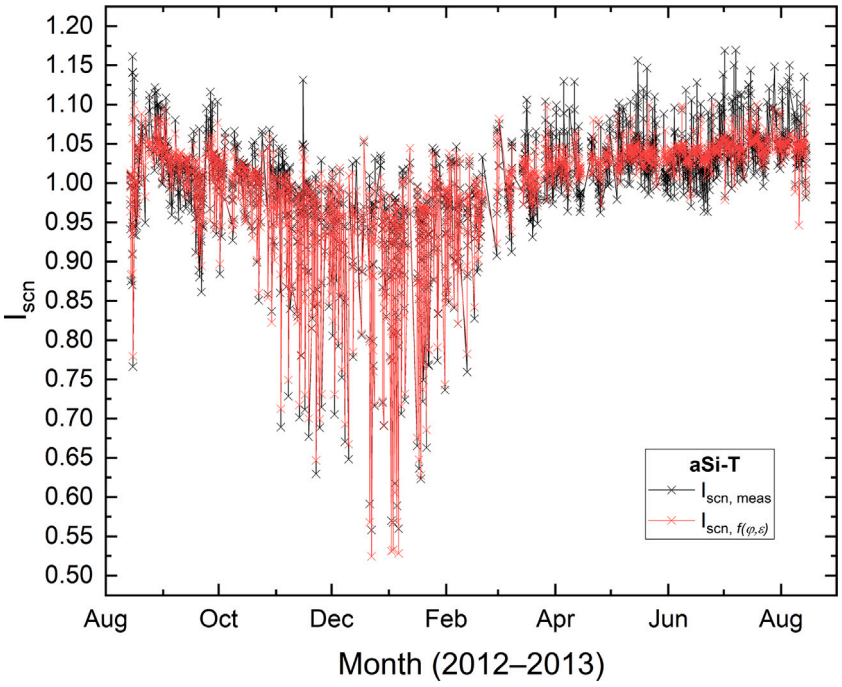

Fig. 10a. Comparison of the measured and predicted values of I_{scn} (using $f(\varphi,\varepsilon)$) as a function of time for the aSi-T device.

Table 9
MAE values calculated from predictions of I_{scn} that result from using different spectral bands in the $f(\varphi,\varepsilon)$ SCF.

Device	ε	MAE
aSi-T	650–670	0.0134
	710–730	0.0156
	810–830	0.0134
	930–950	0.0134
CdTe	650–670	0.0149
	710–730	0.0149
	810–830	0.0152
	930–950	0.0150
mSi	650–670	0.0101
	710–730	0.0107
	810–830	0.0107
	930–950	0.0106

three devices, the 650–670 band is considered to be the best choice for further analysis.

The fact that the same waveband gives the lowest prediction error for all three devices suggests that the optimum band selection to characterise the solar spectrum may depend on a fundamental property of the spectrum rather than a device-specific response to the spectrum. None of the three panels have any particular spectral response characteristics between 650–670 nm. Taking the CdTe device as an example, Fig. 7 shows that its peak spectral response is at around 900 nm. therefore, one may have expected that a water absorption band closer to this peak response wavelength, such as 930–950 nm, would have been a better choice for ε. Although the waveband 930–950 nm generates a relatively low MAE, it ranks third highest out of the four wavebands.

7. Validation of the APE-ε SCF

The prediction accuracy of the highest ranking surface fits for each PV device are tested in this section in order to validate each model. Following on from the analysis in the previous sections, ε is set as 650–670 nm. The prediction accuracy of the proposed model is determined by comparing values of I_{scn} predicted by $f(\varphi,\varepsilon)$, $I_{scn,f(\varphi)}$, with values of I_{scn} derived from measured data, $I_{scn,meas}$. The prediction accuracy of the proposed model is then compared to that of the single-variable APE spectral correction and a traditional proxy-variable approach based on air mass and atmospheric precipitable water content.

7.1. Time series analysis

In this section, the temporal trends in the prediction accuracy and associated uncertainty of the proposed model are analysed. Figs. 10a, 10b, and 10c show time series plots of $I_{scn,calc}$ and $I_{scn,meas}$ for the aSi-T, CdTe, and mSi devices, respectively. For these figures, $I_{scn,calc}$ is calculated using $f(\varphi,\varepsilon)$ and is hence denoted $I_{scn,f(\varphi,\varepsilon)}$.

The seasonal variation in $I_{scn,meas}$ for the aSi-T and CdTe devices, identified in Section 4 is predicted accurately by the proposed model. The mean absolute errors (MAEs) for the winter months of October–February, during which time this drop in efficiency occurs, are 0.00967 and 0.01262 for the aSi-T and CdTe modules, respectively. $I_{scn,meas}$ for the mSi module exhibits a weaker seasonal trend, but this is still captured in the predictions by the model, which yields an MAE of 0.00936.

In addition to being able to model the long-term seasonal shift in efficiency due to the spectrum, high frequency changes in $I_{scn,meas}$ are also modelled accurately. In particular for the aSi-T device, high frequency fluctuations in $I_{scn,meas}$ above and below unity are captured by the model. The same is true for the mSi and CdTe devices, albeit to a lesser degree for the extreme values of $I_{scn,meas}$.

The prediction error increases for all three devices in the summer months. Whereas the model predicts a relatively stable efficiency in these months with only minor variations $\mathcal{O}(3\,\%)$, in reality $I_{scn,meas}$ exhibits variation $\mathcal{O}(10\,\%)$. Figs. 11a, 11b, and 11c show the MAE for each 15-minute measurement and prediction, as well as the mean monthly MAE. Prediction errors in the summer months are the dominant contributor to the annual MAE values, which are 0.01018, 0.01343, and 0.01495 for the mSi, aSi, and CdTe devices, respectively. One cause of this may lie in the methodology used for the normalisation of I_{sc}. The influence of two factors – irradiance and temperature – were removed from the measured short-circuit current in the normalisation process, and it was assumed the resulting difference of I_{sc}^{*} from the reference current was due to the influence of the spectrum. In the

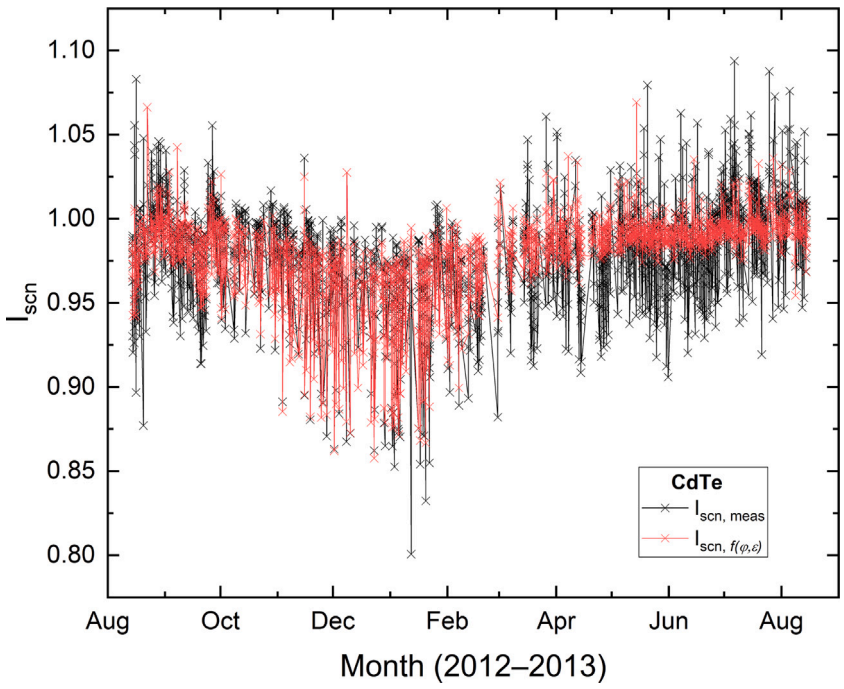

Fig. 10b. Comparison of the measured and predicted values of I_{scn} (using $f(\varphi, \varepsilon)$) as a function of time for the CdTe device.

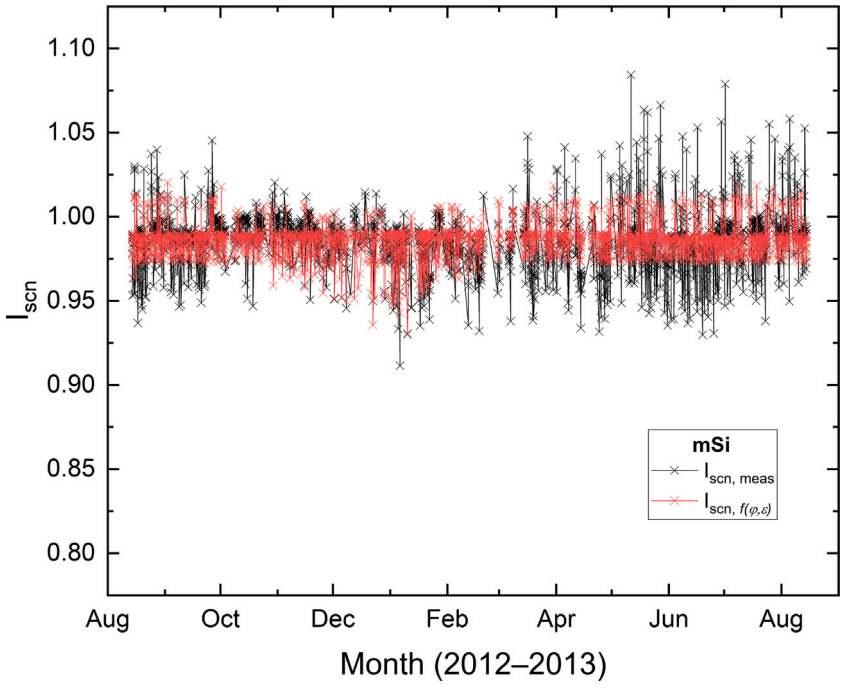

Fig. 10c. Comparison of the measured and predicted values of I_{scn} (using $f(\varphi, \varepsilon)$) as a function of time for the mSi device.

summer, the relative contribution of the direct beam component of irradiance is greater, and thus the angle of incidence (AOI) has a greater effect on module efficiency [12]. However, changes in efficiency due to the AOI have not been considered because to do so would require information on the separate components of irradiance rather than just the global plane of array irradiance. Due to construction work around the meteorological weather station [34] in the summer of 2013, the measurements of the separate components of irradiance suffer from a relatively large number of missing values. Using these data would have impacted the reliability and accuracy of the derived model. Since the primary aim of this study is to present a new method to account for the spectral influence on the performance of PV devices, and demonstrate the relative power of an additional index in the single-variable APE spectral correction method, it was considered more appropriate to use the larger G_{poa}dataset. The increase in prediction error resulting from the exclusion of an AOI correction for the summer months of 2013 is a systematic error present across all models for each device and does not impact the final conclusions of the study in terms of the relative

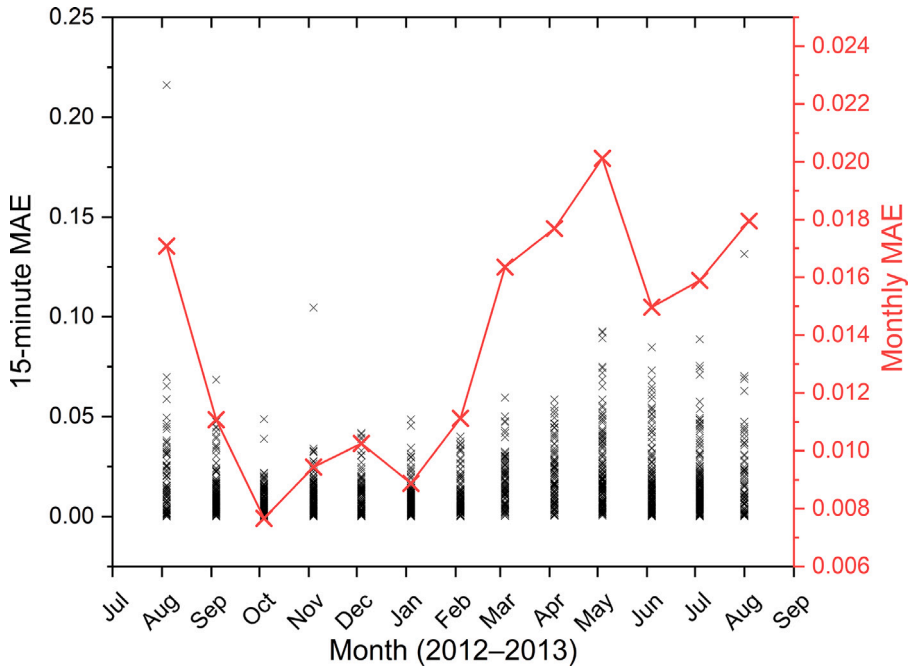

Fig. 11a. 15-minute (crosses) and monthly MAE (line and crosses) values as a function of time for the aSi-T device. The annual MAE value is 0.0134.

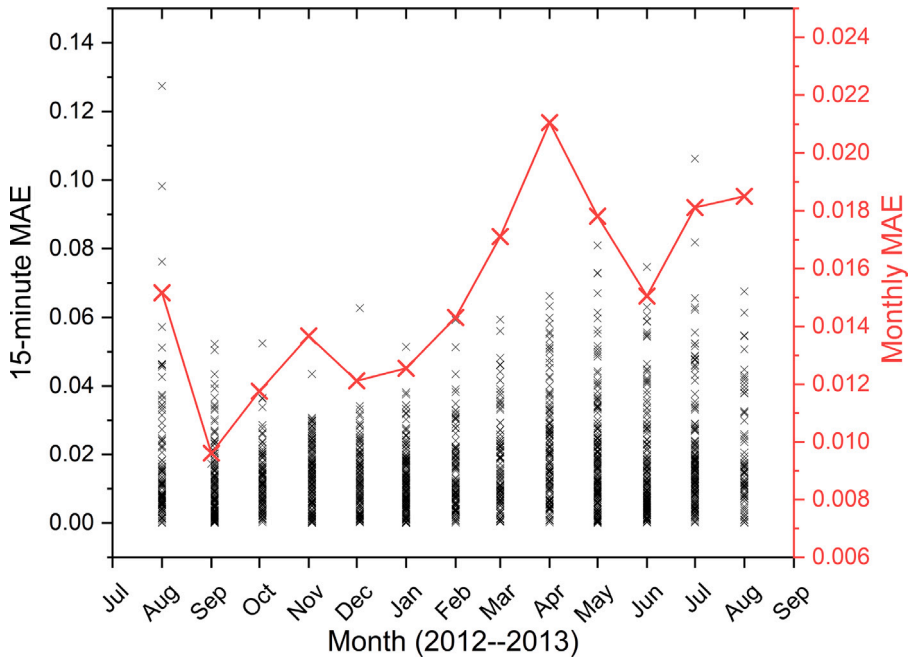

Fig. 11b. 15-minute (crosses) and monthly MAE (line and crosses) values as a function of time for the CdTe device. The annual MAE value is 0.0149.

performance of the different models and the parameters they include. There is another spike in the MAE for all models in the month of April 2013. This is likely to be a result of increased atmospheric aerosol levels during this month, as shown in Fig. 5.

7.2. Comparison with existing models

In this section, the prediction accuracy of the proposed model, $f(\varphi, \varepsilon)$, is compared with that of the APE model, $f(\varphi)$, presented in Section 5.1, and a proxy-based alternative — the air mass and precipitable water content model, $f(AM_a, W)$ [46]. The latter is chosen for two reasons. First, as a proxy variable-based approach, it offers an insight into how the two methods – spectra- and proxy-based – compare in terms of accuracy. Second, it is chosen for its similarity in terms of inclusion of the effects of water vapour. The first subsection introduces the $f(AM_a, W)$ model and presents new coefficients for the model that have been derived for the aSi-T device in this study. The following subsection compares the predictive accuracies of the different SCFs.

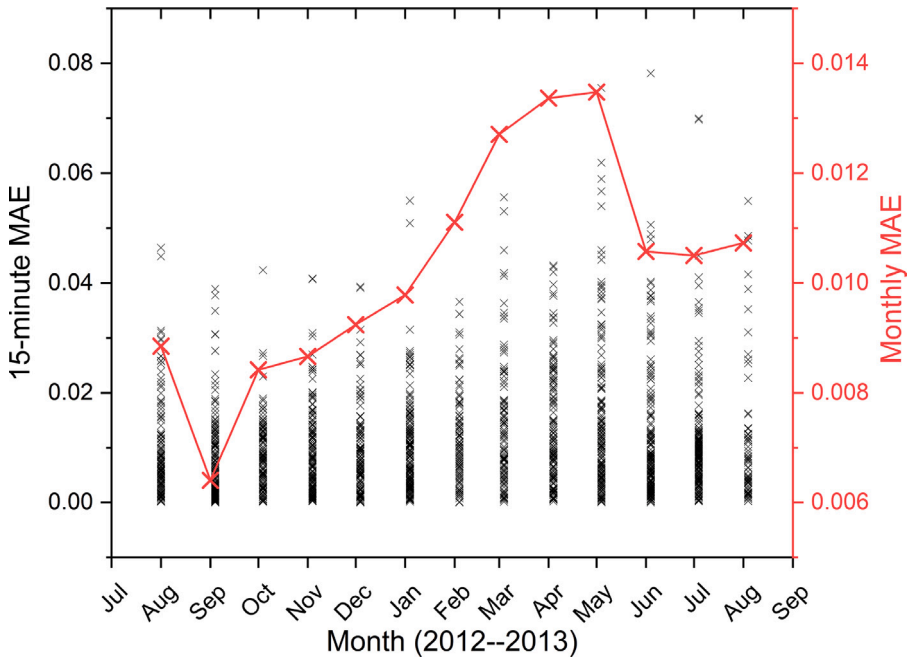

Fig. 11c. 15-minute (crosses) and monthly MAE (line and crosses) values as a function of time for the mSi device. The annual MAE value is 0.0101.

Table 10
$f(AM_a, W)$ model coefficients for the aSi-T PV module.

	b0	b1	b2	b3	b4	b5
aSi-T	0.928	−0.103	−0.0597	0.0939	0.166	0.00656

7.2.1. Air mass and precipitable water SCF

The functional form of $f(AM_a, W)$ is as follows [46]:

$$I_{scn} = b_0 + b_1 \cdot AM_a + b_2 \cdot W + b_3 \cdot \sqrt{AM_a} + b_4 \cdot \sqrt{W} + b_5 \cdot \frac{AM_a}{\sqrt{W}}. \quad (7)$$

Coefficients for the $f(AM_a, W)$ model have only been published for crystalline silicon and CdTe PV devices [46]. For the purpose of this study, an additional set of coefficients has been derived for the aSi-T device using the Golden data and these are summarised in Table 10.

7.3. Performance comparison

Time series plots of the same format as those which were presented in Section 7.1 are presented here for each device, where $I_{scn,calc} = f(AM_a, W)$. Figs. 12a and 12b show that $f(AM_a, W)$ can accurately predict the seasonal shift in I_{scn} for the CdTe and aSi-T devices. However, in the case of mSi device, although a decrease in I_{scn} is predicted in the winter months, this decrease is significantly overestimated. The mSi MAEs for the December, January, and February are 0.0151, 0.0144, and 0.0154, respectively.

A common observation across all three PV devices is that whereas $f(AM_a, W)$ is capable of modelling higher frequency variations in I_{scn} than the simple air mass function, with reference to the results reported in [25], the extreme values of I_{scn} are not accurately predicted. Figs. 12a, 12b, and 12c show that although $f(AM_a, W)$ tends to estimate a change in I_{scn} correctly, the absolute prediction value is typically either an over– or underestimate. The reason for this may be the fact that W is a relatively simplistic indicator of atmospheric water vapour, without specific reference to its impact on the spectrum. Therefore, W may not be sufficiently sensitive to changes in the spectrum at specific wavelengths, caused by the presence of atmospheric water vapour, that are most significant for PV performance. On the other hand, in the proposed $f(\varphi, \varepsilon)$ SCF, the focus of ε on a specific water absorption band enables greater sensitivity to water vapour-induced changes in the spectrum that are the most influential on PV performance. This notion is supported by the annual values of MAE for all three devices, which are summarised to three significant figures in Table 11 for $f(AM_a, W)$, $f(\varphi)$, and $f(\varphi, \varepsilon)$. The proposed model reduces the MAE generated by $f(AM_a, W)$ by approximately 60% for both the mSi and aSi-T modules, and by around 20% for the CdTe module.

Table 11
Annual MAE values for the predictions made by $f(AM_a, W)$ and $f(\varphi, \varepsilon)$ for each DUT.

DUT	MAE		
	$f(\varphi)$	$f(AM_a, W)$	$f(\varphi, \varepsilon)$
mSi	0.0112	0.0250	0.0102
CdTe	0.0152	0.0187	0.0149
aSi-T	0.0149	0.0224	0.0134

For all three PV devices, the new model also reduces the prediction error compared with $f(\varphi)$, albeit by a smaller margin than the improvements on $f(AM_a, W)$. Although the absolute value of the reduction in MAE appears small, the percentage change is around 10% for the mSi and aSi-T devices, and 2% for the CdTe device, which is significant.

While it is not the main focus of this study, it is still useful to compare the $f(\varphi)$ and $f(AM_a, W)$ SCFs. The percentage improvements achieved through the inclusion of ε in the SCF for the aSi-T, CdTe, and mSi modules are around 10%, 2%, and 9%. It is worth bearing in mind that the additional information required for the new model, ε, is already present in data used to calculate φ. Therefore, there is no information cost associated with these improvements since no additional data are required. The only requirement is an extra calculation of the value for ε in the model, but this is relatively simple and does not add any significant computational cost to the overall modelling procedure. On the other hand, the parameterisation of $f(\varphi, \varepsilon)$ is somewhat more complex and if the user is working with a new or customised DUT, greater computational cost would be incurred to derive the model coefficients for this multivariable function. However, in the following section, it is shown that a balance between computational cost and accuracy can easily be struck. Even without the optimal parameterisation of $f(\varphi, \varepsilon)$, where a sub-optimal but simpler parameterisation is used

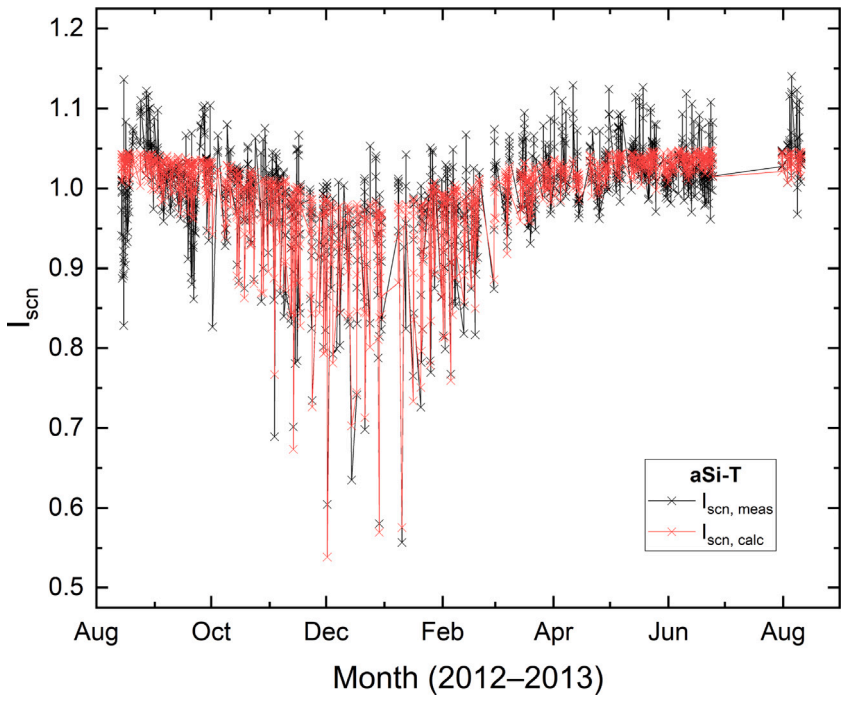

Fig. 12a. Comparison of the measured and predicted values of $I_{scn} = f(AM_a, W)$ as a function of time for the aSi-T device.

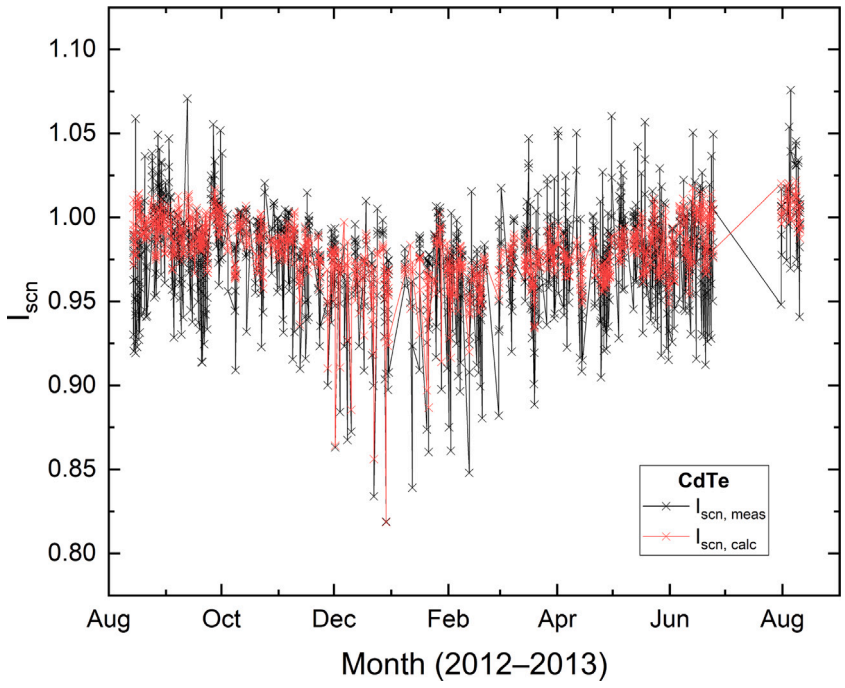

Fig. 12b. Comparison of the measured and predicted values of $I_{scn} = f(AM_a, W)$ as a function of time for the CdTe device.

instead, significant improvements in accuracy for all PV devices are still be achieved relative to existing SCFs.

7.4. Balancing model complexity and accuracy

It is clear that the proposed spectral correction based on the average photon energy and the depth of a water absorption band exceeds the accuracy of existing spectral corrections significantly. However, the analysis thus far finds that different functional forms of $f(\varphi, \varepsilon)$ offer better information criteria scores for different devices. From a computational complexity perspective, the functions are relatively simple and can easily be integrated into commercial software or other PV performance applications. Nevertheless, a simpler expression could save time for large-scale iterative calculations that may be required when analysing multiple modelling scenarios. Furthermore, from an end-user application perspective, it would be simpler to have a single

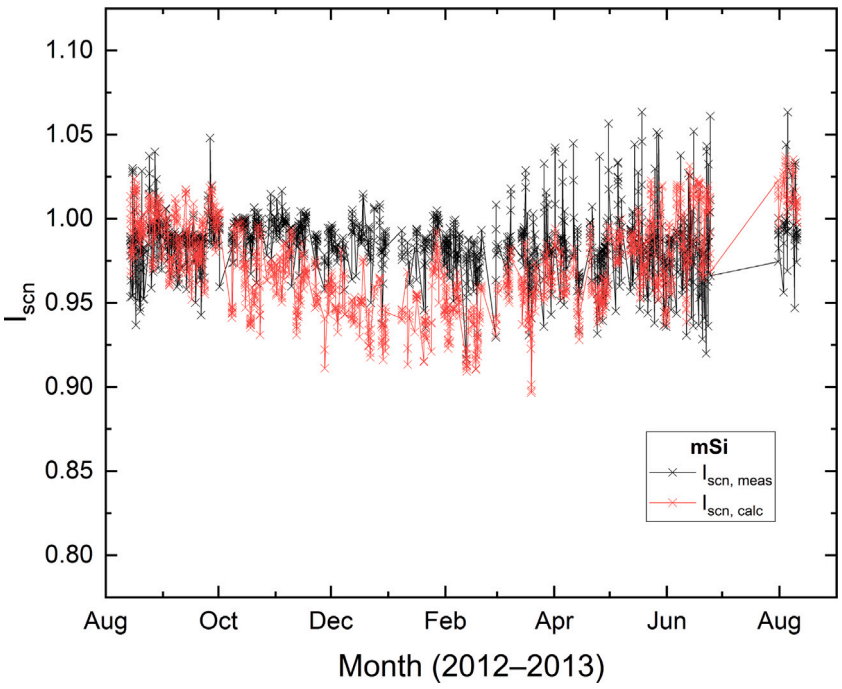

Fig. 12c. Comparison of the measured and predicted values of $I_{scn} = f(AM_a, W)$ as a function of time for the mSi device.

Table 12
Poly2D model coefficients for each DUT to four significant figures. R^2 values for the corresponding surface fitting functions are included to three significant figures.

DUT	R^2	Model coefficients					
		z_0	a	b	c	d	g
aSi-T	0.907	−21.94	22.62	−0.01393	−5.521	1.7341E−4	0.003860
CdTe	0.584	−0.5313	0.7208	0.02232	0.05321	1.629E−4	−0.01445
mSi	0.194	−0.3998	1.101	0.03366	−0.1837	1.493E−4	−0.02046

Table 13
Annual MAE values for the predictions made by $f(\varphi, \varepsilon)$ for each DUT. ε is set as the 650–670 nm spectral band. The model used is described in Eq. (8) and the model coefficients for each DUT are summarised in Table 12.

DUT	MAE
mSi	0.0101
CdTe	0.0150
aSi-T	0.0140

functional form that only has device-specific coefficients, rather than a device-specific functional form with device-specific coefficients. It is also easier and more efficient to record and share the model if its form is standardised. In this section, a single functional form is chosen to demonstrate how a high level of accuracy is still maintained even if the optimal functional expression, according to the statistical ranking, is not adopted for $f(\varphi, \varepsilon)$.

The simplest of the functions tested in Section 6.1 is the "Poly2D" equation, which takes the following form [81]:

$$z = z_0 + ax + by + cx^2 + dy^2 + gxy. \tag{8}$$

The model coefficients (z_0, a, b, c, d, g) for each DUT are summarised in Table 12. The resulting MAE values for the comparison between $I_{scn,calc}$ and $I_{scn,meas}$ are summarised in Table 12.

The Poly2D model is used to predict values of I_{scn} in the same way as that which was presented in Section 7.3. The resulting annual MAE values for each DUT are summarised in Table 13

The results in this section show that the proposed methodology is relatively flexible in terms of its functional form. It is also clear that the selection of the optimal functional form is a topic for further exploration to find a balance between not only model complexity and accuracy, but also usability. The two-dimensional polynomial function is a simple surface expression and, although it does not offer the highest level of accuracy possible for $f(\varphi, \varepsilon)$, it can still be used to derive an SCF for all three DUTs that improves the MAE value compared with $f(\varphi)$ and $f(AM_a, W)$.

8. Conclusion

Changes in spectral irradiance can lead to variations in PV performance of up to 9%, 5%, and 2% on a monthly average timescale for the aSi-T, CdTe, and mSi devices, respectively. These values increase to 40%, 20%, and 5% on a 15-minute average timescale. A single-variable spectral correction function based on the average photon energy can account for most of this variation, but suffers from increased uncertainty for APE values between 1.85 eV and 1.90 eV. In this range of only 0.05 eV (±2.5%), the same or similar values of APE are correlated with a wide range (±10%) of I_{scn} values.

Including an additional parameter in the spectral correction, namely the depth of a water absorption band (ε) enables an association of the different I_{scn} values, which have the same APE value, to unique combinations of APE and ε instead. This work identifies the 650–670 nm spectral band to be optimal for all three devices investigated. Improvements in the prediction accuracy (reductions in the mean absolute error of prediction) of up to 10% are achieved with the new spectral correction. Furthermore, against a comparable two-variable proxy spectral correction, the air mass and precipitable water SCF, improvements in the prediction accuracy of up to 60% are achieved.

The optimal parameterisation for the model appears to be device-dependent but, in this work, it is shown that a compromise can be made between accuracy and complexity by adopting a simple and standard expression for the model for all devices. This expression is a two-dimensional polynomial (Poly2D), the device-specific coefficients for

which are determined in this study. The MAE for all three devices is still either the same or better when using the Poly2D expression rather than the optimal device-specific expression.

This work presents a new spectral correction model that addresses the question of bijectivity with the APE parameter that is used in some existing spectral correction methods. The methodology followed in this work has been demonstrated for three PV devices and the model accuracy has been validated for all three devices. Its benefits with respect to existing approaches have been shown and are manifested in particular through reductions in the normalised short-circuit current prediction error. However, the precise functional form of the model requires further investigation. Although multiple parameterisations have been investigated, further work should continue these investigations in more detail through the use of larger datasets from different climate regions, different statistical analyses, and a wider range of PV devices. The data used in this study are restricted to a single location, hence the presented model coefficients may be limited by site-specific conditions such as the local climate, the specific technical characteristics of the PV devices, local measurement site set up, etc. Therefore, the precise coefficients presented in this study may not be universally applicable. Nevertheless, the standardised comparison between different SCF methodologies within this study shows that the proposed methodology – an SCF based on both φ and ε – addresses the shortfalls of $f(\varphi)$, and is more accurate than a comparable proxy multivariable SCF, namely $f(AM_a, W)$. The proposed method should now be validated in different climate regions, for different PV devices, and so on.

CRediT authorship contribution statement

Rajiv Daxini: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Robin Wilson:** Writing – review & editing, Supervision. **Yupeng Wu:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Yupeng Wu reports financial support was provided by Engineering and Physical Sciences Research Council.

Data availability

I have provided links and references to the data used in this work. The data have been made available to the public

Acknowledgements

This work was supported by the Faculty of Engineering, University of Nottingham, United Kingdom, through a Ph.D. studentship awarded to Rajiv Daxini. The authors also appreciate the contribution of the team at the National Renewable Energy Laboratory (NREL), USA, for their publication of high quality photovoltaic performance and meteorological data in the public domain.

Appendix

The functional forms of the six parameterisations used in this study, which are referenced in Tables 6, 7, and 8, are summarised this appendix. Further information on the fitting functions listed here, including the meanings behind each of the coefficients in the models, can be found in Ref. [82].

1. LogNormal2D :

$$z(x,y) = z_0 + B\exp\left\{-\frac{\left(\ln\frac{x}{C}\right)^2}{2D^2}\right\} + E\exp\left\{-\frac{\left(\ln\frac{y}{F}\right)^2}{2G^2}\right\} + H\exp\left\{-\frac{\left(\ln\frac{x}{C}\right)^2}{2D^2} - \frac{\left(\ln\frac{y}{F}\right)^2}{2G^2}\right\} \quad (A.1)$$

2. ExtremeCum:

$$z(x,y) = z_0 + B\exp\left\{-\exp\left\{\frac{C-x}{D}\right\}\right\} + E\exp\left\{-\exp\left\{\frac{F-y}{G}\right\}\right\} + H\exp\left\{-\exp\left\{\frac{C-x}{D}\right\} - \exp\left\{\frac{F-y}{G}\right\}\right\} \quad (A.2)$$

3. RationalTaylor:

$$z(x,y) = \frac{z_0 + A_{01}x + B_{01}y + B_{02}y^2 + C_{02}xy}{1 + A_1x + B_1y + A_2x^2 + B_2y^2 + C_2xy} \quad (A.3)$$

4. DoseResp2D:

$$z(x,y) = z_0 + \frac{B}{\left[1+\left(\frac{x}{C}\right)^{-D}\right]\left[1+\left(\frac{y}{E}\right)^{-F}\right]} \quad (A.4)$$

5. Parabola2D:

$$z(x,y) = z_0 + ax + by + cx^2 + dy^2 \quad (A.5)$$

6. Poly2D

$$z(x,y) = z_0 + ax + by + cx^2 + dy^2 + gxy \quad (A.6)$$

References

[1] Kalogirou SA. Solar energy engineering: processes and systems. Academic Press; 2013.
[2] Chikate BV, Sadawarte Y, Sewagram B. The factors affecting the performance of solar cell. Int J Comput Appl 2015;1(1):0975–8887.
[3] Seera M, Tan CJ, Chong K-K, Lim CP. Performance analyses of various commercial photovoltaic modules based on local spectral irradiances in Malaysia using genetic algorithm. Energy 2021;223:120009.
[4] Nann S, Emery K. Spectral effects on PV-device rating. Sol Energy Mater Sol Cells 1992;27(3):189–216.
[5] Chivelet N. Analysis of spectral factor of different commercial PV modules based on measured data. In: 14 Th European PV conference. 1997, p. 282–3.
[6] Virtuani A, Fanni L. Seasonal power fluctuations of amorphous silicon thin-film solar modules: Distinguishing between different contributions. Prog Photovolt, Res Appl 2014;22(2):208–17.
[7] Dirnberger D, Blackburn G, Müller B, Reise C. On the impact of solar spectral irradiance on the yield of different PV technologies. Sol Energy Mater Sol Cells 2015;132:431–42.
[8] Gottschalg R, Betts T, Infield D, Kearney M. On the importance of considering the incident spectrum when measuring the outdoor performance of amorphous silicon photovoltaic devices. Meas Sci Technol 2004;15(2):460.
[9] Lindsay N, Libois Q, Badosa J, Migan-Dubois A, Bourdin V. Errors in PV power modelling due to the lack of spectral and angular details of solar irradiance inputs. Sol Energy 2020;197:266–78.
[10] Hirata Y, Tani T. Output variation of photovoltaic modules with environmental factors—I. The effect of spectral solar radiation on photovoltaic module output. Sol Energy 1995;55(6):463–8.
[11] Fernández EF, Soria-Moya A, Almonacid F, Aguilera J. Comparative assessment of the spectral impact on the energy yield of high concentrator and conventional photovoltaic technology. Sol Energy Mater Sol Cells 2016;147:185–97.
[12] King DL, Kratochvil JA, Boyson WE. Measuring solar spectral and angle-of-incidence effects on photovoltaic modules and solar irradiance sensors. In: Conference record of the twenty sixth IEEE photovoltaic specialists conference-1997. IEEE; 1997, p. 1113–6.
[13] Gottschalg R, Infield D, Kearney M. Experimental study of variations of the solar spectrum of relevance to thin film solar cells. Solar Energy Mater Solar Cells 2003;79(4):527–37.
[14] Lee M, Panchula AF. Variation in spectral correction of PV module performance based on different precipitable water estimates. In: 2016 IEEE 43rd photovoltaic specialists conference. IEEE; 2016, p. 2692–7.

[15] Passow K, Lee M. Effect of spectral shift on solar PV performance. In: 2016 IEEE conference on technologies for sustainability. IEEE; 2016, p. 246–50.
[16] Jardine CN, Betts T, Gottschalg R, Infield D, Lane K. Influence of spectral effects on the performance of multijunction amorphous silicon cells. In: Proc. photovoltaic in Europe conference. 2002, p. 1756–9.
[17] Moreno-Sáez R, Mora-López L. Modelling the distribution of solar spectral irradiance using data mining techniques. Environ Model Softw 2014;53:163–72.
[18] Rodziewicz T, Rajfur M. Numerical procedures and their practical application in PV modules' analyses. Part II: Useful fractions and APE. Opto-Electron Rev 2019;27(2):149–60.
[19] Piliougine M, Elizondo D, Mora-López L, Sidrach-de Cardona M. Multilayer perceptron applied to the estimation of the influence of the solar spectral distribution on thin-film photovoltaic modules. Appl Energy 2013;112:610–7.
[20] Conde LA, Angulo JR, Sevillano-Bendezú MÁ, Nofuentes G, Töfflinger JA, de la Casa J. Spectral effects on the energy yield of various photovoltaic technologies in Lima (Peru). Energy 2021;223:120034.
[21] Chantana J, Imai Y, Kawano Y, Hishikawa Y, Nishioka K, Minemoto T. Impact of average photon energy on spectral gain and loss of various-type PV technologies at different locations. Renew Energy 2020;145:1317–24.
[22] Cornaro C, Andreotti A. Influence of Average Photon Energy index on solar irradiance characteristics and outdoor performance of photovoltaic modules. Prog Photovolt, Res Appl 2013;21(5):996–1003.
[23] Gottschalg R, Betts T, Infield D, Kearney M. The effect of spectral variations on the performance parameters of single and double junction amorphous silicon solar cells. Solar Energy Mater Solar Cells 2005;85(3):415–28.
[24] Sirisamphanwong C, Ketjoy N, Sirisamphanwong C. The effect of average photon energy and module temperature on performance of photovoltaic module under thailand's climate condition. Energy Procedia 2014;56:359–66.
[25] Daxini R, Sun Y, Wilson R, Wu Y. Direct spectral distribution characterisation using the average photon energy for improved photovoltaic performance modelling. Renew Energy 2022.
[26] IEC. Photovoltaic devices: Solar simulator performance requirements. IEC 60904-9, IEC Central Office Geneva; 2007.
[27] Minemoto T, Nakada Y, Takahashi H, Takakura H. Uniqueness verification of solar spectrum index of average photon energy for evaluating outdoor performance of photovoltaic modules. Sol Energy 2009;83(8):1294–9.
[28] Tsuji M, Rahman MM, Hishikawa Y, Nishioka K, Minemoto T. Uniqueness verification of solar spectrum obtained from three sites in Japan based on similar index of average photon energy. Sol Energy 2018;173:89–96.
[29] Polo J, Alonso-Abella M, Ruiz-Arias JA, Balenzategui JL. Worldwide analysis of spectral factors for seven photovoltaic technologies. Sol Energy 2017;142:194–203.
[30] Nofuentes G, Gueymard C, Aguilera J, Pérez-Godoy M, Charte F. Is the average photon energy a unique characteristic of the spectral distribution of global irradiance? Sol Energy 2017;149:32–43.
[31] Ishii T, Otani K, Itagaki A, Utsunomiya K. A methodology for estimating the effect of solar spectrum on photovoltaic module performance by using average photon energy and a water absorption band. Japan J Appl Phys 2012;51(10S):10NF05.
[32] King DL, Kratochvil JA, Boyson WE. Photovoltaic array performance model. United States. Department of Energy; 2004.
[33] Stoffel T, Andreas A. NREL solar radiation research laboratory (SRRL): Baseline Measurement System (BMS); Golden, Colorado (data). Tech. rep., National Renewable Energy Lab (NREL), Golden, CO (United States); 1981.
[34] Marion W, Anderberg A, Deline C, Glick S, Muller M, Perrin G, et al. User's manual for data for validating models for PV module performance. Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States); 2014.
[35] Marion B, Anderberg A, Deline C, Muller M, Perrin G, Rodriguez J, et al. Data for validating models for PV module performance. Tech. rep., EMN-DURMAT (EMN-DuraMAT); National Renewable Energy Lab.(NREL), Golden, CO; 2021.
[36] Eke R, Senturk A. Monitoring the performance of single and triple junction amorphous silicon modules in two Building Integrated Photovoltaic (BIPV) installations. Appl Energy 2013;109:154–62.
[37] Qiu C, Yi YK, Wang M, Yang H. Coupling an artificial neuron network daylighting model and building energy simulation for vacuum photovoltaic glazing. Appl Energy 2020;263:114624.
[38] Liao W, Xu S. Energy performance comparison among see-through amorphous-silicon PV (photovoltaic) glazings and traditional glazings under different architectural conditions in China. Energy 2015;83:267–75.
[39] Cheng Y, Gao M, Dong J, Jia J, Zhao X, Li G. Investigation on the daylight and overall energy performance of semi-transparent photovoltaic facades in cold climatic regions of China. Appl Energy 2018;232:517–26.
[40] Uddin MM, Wang C, Zhang C, Ji J. Investigating the energy-saving performance of a CdTe-based semi-transparent photovoltaic combined hybrid vacuum glazing window system. Energy 2022;253:124019.
[41] Li J, Ren X, Yuan W, Li Z, Pei G, Su Y, et al. Experimental study on a novel photovoltaic thermal system using amorphous silicon cells deposited on stainless steel. Energy 2018;159:786–98.
[42] Ren X, Li J, Hu M, Pei G, Jiao D, Zhao X, et al. Feasibility of an innovative amorphous silicon photovoltaic/thermal system for medium temperature applications. Appl Energy 2019;252:113427.
[43] Lorenzi B, Mariani P, Reale A, Di Carlo A, Chen G, Narducci D. Practical development of efficient thermoelectric–Photovoltaic hybrid systems based on wide-gap solar cells. Appl Energy 2021;300:117343.
[44] Yang Y, Yu A, Hsu B, Hsu W, Yang A, Lan C. Development of high-performance multicrystalline silicon for photovoltaic industry. Prog Photovolt, Res Appl 2015;23(3):340–51.
[45] Meral ME, Dincer F. A review of the factors affecting operation and efficiency of photovoltaic based electricity generation systems. Renew Sustain Energy Rev 2011;15(5):2176–84.
[46] Lee M, Panchula A. Spectral correction for photovoltaic module performance based on air mass and precipitable water. In: 2016 IEEE 43rd photovoltaic specialists conference. IEEE; 2016, p. 1351–6.
[47] Batzelis E. Non-iterative methods for the extraction of the single-diode model parameters of photovoltaic modules: A review and comparative assessment. Energies 2019;12(3):358.
[48] Hao P, Zhang Y, Lu H, Lang Z. A novel method for parameter identification and performance estimation of PV module under varying operating conditions. Energy Convers Manage 2021;247:114689.
[49] Gholami A, Ameri M, Zandi M, Ghoachani RG. Electrical, thermal and optical modeling of photovoltaic systems: Step-by-step guide and comparative review study. Sustain Energy Technol Assess 2022;49:101711.
[50] Platon R, Martel J, Woodruff N, Chau TY. Online fault detection in PV systems. IEEE Trans Sustain Energy 2015;6(4):1200–7.
[51] Theocharides S, Makrides G, Venizelou V, Kaimakis P, Georghiou G. PV production forecasting model based on Artificial Neural Networks (ANN). In: 33rd Eur. photovolt. sol. energy conf, no. September. 2017, p. 1830–94.
[52] Theocharides S, Makrides G, Livera A, Theristis M, Kaimakis P, Georghiou GE. Day-ahead photovoltaic power production forecasting methodology based on machine learning and statistical post-processing. Appl Energy 2020;268:115023.
[53] Halabi S, Lin C-Y, Kelly WK, Fizazi KS, Moul JW, Kaplan EB, et al. Updated prognostic model for predicting overall survival in first-line chemotherapy for patients with metastatic castration-resistant prostate cancer. J Clin Oncol 2014;32(7):671.
[54] Lyman GH, Kuderer NM, Crawford J, Wolff DA, Culakova E, Poniewierski MS, et al. Predicting individual risk of neutropenic complications in patients receiving cancer chemotherapy. Cancer 2011;117(9):1917–27.
[55] Malof JM, Bradbury K, Collins LM, Newell RG. Automatic detection of solar photovoltaic arrays in high resolution aerial imagery. Appl Energy 2016;183:229–40.
[56] Kriegeskorte N. Crossvalidation. In: Toga AW, editor. Brain mapping. Waltham: Academic Press; 2015, p. 635–9. http://dx.doi.org/10.1016/B978-0-12-397025-1.00344-4, URL https://www.sciencedirect.com/science/article/pii/B9780123970251003444.
[57] Hawkins DM. The problem of overfitting. J Chem Inf Comput Sci 2004;44(1):1–12.
[58] Cook RD, Weisberg S. Applied regression including computing and graphics. John Wiley & Sons; 2009.
[59] Reitan CH. Surface dew point and water vapor aloft. J Appl Meteorol (1962-1982) 1963;776–9.
[60] Gueymard C. Assessment of the accuracy and computing speed of simplified saturation vapor equations using a new reference dataset. J Appl Meteorol Climatol 1993;32(7):1294–300.
[61] Kasten F, Young AT. Revised optical air mass tables and approximation formula. Appl Optics 1989;28(22):4735–8.
[62] King BH, Hansen CW, Riley D, Robinson CD, Pratt L. Procedure to determine coefficients for the Sandia Array Performance Model (SAPM). Tech. rep., Sandia National Lab.(SNL-NM), Albuquerque, NM (United States); 2016.
[63] Riordan CJ, Hulstrom R, Myers D. Influences of atmospheric conditions and air mass on the ratio of ultraviolet to total solar radiation. Tech. rep., Solar Energy Research Inst., Golden, CO (USA); 1990.
[64] Viswanadham Y. The relationship between total precipitable water and surface dew point. J Appl Meteorol Climatol 1981;20(1):3–8.
[65] NASA Jet Propulsion Laboratory. Understanding the climate: air and water. 2023, Online https://sealevel.jpl.nasa.gov/ocean-observation/understanding-climate/air-and-water/. [Last Accessed 16 February 2023].
[66] Spark W. The weather year round anywhere on earth. 2021.
[67] Nofuentes G, De la Casa J, Torres-Ramírez M, Alonso-Abella M. Solar spectral and module temperature influence on the outdoor performance of thin film PV modules deployed on a sunny inland site. Int J Photoenergy 2013;2013.
[68] Amillo AMG, Huld T, Vourlioti P, Müller R, Norton M. Application of satellite-based spectrally-resolved solar radiation data to PV performance studies. Energies 2015;8(5):3455–88.
[69] Sai H, Matsui T, Matsubara K. Stabilized 14.0%-efficient triple-junction thin-film silicon solar cell. Appl Phys Lett 2016;109(18):183506.
[70] Soga T. Nanostructured materials for solar energy conversion. Elsevier; 2006.
[71] Williams SR, Betts TR, Helf T, Gottschalg R, Beyer H, Infield D. Modelling long-term module performance based on realistic reporting conditions with consideration to spectral effects. In: 3rd world conference onphotovoltaic energy conversion, 2003. proceedings of, vol. 2. IEEE; 2003, p. 1908–11.
[72] Peharz G, Siefer G, Bett A. A simple method for quantifying spectral impacts on multi-junction solar cells. Sol Energy 2009;83(9):1588–98.

[73] King DL, Kratochvil JA, Boyson WE. Stabilization and performance characteristics of commercial amorphous-silicon PV modules. In: Conference record of the twenty-eighth IEEE photovoltaic specialists conference-2000 (Cat. No. 00CH37036). IEEE; 2000, p. 1446–9.
[74] Gottschalg R, Del Cueto J, Betts TR, Williams SR, Infield D. Investigating the seasonal performance of amorphous silicon single-and multi-junction modules. In: 3rd world conference onphotovoltaic energy conversion, 2003. Proceedings of, vol. 2. IEEE; 2003, p. 2078–81.
[75] Ramgolam YK, Soyjaudah KMS. Modelling the impact of spectral irradiance and average photon energy on photocurrent of solar modules. Sol Energy 2018;173:1058–64.
[76] Peters IM, Haohui L, Reindl T, Buonassisi T. Global comparison of the impact of temperature and precipitable water on CdTe and silicon solar cells. In: 2017 IEEE 44th photovoltaic specialist conference. IEEE; 2017, p. 1140–2.
[77] Faine P, Kurtz SR, Riordan C, Olson J. The influence of spectral solar irradiance variations on the performance of selected single-junction and multijunction solar cells. Solar Cells 1991;31(3):259–78.
[78] Guechi A, Chegaar M, Merabet eA. The effect of water vapor on the performance of solar cells. Physics Procedia 2011;21:108–14.
[79] Sierk B, Solomon S, Daniel J, Portmann R, Gutman S, Langford A, et al. Field measurements of water vapor continuum absorption in the visible and near-infrared. J Geophys Res: Atmos 2004;109(D8).
[80] OriginLab Corporation. OriginPro 2020. 2023, Online https://www.originlab.com/. [Accessed 29 December 2022] [Last Accessed 16 February 2023].
[81] OriginLab Corporation. 30.3.16 Poly2D. 2023, Online https://www.originlab.com/doc/Origin-Help/Poly2D-FitFunc. [Accessed 29 December 2022] [Last Accessed 16 February 2023].
[82] OriginLab Corporation. 30.3 Non-linear surface fitting. 2023, Online https://www.originlab.com/doc/Origin-Help/Non-linear-Surface-Fitting-FitFunc. [Accessed 02 June 2023] [Last Accessed 16 February 2023].