IET Journals
The Institution of Engineering and Technology

Compensation of temporal averaging bias in solar irradiance data

Keith Gibson[1], Ian R. Cole[1], Brian Goss[1] ✉, Thomas R. Betts[1], Ralph Gottschalg[1]

[1]*Centre for Renewable Energy Systems Technology (CREST), School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, LE11 3TU, UK*
✉ *E-mail: bgoss@theiet.org*

Abstract: Solar irradiance data is used for the prediction of solar energy system performance but is presently a significant source of uncertainty in energy yield estimation. This also directly affects the expected revenue, so the irradiance uncertainty contributes to project risk and therefore the cost of finance. In this study, the combined impact of temporal averaging, component deconstruction and plane translation mechanisms on uncertainty is analysed. A new method to redistribute (industry standard) hourly averaged data is proposed. This clearness index redistribution method is based on the statistical redistribution of clearness index values and largely corrects the bias error introduced by temporal averaging. Parameters for the redistribution model were derived using irradiance data measured at high temporal resolution by CREST, Loughborough University, over a 5-year period. The root mean square error of example net annual (2014) diffuse, beam and global yield of hourly averaged data were reduced from ~15 to 1, 14 to 3 and 4 to 1%, respectively.

1 Introduction

High-quality datasets containing beam and diffuse irradiance are used in many fields of engineering and research, including solar energy, climate modelling, building performance, thermodynamics, material science and the study of transmittance and reflectance. In particular, developers of solar photovoltaic (PV) and solar thermal systems require accurate predictions of system energy output to develop robust business cases and secure project investment. The uncertainty of irradiance data varies from 5 to 25% [1] and therefore is the predominant variable contributing to uncertainty in system performance modelling [2, 3]. A reduction in the uncertainty of system output predictions equates to a reduction in investment risk.

Many weather monitoring organisations including the UK Met Office record irradiance as hourly averages measured with good quality pyranometers. For solar performance modelling, a minimum of 10, but ideally 20, years of irradiance data are aggregated into a typical meteorological year (TMY). A TMY is considered the best available approximation to the irradiance

profiles of the forthcoming 20 years. Fifteen-minute resolution data is available for some regions and is typical for irradiance data derived from satellite imaging.

The meteorological standard measurement for solar energy is termed 'global horizontal irradiance'. This quantifies the power density received by a horizontal plane from the whole sky. The horizontal plane is sub-optimal for PV installation and optimal installation planes vary from location to location. Generally, for the UK, the optimal installation plane is approximately south facing with a 30–35° inclination [4]. In order to assess irradiance available to a given plane, translation algorithms are applied to the horizontal irradiance data. The outputs of these translation algorithms are typically non-linear to input irradiance, so the use of average values can lead to errors with a bias element. This paper assesses the consequences of the averaging of irradiance measurements and investigates a proposed solution to correct the bias introduced by averaging.

2 Data used for analysis and validation

The data used in this paper was collected at the Centre for Renewable Systems Technology (CREST), Loughborough University. The measurement specification is summarised in Table 1.

For quality control purposes, filters were applied to all input data before analysis to remove data associated with very high measurement uncertainty which might introduce a bias into the analysis. The data quality filters are summarised in Table 2.

The Solys2 sun tracker and sensors at Loughborough University used to collect data to validate the method in this paper are shown in Fig. 1.

3 Overview of horizontal to in-plane irradiance translation

The process to convert global horizontal irradiance to global in-plane irradiance consists of two main stages: first, the deconstruction model which separates global horizontal irradiance into beam and diffuse components; second, the in-plane irradiance model which translates horizontal diffuse to in-plane diffuse. Translation of the beam component requires simple geometry whereas the translation of diffuse irradiance requires a model of

Table 1 Specification of irradiance measurements

location	CREST, Loughborough University, UK
latitude, longitude, deg	52.7616, −1.2406
altitude, m	79
measurement	global horizontal irradiance (G_h)
reference years	2010–2014
pyranometer	Multiple Kipp and Zonen CMP11 s with anti-condensation heating and ventilation
datalogger	Campbell CR1000
resolution at source	1 s samples (with a response delay of 5–15 s)
calibration	traceable to international standards

Table 2 Data filtering parameters

Parameter	Rule
global irradiance (G_h)	<20 W/m^2
clearness index (k_t)	>1.2
solar elevation angle (h)	>0°

directional distribution of irradiance from the sky dome, which can be isotropic or anisotropic.

Deconstruction models typically determine the diffuse fraction of light through an empirical model of the form $k_d = f(k_t)$, where k_d is the ratio of diffuse irradiance to global irradiance. k_t is the 'clearness index' which is the ratio of global horizontal irradiance upon the Earth's surface to the total irradiance received by an extra-terrestrial horizontal plane [5]. The overall translation from global horizontal (G_h) to in-plane irradiance (G_k) takes the form as follows:

$$G_k = f(G_h, k_t, \text{AST}, \alpha_a, \varepsilon_a) \qquad (1)$$

where α_a and ε_a are the azimuth and elevation angles of the inclined plane and AST is the apparent solar time.

The Orgill and Hollands model [6] extended the analysis from daily averages to hourly values and developed piecewise relationships for determining k_d based on k_t bands. The work of Boland *et al.* [7] was improved by Ridley *et al.* [8] by introducing the multiple predictors; apparent solar time, solar height and persistence index; into the Boland, Scott, Luther (BRL) model [9]. The empirical nature of these models means that their accuracy depends on local climatic and topographic features. Therefore,

Fig. 1 *Photograph showing instruments for measurement of horizontal, beam and diffuse irradiance measurement on a Solys2 sun tracker at Loughborough University*

deconstruction and translation models are often chosen based on the relative success of their validation for a given region. There have been various reviews of the accuracy of diffuse irradiance models [10, 11] including for the UK climatic conditions a review by a Loughborough University postgraduate student. An optimised form of the BRL model with coefficients most suitable for use at Loughborough and other locations across the UK was created [12]. The resulting equation (2) is used throughout this paper and is referred to as the modified BRL equation (see (2)) where k_d is the diffuse fraction, k_t is the instantaneous clearness index, AST is the apparent solar time, K_t is the daily average clearness index and α is the humidity fraction.

The primary cause of uncertainty when calculating diffuse irradiation upon an inclined plane lies not within the horizontal to in-plane translation model but in the separation of the diffuse component from global horizontal irradiance (G_h) [13]. The uncertainty in G_h in turn depends on the irradiance measurement method [14–17].

4 Impact of temporal averaging on uncertainty

Fig. 2 demonstrates how lower temporal resolution measurements result in a loss of detail, in particular time averaged data smooths out many dynamic effects and reduces the variance.

The modified BRL equation is non-linear, therefore any loss of extreme values in the input data will not necessarily balance out in the final output and will thus introduce a bias between the calculated diffuse and beam components.

Furthermore, validation of the DC electrical performance of PV arrays compared with modelled performance has identified that high temporal resolution irradiance data is required to accurately model performance during periods of high irradiance (>1000 W/m^2) [17]. High irradiance values with clearness indexes in excess of 0.9 are likely to be due to cloud enhancement.

Fig. 3 shows that the frequency distributions of the clearness index for hourly averaged values compared with mid-hourly spot values shows fewer high and low range values and more mid-range values. This results in higher diffuse irradiance and lower beam irradiance when the global irradiance is deconstructed, as shown in Fig. 5.

It should be noted that the different form of the hourly average distributions relative to the spot value distributions are often exaggerated in irradiance modelling due the use of fitted probability distribution functions (mostly Weibull or beta
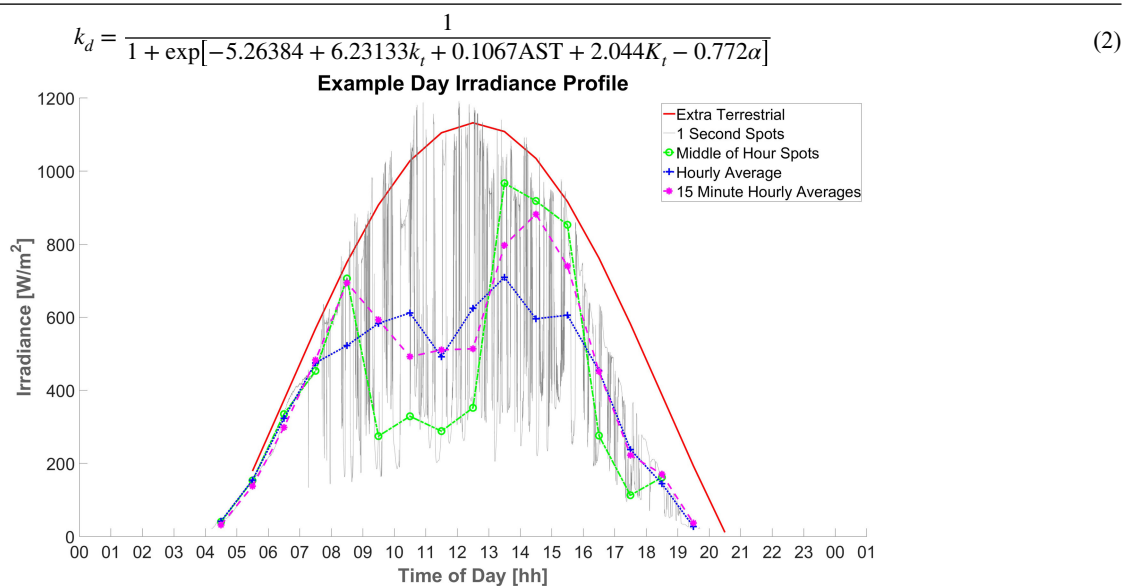
$$k_d = \frac{1}{1 + \exp\left[-5.26384 + 6.23133 k_t + 0.1067\text{AST} + 2.044 K_t - 0.772\alpha\right]} \qquad (2)$$

Fig. 2 *Comparison of global horizontal irradiance measured at Loughborough University at 1 s intervals and hourly averages with extra-terrestrial irradiance for comparison on a day with frequent changes in cloud cover. Note that when satellite data is used in an hourly simulation the hourly averaged irradiance is taken from satellite images at 15 min intervals (hence the average of four spot-values)*
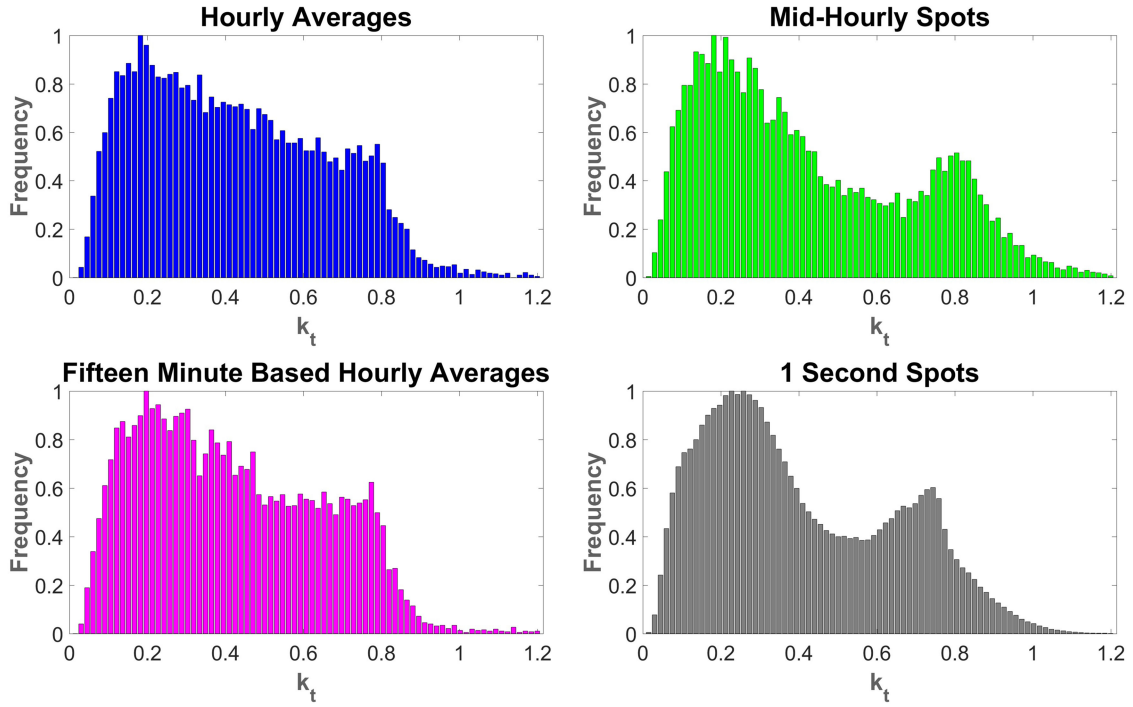
Fig. 3 k_t distributions of four different measuring methods: hourly averages (top left), mid-hourly spot values (top right), hourly averages based on 15 min spot values (bottom left) and the full high-resolution dataset (bottom right)
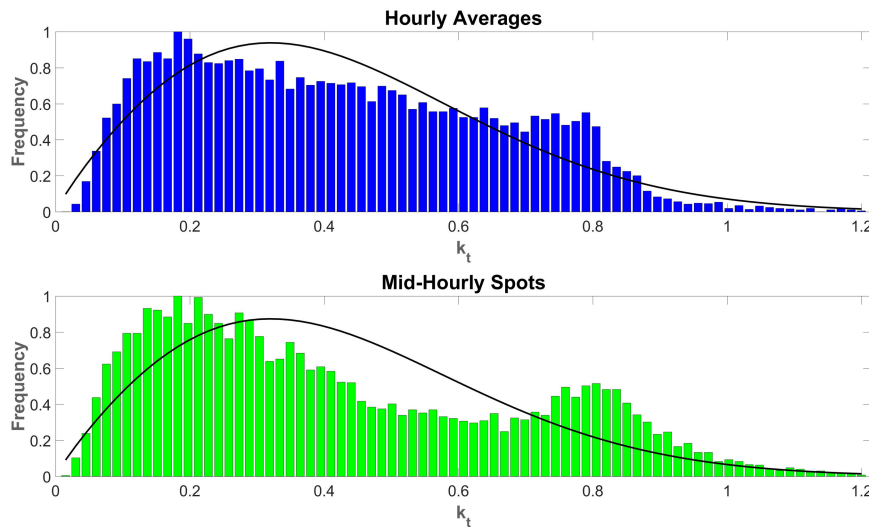

Fig. 4 k_t distributions of two different measuring methods: hourly averages (top), mid-hourly spot values (bottom)

functions) over such datasets [18]. As an example of this, hourly averages and mid-hourly spot distributions from Fig. 3 are shown with fitted Weibull functions superimposed in Fig. 4.

The translation of beam and diffuse horizontal to inclined plane irradiance accentuates the difference between averaged and spot values as shown in Fig. 5. Averaging underestimates beam and overestimates diffuse irradiation on an annual basis for south-facing planes in the UK at any tilt angle.

5 CREST method of temporal bias compensation

Analysis of the variance of high-resolution data suggests that solar elevation angle and hourly averaged clearness index are the two most suitable variables for approximating the true hourly variance of averaged irradiance data. In Fig. 6, the greatest standard deviation occurs within hours with the mid-range of averaged clearness index, whereas there is less variation during hours with very clear or very overcast skies. The root mean square error (RMSE) is 0.065 and R^2 is 0.386, it is recognised that this fit does not accurately represent the chaotic spread of the points. However, it captures a significant trend in the data which is otherwise

ignored, and which leads to the conclusive results described below.

An empirically fitted polynomial surface shown in Fig. 6 generated (3) and (4) for the approximated standard deviation (σ_{approx}) of the hourly average clearness index ($k_{t,h}$) which must be calculated for each time step

$$\sigma_{\text{approx}} = f(\overline{k_t}, h) \tag{3}$$

$$\sigma_{\text{approx}} = p_{00} + p_{10}k_t + p_{01}h + p_{20}k_t^2 + p_{11}k_t h \\ + p_{02}h^2 + p_{30}k_t^3 + p_{21}k_t^2 h + p_{12}k_t h^2 \tag{4}$$

where for this 5-year dataset at Loughborough: $p_{00} = 0.04997$; $p_{10} = -0.09304$; $p_{01} = -0.1554$; $p_{20} = 0.2878$; $p_{11} = 1.676$; $p_{02} = -0.05915$; $p_{30} = -0.1638$; $p_{21} = -1.667$; $p_{12} = -0.07647$; $h = $ solar elevation angle (deg).

The clearness index is the most relevant parameter when considering the relative intensities of diffuse and beam components of irradiance, however, a single hourly-averaged value is not
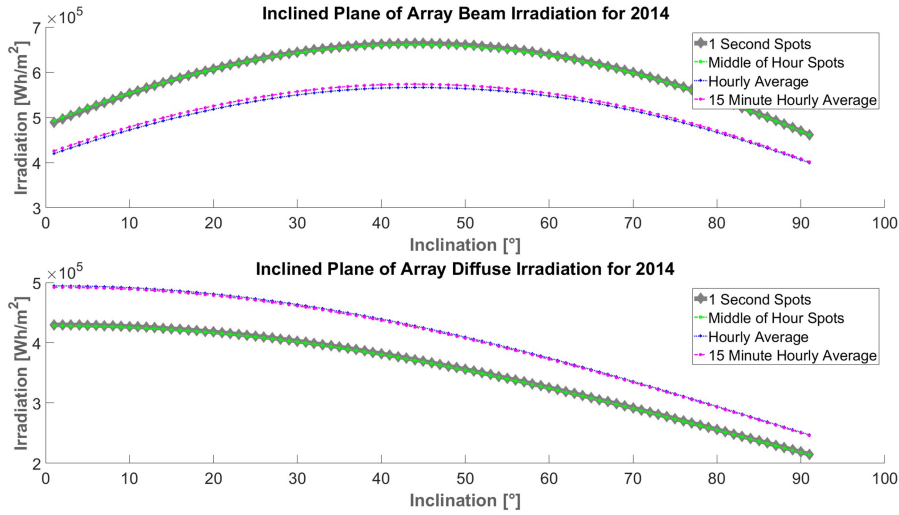
Fig. 5 *Graph showing the irradiation yield against array inclination angle for a south-facing plane of array*
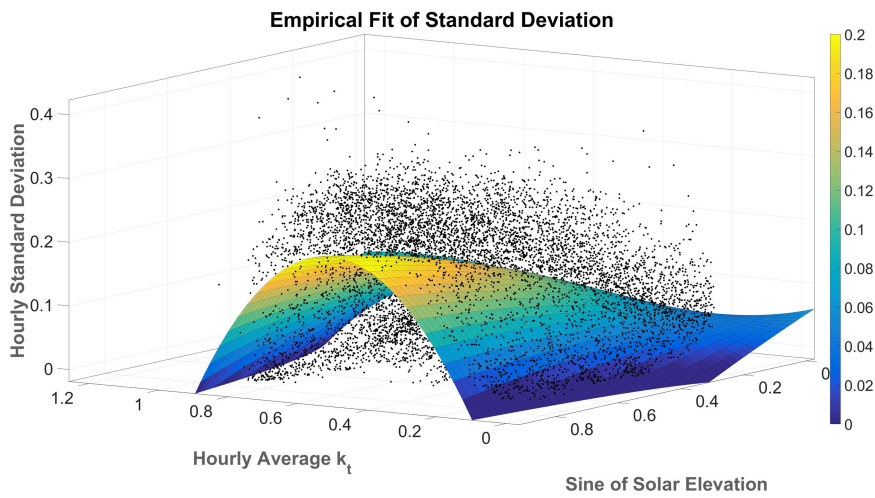

Fig. 6 *Contour plot showing the standard deviation of clearness index as measured using the high-resolution data (z-axis), plotted against the sine of the solar elevation angle computed for the mid-hourly point of each hour (x-axis) and the hourly averaged clearness index for each hour of high-resolution data (y-axis)*
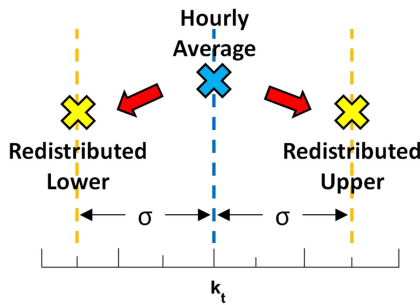

Fig. 7 *Diagram showing how the averaged value is redistributed into two new values*

sufficient for representing the two components. Each hourly standard deviation of k_t values can be used as upper and lower limits to redistribute the hourly averaged value of k_t ($k_{t,h}$) into two new values as shown in Fig. 7.

The relevance of the standard deviation to the two distinct maxima in the distribution of clearness index is shown in Fig. 8. The mean irradiance is shown as the red broken line with a k_t of 0.4080. The mean with standard deviation subtracted and added are shown as the black broken lines with k_ts of 0.1675 and 0.6484, respectively.

The averaged values are redistributed by the approximated standard deviation σ_{approx} as shown in the following equations:

$$k_{t,h1} = k_{t,h} + \sigma_{\mathrm{approx}} \tag{5}$$

$$k_{t,h2} = k_{t,h} - \sigma_{\mathrm{approx}} \tag{6}$$

where $k_{t,h1}$ and $k_{t,h2}$ are the redistributed upper and lower values of k_t for each hour. The computed irradiation yields will each have a weighting of exactly half the original to ensure that the total irradiation of the dataset remains the same. The $k_{t,h1}$ and $k_{t,h2}$ are then used instead of $k_{t,h}$ to calculate in-plane irradiance using the following equation:

$$G_{k1} = f(G_{\mathrm{h}}, k_{t,h1}, \mathrm{AST}, \alpha_a, \varepsilon_a,) \tag{7}$$

$$G_{k2} = f(G_{\mathrm{h}}, k_{t,h2}, \mathrm{AST}, \alpha_a, \varepsilon_a,) \tag{8}$$

The modified in-plane irradiance $G_{k,\mathrm{mod}}$ is then calculated as the arithmetic mean of the upper and lower limits of G_k as shown in the following equation:

$$G_{k,\mathrm{mod}} = \frac{G_{k1}, G_{k2}}{2} \tag{9}$$

The complete CREST method is shown in Fig. 9.
The distribution of the redistributed clearness index compared with the hourly average values is shown in Fig. 10.
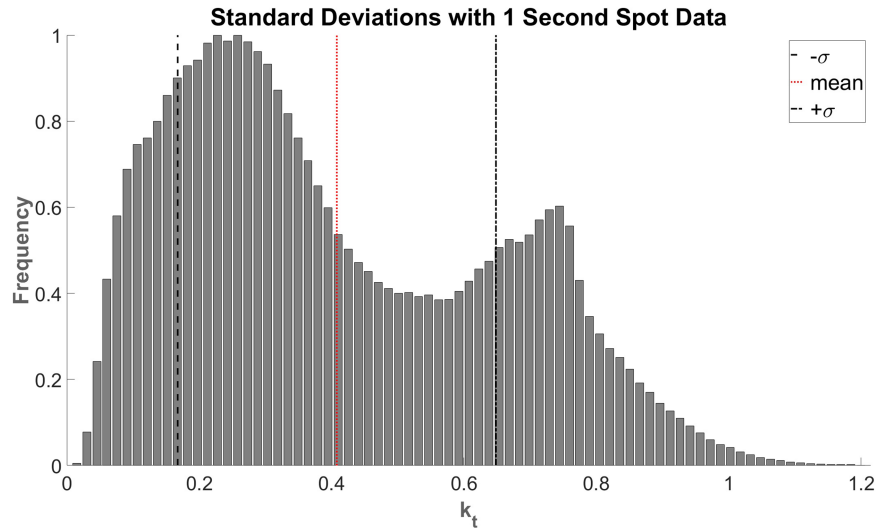
Fig. 8 *Confidence limits of the bimodal k_t distribution – it is assumed that the two distinct maxima correspond to a diffuse distribution combined with a beam distribution. ~99% of points exist within -1.5σ and $+2.6\sigma$*
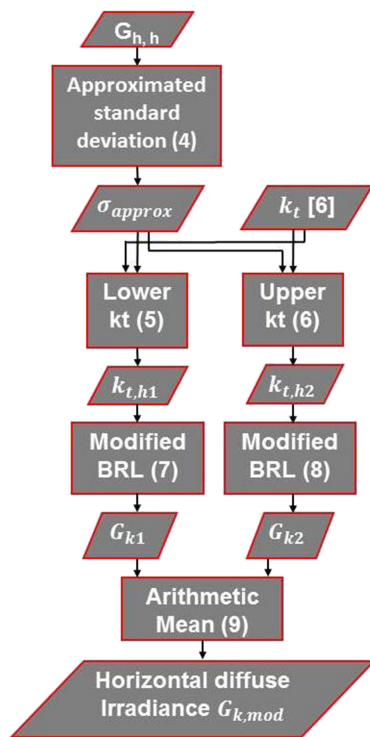

Fig. 9 *Flow diagram showing the complete CREST method described in this paper*

The resulting redistributed values of in-plane irradiance are shown in Fig. 11 alongside the original hourly averages. The new redistributed values show greater variance and better represent the high-resolution measurements.

The redistribution of singular averaged values into two new values results in a more accurate calculation of beam irradiation and diffuse irradiation. Fig. 12 shows that the new redistributed values (yellow line with diamond markers) have been shifted away from their original yields (blue line) toward the values obtained through high-resolution measurements for both beam and diffuse.

Fig. 12 shows the effectiveness of the k_t redistribution procedure in correcting for irradiation yields. It is clearly seen that the procedure brings both the beam and diffuse irradiation yields much closer to those calculated using the high-resolution data, for all collector tilt angles. This not only improves the in-plane global power estimation (see Table 3) but would also improve the estimation of the incident spectrum for spectrally resolved simulations.

Table 3 provides a summary of the RMSE values for 2010–2014 as calculated across 0–90° of array inclination angle as compared with high temporal resolution measurements. All inclined irradiation calculations are for a south-facing plane of array. The main source of uncertainty addressed in this paper is the combination of in-plane translations of pre-averaged data. Therefore, the reference dataset for calculation of RMSE is global horizontal irradiance measured at 1 s intervals which was translated to in-plane using the original method without redistribution, hourly averages were then taken from the total in-plane irradiance. As seen in Fig. 12, the deviation in annual irradiation between the hourly averaged and redistributed values is proportional to the irradiation, therefore the RMSE values shown in Table 3 were taken as an average of the RMSEs for all angles of inclination from 0 to 90°. For locations where spot values of hourly data are available then these should be used rather than hourly averages. However, hourly averages are the meteorological standard for recording solar irradiation in many regions, where it is recommended that the new method described above is used.

It should be noted that the relative reduction in error of the total in-plane irradiation from the hourly averages data comes from the combination of overestimated diffuse and underestimated beam components. For spectrally resolved simulations this error is effectively amplified as the spectral composition of the estimated energy is skewed toward the diffuse spectrum. The effects of spectral deviations on the performance of various PV systems are analysed in more detail in [20–21].

6 Conclusion

The bias introduced by the averaging of solar irradiance data has been analysed over an extended dataset. Diffuse and beam components each have a significant bias error when calculated from hourly averaged global irradiance measurements, with beam irradiance underestimated and diffuse irradiance overestimated. In terms of annual irradiation yield these bias errors were ~15% for the UK case study. A method to compensate for this bias by clearness index redistribution was proposed and investigated. The application of the method reduced the error in the in-plane irradiance calculations such that the higher resolution measurement data was better approximated. Hence, the clearness index redistribution method proposed in this paper can be used to reduce the uncertainty of irradiance data for solar energy simulations derived from hourly averages (which is the industry standard approach).

The new clearness index redistribution method improved the RMSE of the extracted beam component on average from 14.38 to 2.51% and for the diffuse component from 15.08 to 0.79%. The improvement in the accuracy of the total in-plane irradiation was more moderate with a reduction in RMSE of global in-plane
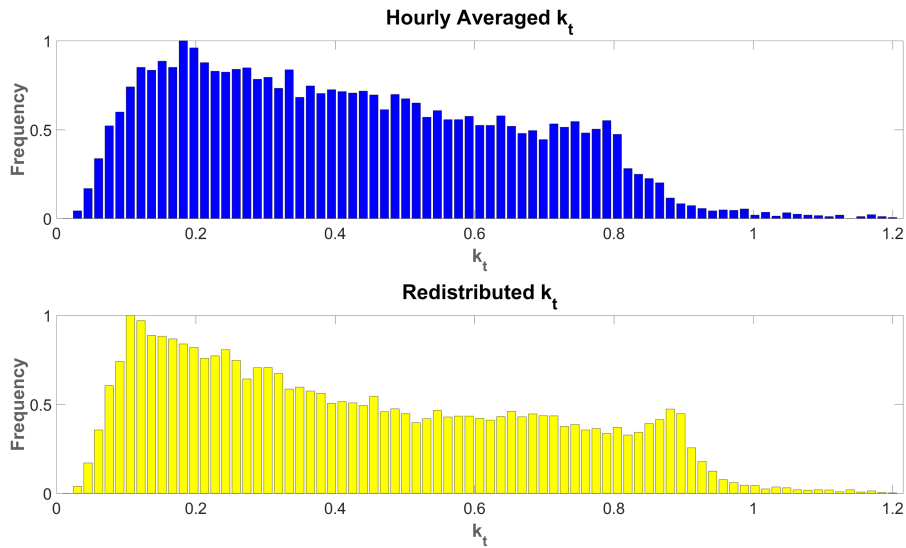
Fig. 10 k_t distribution comparison – original hourly averaged values (top) and redistributed values (bottom)
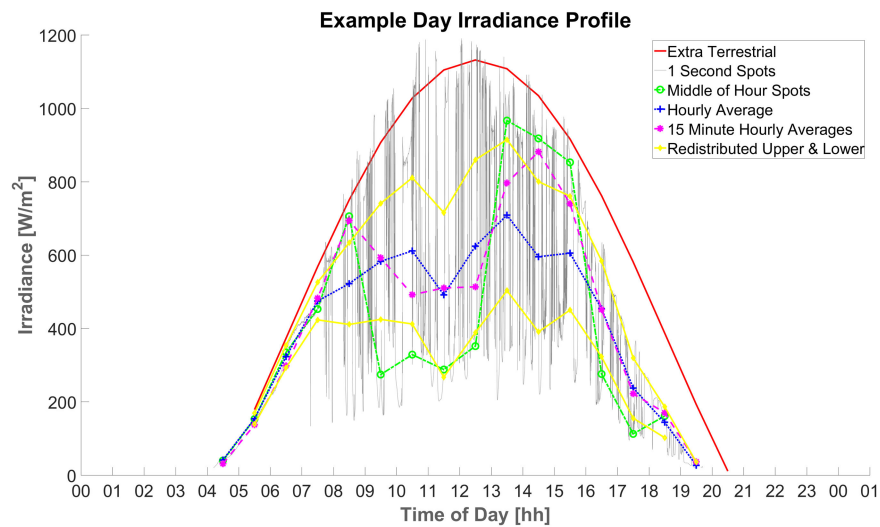

Fig. 11 Daily irradiance profile showing irradiance measured at Loughborough University with different averaging approaches alongside the original 1 Hz measurements
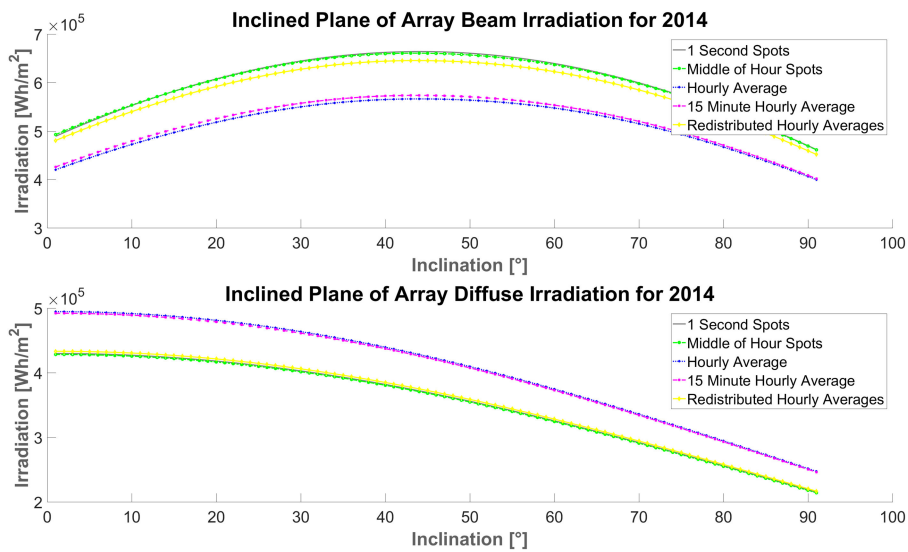

Fig. 12 Graph showing irradiation yield against array inclination angle for a south-facing plane of array

irradiation from 3.73 to 1.31%. However, for spectrally resolved simulations the improvement is effectively amplified due to a reduction in spectral skew.

The difference in effective correction of beam and diffuse components suggests that an asymmetric redistribution procedure may offer a further improvement in performance. This will be investigated by the authors in future work.

Table 3 RMSEs of averaged irradiance measured at Loughborough University from 2010 to 2014 with 0° to 90° inclination

Data sampling technique	RMSE, %		
	Beam in-plane irradiation	Diffuse in-plane irradiation	Total in-plane irradiation (recombined)
15 min hourly average	13.39	14.49	3.36
mid-hourly spot samples	0.35	0.35	0.31
hourly averages	14.38	15.08	3.73
hourly averages redistributed	2.51	0.79	1.31

7 Acknowledgments

8 References

[1] Goss, B., Gottschalg, R., Betts, T.R.: 'Uncertainty analysis of photovoltaic energy yield prediction'. 8th Photovoltaic Science, Applications and Technology Conf. (PVSAT-8), 2012

[2] Strobel, M.B., Betts, T.R., Friesen, G., *et al.*: 'Uncertainty in photovoltaic performance parameters – dependence on location and material', *Sol. Energy Mater. Sol. Cells*, 2009, **93**, (6-7), pp. 1124–1128

[3] Goss, B., Cole, I.R., Koumpli, E., *et al.*: '*The performance of photovoltaic (PV) systems: modelling, measurement and assessment*', in Pearsall, N. (Ed.): (Woodhead Publishing, 2017, 1st edn.), ch. 4, p. 104

[4] Goss, B.: '*Design process optimisation of solar photovoltaic systems*' (Loughborough University, 2015), ch. 6

[5] Liu, B.Y.H., Jordan, R.C.: 'The interrelationship and characteristic distribution of direct, diffuse and total solar radiation', *Sol. Energy*, 1960, **4**, (3), pp. 1–19

[6] Orgill, J.F., Hollands, K.G.T.: 'Correlation equation for hourly diffuse radiation on a horizontal surface', *Sol. Energy*, 1977, **19**, (4), pp. 357–359

[7] Boland, J., Ridley, B., Brown, B.: 'Models of diffuse solar radiation', *Renew. Energy*, 2008, **33**, (4), pp. 575–584

[8] Ridley, B., Boland, J., Lauret, P.: 'Modelling of diffuse solar fraction with multiple predictors', *Renew. Energy*, 2010, **35**, (2), pp. 478–483

[9] Boland, J., Scott, L., Luther, M.: 'Modelling the diffuse fraction of global solar radiation on a horizontal surface', *Environmetrics*, 2001, **12**, (2), pp. 103–116

[10] Torres, J.L., De Blas, M., García, A., *et al.*: 'Comparative study of various models in estimating hourly diffuse solar irradiance', 2010

[11] Jacovides, C.P., Tymvios, F.S., Assimakopoulos, V.D., *et al.*: 'Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation', *Renew. Energy*, 2006, **31**, (15), pp. 2492–2504

[12] Smale, N.: '*Assessment and development of diffuse irradiance models for application within the UK, an MSc project report*' (Loughborough University, 2011), pp. 1–10

[13] Smietana, P.J., Flocchini, R.G., Kennedy, R.L., *et al.*: 'A new look at the correlation of K_d and K_t ratios and at global solar radiation tilt models using one-minute measurements', *Sol. Energy*, 1984, **32**, (1), pp. 99–107

[14] Rösemann, R.: '*A guide to solar radiation measurement*' (Gengenbach Messtechnik, 2011, 2nd edn.)

[15] Gueymard, C.A., Myers, D.R.: 'Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling', *Sol. Energy*, 2009, **83**, (2), pp. 171–185

[16] Sengupta, M., Habte, A., Kurtz, S.: '*Best practices handbook for the collection and use of solar resource data for solar energy applications*' (National Renewable Energy Laboratory, 2015)

[17] Ransome, S., Funtan, P.: 'Why hourly averaged measurement data is insufficient to model PV system performance accurately'. 20th European Photovoltaic Solar Energy Conf., 2005, pp. 2752–2755

[18] Tina, G., Gagliano, S.: 'Probabilistic analysis of weather data for a hybrid solar/wind energy system', *Int. J. Energy Res.*, 2011, **35**, (3), pp. 221–232

[19] Krawczynski, M.: '*Solar spectral irradiance: measurement and application in photovoltaics*' (Loughborough University, 2014)

[20] Cole, I.R.: '*Modelling CPV*' (Loughborough Unversity, 2015)

[21] Betts, T.R.: '*Investigation of photovoltaic device operation under varying spectral conditions*' (Loughborough University, 2005)