Article:

Munoz, M.N., Ballantyne, E.E.F. orcid.org/0000-0003-4665-0941 and Stone, D.A. orcid.org/0000-0002-5770-3917 (2022) Development and evaluation of empirical models for the estimation of hourly horizontal diffuse solar irradiance in the United Kingdom. Energy, 241. 122820. ISSN 0360-5442

https://doi.org/10.1016/j.energy.2021.122820

Development and evaluation of empirical models for the estimation of hourly horizontal diffuse solar irradiance in the United Kingdom.

Maria Nunez Munoz[a,b], Erica E.F. Ballantyne[a], David A. Stone[b]

[a] University of Sheffield, Management school

[b] University of Sheffield, Department of Electronic and Electrical Engineering

Abstract

Solar irradiation data is required in many applications to obtain the solar energy output. However, solar irradiation data is not always available, especially horizontal surface diffuse solar irradiation. This has led to the creation of empirical solar models to predict these values. The aim of this paper is to develop an empirical model for the estimation of hourly diffuse solar irradiance on horizontal surfaces. Initially, a set of existing empirical models have been evaluated to test their accuracy. New models have been developed based on ground level measurements of global and diffuse horizontal solar irradiation that have been obtained from three different regions of the United Kingdom: South Yorkshire, Norfolk, and West Sussex for a period between 1982 and 1999. The models created for each region correlate the diffuse fraction (kd) and the clearness index (kt), for the estimation of hourly diffuse solar irradiance on horizontal surfaces. The models developed for the three regions were evaluated to test their accuracy using error histograms and by contrasting different datasets. Finally, a comparison of the new models to the existing empirical models showed that the new developed correlations significantly improved the existing empirical models for the three regions evaluated in this study.

Keywords: Solar radiation; Clearness index, Hourly diffuse irradiance, empirical model

1. Introduction

Solar energy provides a limitless, clean, and environmentally friendly alternative to fossil fuels. It has a wide range of uses in applications such as solar photovoltaics (PV), solar water heating, solar ventilation, and lighting. Predominantly, solar energy systems are placed on an inclined surface where the solar radiation received is the sum of the direct, diffuse, and reflected irradiation on tilted surfaces. When calculating or predicting these values, these three parameters can be obtained by from knowledge of the global horizontal solar irradiation (H) and the diffuse horizontal solar irradiation (H_d). Therefore, accurate data on horizontal solar irradiation is an essential precondition for the design and modelling of any solar system [1].

Horizontal solar irradiation can be measured via a satellite or a pyranometer/pyrheliometers (at ground level). Whilst there is more data available from satellite measurements [2], ground level measurements provide more accurate values when compared to satellite data [3] due to fewer large systematic errors. To identify and reduce those errors, satellite data has to be compared with ground data [4–6]. However, ground measurements of diffuse and direct horizontal irradiation are not only expensive [7,8] but also prove to be complicated values to measure [9]. Thus, the availability of ground level solar radiation measurements is scarce, especially in developing countries [10,11]. Out of all the registered parameters, the horizontal global irradiation is the most frequently found [2,12,13]. In the UK, the horizontal global irradiation data spans from 1947 to 2018 and remains the most frequent form of solar irradiation data measured in the country. However, it is difficult to find a weather station that reports measurements for the full period. The weather stations located in Greater London have been measuring horizontal global irradiation data between 1958 and 2018. In West Sussex, weather stations have reported horizontal global irradiation from 1992 until 2018; and similarly, in Norfolk, this data is available from 1981 to 2006. Although there are other locations in the UK that have measured the horizontal global irradiation, these have generally been over a much shorter period of time (i.e. South Yorkshire from 1982 to 1995). Conversely, ground horizontal diffuse irradiation values have only been measured in two locations in the UK since 2002, Camborne in Cambridgeshire and Lerwick, Scotland [8]. However, the London weather station is the only one that has reported values of ground horizontal diffuse irradiation in the entire UK until 2005 [14].

In the absence of actual horizontal diffuse solar irradiation data, empirical solar models have been extensively used to accurately predict diffuse irradiance in locations where solar radiation data is not measured or available [15]. Empirical models correlate values of diffuse fraction (kd) or diffuse transmittance with other available variables. [12] classified empirical models in two groups. The first group includes models that predict horizontal diffuse irradiance from global horizontal irradiation (H- based models), whilst the second group includes models that depend on different weather variables (non H-based models) such as air temperature, relative humidity, or ratio of sunshine duration. The most common approach to obtain values of horizontal diffuse solar irradiance is using H- based models, particularly, a correlation between diffuse fraction (kd) and clearness index (kt) [12,16]. Due to their simplicity and accuracy, these H- based empirical models are integrated in specialised software for solar energy applications and forecasting [17].

Diffuse fraction (kd) is the ratio between horizontal diffuse solar irradiance (G_d) and the global horizontal solar irradiance (G_H) [18].

$$kd = \frac{G_d}{G_H}$$

Equation 1

Clearness index (kt) is defined as the ratio between G_H and the extra-atmospheric irradiance (G) [18].

$$kt = \frac{G_H}{G}$$

Equation 2

The first empirical model to predict G_d from G_H with a correlation between kd and kt was developed by [18] and since then many researchers have been studying and modifying the correlation to adapt it to different locations. [19] developed the diffuse fraction correlation using ground solar irradiation data for four different cities in the U.S.A. between 1961 and 1976. [20] developed a set of regression models for five locations within the UK. [21] correlated values of diffuse fraction with clearness index with measurements of solar irradiation between 1979-1982 and collected one year's worth of data from three European cities. [22] developed a correlation using ground measured hourly values of global and diffuse solar irradiation on horizontal surfaces from Greece, Portugal, France, and Spain. [7] developed a sigmoid function using kt as a predictor to calculate kd using data from 21 locations around Europe and the United States.

All of the previously mentioned models are based on a correlation between the diffuse fraction (kd) and the clearness index (kt). However, some authors have been simultaneously investigating the effect of additional predictors to calculate diffuse solar irradiance. [21] identified the clearness index, ambient temperature, relative humidity, and solar altitude (γ_S) with the highest relevance in diffuse solar irradiance estimation when compared with 24 other predictors. For a city in the UK, [1] concluded that the correlation between diffuse fraction (kd) and clearness index (kt) improves when sunshine fraction (SF), cloud cover and air mass (m) are included. A similar result was obtained in Tamanrasset (Algeria) where [15] evaluated eighty empirical models of which the most accurate model correlated kd with kt and SF.

Independent of the number of predictors used, many authors have highlighted the dependency of empirical models with the location of the dataset used to develop the correlation [1,19,22]. Thus, given the relevance that the estimation of horizontal diffuse solar irradiance has in solar energy projects, many researchers around the world are nowadays developing empirical correlations to obtain accurate values at specific locations. [17] compared twenty-three existing empirical models developed in different countries with new correlations using datasets from Athalassa (Cyprus). The authors found that among the existing models analysed, those that had been developed in cities which have similar climatological conditions to Athalassa achieved better results. However, the most accurate estimation of diffuse horizontal solar irradiance was obtained by the new correlations created. [23] constructed a set of empirical correlations for the estimation of monthly average daily diffuse horizontal solar radiation over the Algerian Sahara. The authors demonstrated that one of the proposed correlations achieved accurate results when compared to existing models developed in different locations. In a recent study, [2] found that it was not possible to use a unique correlation for all the 19 worldwide locations studied and so the authors arrange each location in an increasing order of latitude to develop three monthly-averaged hourly correlations. One of the three correlations the authors developed spans the entire UK.

For the purpose of this study, six empirical correlations, to estimate hourly horizontal diffuse solar irradiance, have been selected considering the location of its datasets to have similar climatological conditions to the three regions studied (South Yorkshire, Norfolk and West Sussex). A summary of the models mentioned is presented in Table 1 together with detailed information regarding the location, predictors studied, and correlations.

Considering the scarcity of horizontal diffuse solar irradiation ground level measurements in the UK, the accuracy of these type of measurements in comparison with satellite measurements, and the location dependency of empirical models, the objectives of the present study are:

1) To evaluate existing H- based empirical models in different regions of the United Kingdom to determine a best fit for the estimation of horizontal diffuse solar irradiance (G_d).
2) To develop specific site correlations for every region studied within the UK.

The remainder of this paper is structured as follows: firstly, a description of the datasets used in this study is introduced. The datasets are pre-processed, and a quality control is performed to validate them. At the end of this section, the statistical error metrics used to evaluate the models are described. Following, the study and evaluation of existing empirical H-based models, introduced in Table 1, is completed to estimate its accuracy in three different locations: South Yorkshire, Norfolk, and West Sussex. According to results obtained by the evaluation of existing empirical models, a new correlation is developed for each location considered in this study. The new model correlations are evaluated and validated using statistical error metrics, error histograms and a comparative study between different datasets. The findings and conclusions are summarised at the end of the paper.

Table 1. H- based models selected for evaluation

Model		Location	Predictors	Correlation	
Torre's model [24]	Model 1	Spain *(1)*	kt	$kt \leq 0.225$	$kd = 0.9943 - 0.1165 \cdot kt$
				$0.225 < kt < 0.755$	$kd = 1.4101 - 2.9918kt + 6.4599kt^2 - 10.329kt^3 + 5.514kt^4$
				$kt > 0.755$	$kd = 0.18$
De Miguel's model [22]	Model 2	France *(3)*, Portugal *(4)*, and Spain *(1)*	kt	$kt \leq 0.21$	$kd = 0.995 - 0.081 \cdot kt$
				$0.21 < kt < 0.76$	$kd = 0.724 + 2.738kt - 8.32kt^2 + 4.967kt^3$
				$kt > 0.76$	$kd = 0.180$
Reindl's model (1) [21]	Model 3	Denmark *(1)*, Germany *(1)*, Ireland *(1)* and United States *(2)*	kt	$kt \leq 0.3$	$kd = 1.020 - 0.248 \cdot kt$
				$0.3 < kt < 0.78$	$kd = 1.45 - 1.67kt$
				$kt > 0.78$	$kd = 0.147$
Muneer's model (1) [20]	Model 4	United Kingdom *(5)*	kt	$kt > 0.2$	$kd = 0.687 + 2.932kt - 8.546kt^2 + 5.227kt^3$
Reindl's model (2) [21]	Model 5	Denmark *(1)*, Germany *(1)*, Ireland *(1)* and United States *(2)*	kt and γ_S	$kt \leq 0.3$	$kd = 1.020 - 0.254 \cdot kt + 0.0123 \sin(\gamma_S)$
				$0.3 < kt < 0.78$	$kd = 1.4 - 1.749kt + 0.177 \sin(\gamma_S)$
				$kt > 0.78$	$kd = 0.486kt - 0.182 \sin(\gamma_S)$
Muneer's model (2) [25]	Model 6	United Kingdom *(2)*	kt, SF and m	-	$kd = (0.899 - 0.683SF + 0.648SF^2 + 0.028m - 0.002m^2) + (0.880 - 0.666SF - 0.314SF^2 - 0.158m + 0.003m^2)kt + (-1.751 + 2.786SF - 1.924SF^2 + 0.044m + 0.012m^2)kt^2$

**The number in parenthesis in the location column, reflects the number of cities studied in each location for the development of the correlation.*

2. Quality control of UK hourly solar irradiation data

2.1 Dataset information and climatological conditions at the weather stations.

For this study, the three datasets were obtained through the Centre for Environmental Data Analysis (CEDA) Archive. The CEDA Archive is the UK national data centre for atmospheric and earth observation research that ensures easy access to horizontal solar irradiation data from the open data version of Met Office Integrated Data Archive System (MIDAS) [14,26].

The datasets correspond to hourly measurements of horizontal global and diffuse solar irradiation in kJ/m^2. The measurements contain the amount of solar irradiance received during the hour ending at the specified time. The suggestion of [27] that a minimum of three years of data are needed to evaluate and validate solar radiation models was considered for each location. The location of each weather station, the period in which measurements were taken and the number of measurements recorded for each location are shown in Table 2.

Table 2. Information of raw data values of horizontal solar irradiation measured by Met Office, both hourly horizontal global and diffuse solar irradiation.

Location	Latitude [degrees]	Period of measurements	Number of values recorded
Finningley (South-Yorkshire)	53.4845	1982-1995	113,154
Hemsby (Norfolk)	52.6953	1982-1999	134,605
Crawley (West Sussex)	51.1059	1982-1992	91,529

As it is shown in Figure 1, Finningley, Hemsby and Crawley weather stations are sited within three different UK regional areas; the North East, the East of England, and the South, respectively [28].

The climate on the north east of England [28], where Finningley station is located, is characterised by cool temperatures throughout the year when compared to elsewhere in England with temperatures varying from -0.5°C to about 2°C during the coldest month of the year (January) and between 17°C and 21.5°C during the warmest months (July and August.). As we move further south, the mean temperatures, both for the coldest and warmest months, increase.

In the eastern climate [28], the mean temperatures vary between 0°C to 2°C in winter to 20°C -23°C in summer and it is one of the driest areas in the country with an average rainfall of 700 mm per year.

Temperatures in the southern England climate fluctuate from 0.5°C to about 3°C in January and 21°C -23.5°C in July [28]. The southern region has the sunniest places in mainland UK, with an average annual sunshine duration between 1550 to 1600 hours.

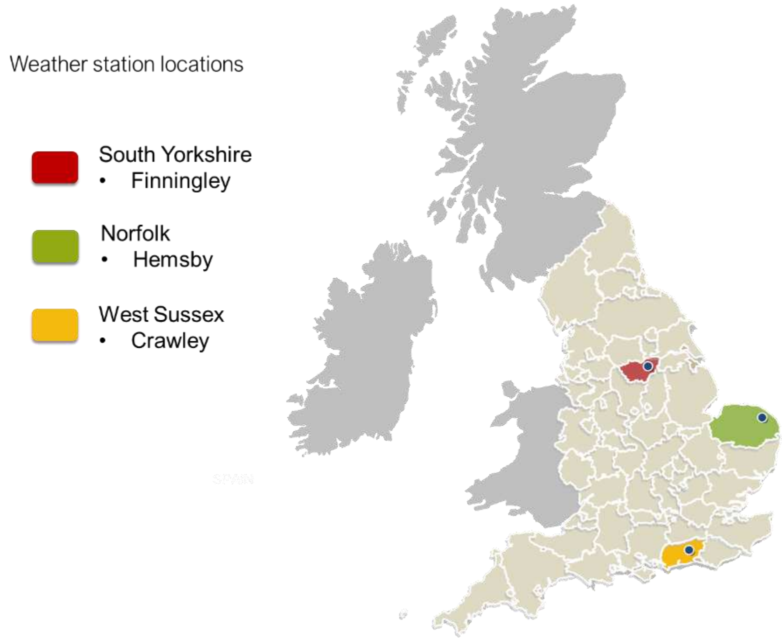

Figure 1. Location of Met Office weather stations used for the study

2.2 Dataset pre-processing

A number of pre-processing steps are completed so that raw data can be used, these steps are described here and shown in Figure 2. The raw datasets from Table 2 were pre-processed to discard unnecessary values (i.e. invalid values recorded at 23:59, that correspond to daily irradiation). Through the examination of the raw data, sections of missing data were identified. These gaps were due to equipment error, operation related problems or diffuse irradiance data processing errors [29] and these values were also discarded.

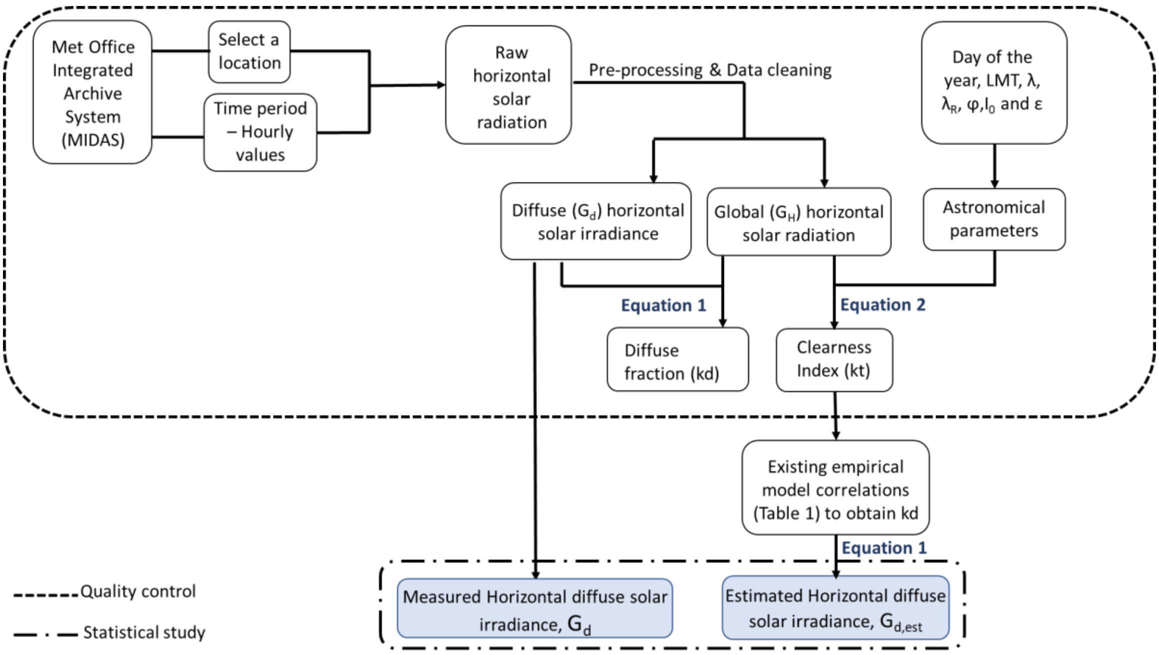

Figure 2. Complete process to evaluate the existing empirical models based on measured data from South Yorkshire, Norfolk, and West Sussex.

Next, the astronomical parameters were calculated in a time period of one minute. The calculation of astronomical parameters is essential for the estimation of the hourly horizontal extra-atmospheric irradiance (G), the unknown parameter in Equation 2, and thus to progress with the comparative study and development of new correlations.

The horizontal e-atmospheric irradiance, G_0 is calculated from the following Equation 3 [30]:

$$G_0 = I_0 \cdot (1 + 0.033 \cdot cos\,(\frac{360 \cdot DOY}{365}))\ \cdot sin\ \gamma_s\ \left[\frac{W}{m^2}\right]$$

Equation 3

where I_0 is the solar constant with a value of 1,367 W/m^2 [30] and γ_s is the solar altitude angle. In order to keep a consistency in the measurements of solar angles, the Standardised ISO system has been used [31]. The ISO system assumes a north orientated system with angles measured clockwise (0°-360°). Following that system, γ_s is calculated following Equation 4 [31]:

$$\gamma_s = sin\ \varphi \cdot sin\ \delta - cos\ \varphi \cdot cos\ \delta \cdot cos\ \omega \quad [degrees]$$

Equation 4

where φ is the latitude, δ is the declination angle and ω is the hour angle. δ is defined by [30], following the approximate equation of Cooper (1969) [32], and it can be seen in Equation 5:

$$\delta = 23.45 \cdot sin\left[360 \cdot \frac{(DOY + 284)}{365}\right]\ \ [degrees]$$

Equation 5

where DOY is day-of-year, and it represents the nth day of the year (i.e. 1st of January is DOY=1; 1st of February is DOY=32, 1st of March is DOY=60, etcetera. Leap year not included).

The hour angle ω, considering a north orientated system is calculated mathematically as follows [31]:

$$\omega = t \cdot 15°\ \ [degrees]$$

Equation 6

where t is local solar time. According to [33], the standard time used for the UK observations is Universal Time Coordinated (UTC) and daylight savings are not considered. Accordingly, t is estimated following [34] as;

$$t = LMT + \frac{\lambda - \lambda_R}{15} + EOT \quad [hours]$$

Equation 7

where LMT is the local mean time or civil time. λ is the longitude of the standard time meridian, λ_R is the longitude of the location and EOT is the equation of time [30]. EOT can be estimated following Equation 8:

$$EOT = 229.2 \cdot (0.000075 + 0.001868 \cdot cos\ \mathrm{B} - 0.032077 \cdot sin\ \mathrm{B} - 0.014615 \cdot cos\ 2\mathrm{B} - 0.04089 \cdot \sin 2B)\ \ [minutes]$$

Equation 8

where B is a coefficient calculated by [30] following Equation 9;

$$B = (DOY - 1) \cdot \frac{360}{365}$$

Equation 9

Once all the astronomical parameters were computed, the horizontal e- atmospheric irradiance (G_0) was calculated at time intervals of one minute. For consistency with hourly dataset measurements, G_0 values have then been averaged every 60 minutes to obtain an hourly value, (G [kW/m^2]). Then, dataset hourly irradiation measurements in kJ/m^2 were converted to hourly irradiance kW/m^2 to obtain the clearness index (kt), using Equation 2, and the diffuse fraction (kd) with Equation 1.

Two additional parameters have been calculated for the purpose of obtaining diffuse fraction (kd) using Model 6 (Table 1).

- Sunshine fraction (SF) is the ratio of sunshine duration and day length [3]. Daily values of sunshine duration were provided from the CEDA archive for each location from 1982 to 1999. Day length can be calculated following Equation 10 [34]:

$$Day\ length = \frac{2}{15} \cdot cos^{-1}(-tan\varphi \cdot tan\delta)\ \ [hours]$$

Equation 10

- Air mass (m) was calculated following the steps performed by [1] following Equation 11:

$$m = [sin\gamma_s + 0.50572(\gamma_s + 6.07995)^{1.6364}]^{-1}$$

Equation 11

2.3 Quality control

After the pre-processing of the raw data, a quality control is performed. The objective of the quality control of hourly solar datasets is to test their validity. Four tests were considered for the quality control, three of these were established by [35] and the fourth test was proposed by [29].

The first test eliminates the values of solar altitude (γ_s) lower than 7°. By doing so, data at sunrise and sunset affected by the cosine effect (the error in the sensor's response to the angle at which radiation strikes the sensing area) is deleted.

The second test is a logical test to avoid any value of kt and kd that is not between zero and one [7].

$$0 < kt < 1$$

$$0 < kd < 1$$

For the third test the diffuse horizontal irradiance on a clear day ($G_{d,c}$) and horizontal diffuse irradiance in an overcast day ($G_{d,oc}$) are calculated. Values of horizontal diffuse irradiance (G_d) should be within the values of $G_{d,c}$ and $G_{d,oc}$ [35].

$$G_{d,oc} \leq G_d \leq G_{d,c}$$

To calculate the diffuse horizontal radiation on a clear day ($G_{d,c}$) and on an overcast day ($G_{d,oc}$), [35] followed Page model (for overcast skies using Equation 12 and for clear skies using Equation 13).

$$G_{d,oc} = 572 \cdot \gamma_s$$

Equation 12

$$G_{d,c} = kd \cdot T_{rd} \cdot F(\gamma_s)$$

Equation 13

Where T_{rd} is the theoretical diffuse irradiance on a horizontal surface when the sun is at the zenith [35], and it is calculated following Equation 14:

$$T_{rd} = -21.657 + 41.752 \cdot T_L + 0.51905 \cdot T_L^2$$

Equation 14

T_L is the Linke turbidity factor. The turbidity factor indicates the concentration of aerosols in the atmosphere and reflects the effect of atmospheric scattering and absorption [36]. The higher the concentration on aerosols, the higher the atmospheric scattering.

[37] built a worldwide T_L database that includes values for seven different cities in the United Kingdom. The values of turbidity factor used for this study correspond to monthly values measured in Aughton, London, and Brooms Barn (United Kingdom) between 1981 and 1990. Since the turbidity factor in the database that corresponds to October for all the UK locations is missing, an averaged value between September and November was used.

$F(\gamma_s)$ is the solar elevation function, calculated by [35] following Equation 15:

$$F(\gamma_s) = 3.8175 \cdot 10^{-2} + 1.5458 \cdot sin\ (\gamma_s)\ - 0.59980 \cdot sin\ (\gamma_s)\ ^2$$

Equation 15

In addition, a *fourth test* was introduced for the quality control of the remaining data. The test was proposed by [29], who developed a boundary that considers the ratio of the clearness index (kt) with the diffuse fraction (kd). A representation of the boundaries can be seen in Figure 3.

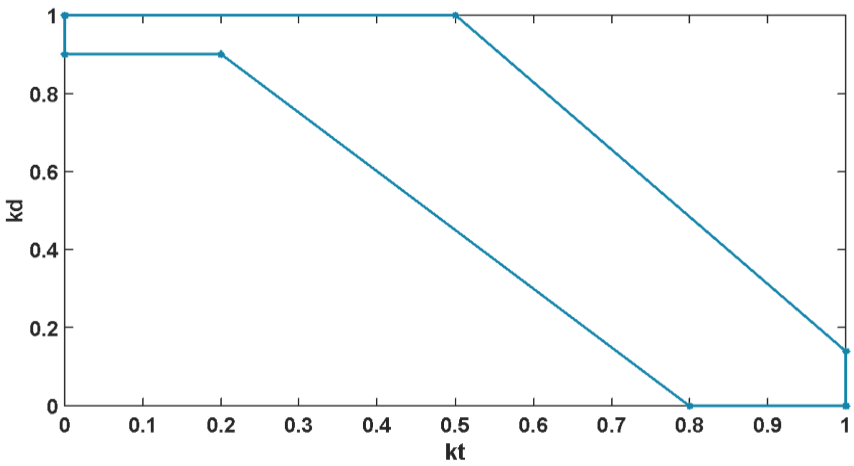

Figure 3. Boundaries developed by [29] *applied in the test 4 of the present study*

The number of values within the remaining datasets after each quality test are shown in Table 3 for each of the study locations. Test 1 deletes values as a result of an equipment error. As it can be seen, this type of error reports the highest impact on the quality control. Specifically, Test 1 reduced the number of data points considered by approximately a 57% at each location. On the other hand, results from Test 2, Test 3 and Test 4 show a smaller impact on the quality control of each dataset. Those tests are designed to keep consistency between different values of solar radiation (e.g. G_d cannot be higher than G) [38].

Table 3. Information on number of values of hourly horizontal solar irradiance after each quality control test.

		South-Yorkshire		Norfolk		West-Sussex	
Test	**Condition applied**	**Before test**	**After test**	**Before test**	**After test**	**Before test**	**After test**
Pre-processing	Deletion of 23:59 data and values of zero solar radiation	113,154	108,016	134,605	130,145	91,529	91,519
Test 1	Solar altitude<7°	108,324	46,053	130,145	54,419	91,519	39,664
Test 2	0<kd<1 0<kt<1	46,053	41,337	54,419	53,037	39,664	32,825
Test 3	$G_{d,oc} \leq G_d \leq G_{d,c}$	41,337	40,316	53,037	52,174	32,825	32,108
Test 4	Boundaries kd and kt	40,316	35,397	52,174	45,627	32,108	27,497

2.4 Statistical error metrics for the evaluation of empirical model

Statistical error metrics have been used extensively in literature to evaluate the performance of empirical solar models. [17] evaluated the empirical models to estimate hourly diffuse fraction for the region of Nicosia (Cyprus) by using the mean bias error (MBE), the root mean square error (RMSE) and the coefficient of determination (r^2). Mean absolute percentage error (MAPE), mean absolute bias error (MABE), RMSE, relative standard error (RSE), and correlation coefficient (r) were used by [39] to assess the accuracy of a set of empirical models, based on the correlation between diffuse solar radiation and clearness index, for the region of Kerman (Iran). [40] tested the performance of a solar model to calculate the diffuse solar radiation using the RMSE metric and [41] appraised daily diffuse solar radiation models for different regions in China by statistical analysis using MBE, RMSE and r amongst other metrics. [7] proposed a regressive model for the estimation of hourly diffuse solar irradiation under all sky conditions that was validated by r^2, MBE and RMSE metrics.

Three of the most commonly used statistical error metrics [42,43] have been selected to assess the performance of the empirical models developed for the study; MBE, RMSE and r^2. Both the MBE and RMSE conserved the units of the variables (kW/m^2).

The MBE metric is extensively applied for the evaluation of solar model performance [44,45]. The MBE metric calculates the arithmetic mean of estimated and measured values of hourly diffuse horizontal irradiance and is defined in Equation 16 as:

$$MBE = \frac{\sum_i^n (G_{d,est,i} - G_{d,meas,i})}{n}$$

Equation 16

Where the $G_{d,est}$ is the estimated hourly diffuse horizontal irradiance (kW/m^2), $G_{d,meas}$ is the measured hourly diffuse horizontal irradiance (kW/m^2) and n is the total number of data points.

Negative or positive values of MBE indicate the model is under-predicting or over-predicting estimated data, respectively. For the purpose of this study, MBE is selected to provide an overall indication of the accuracy of the model. A MBE nearest to zero is desired

The RMSE is the most widely used statistical error metric in the study of the reliability and the degree of accuracy of a solar model [46,47]. More precisely, RMSE metric allows a term-by-term comparison of the difference between estimated and measured values [48]. RMSE is given by Equation 17:

$$RMSE = \sqrt{\frac{\sum_i^n (G_{d,est,i} - G_{d,meas,i})^2}{n}}$$

Equation 17

RMSE metric gives more weight to the largest errors [42] that are a consequence of the cosine effect [49] however, in this study, those are significantly reduced through the first test of the quality control. The lower the absolute value of RMSE, the higher the accuracy of the model.

Finally, the coefficient of determination (r^2) was selected to assess the linearity between measured values and values obtained through the models' correlation (Equation 18).

$$r^2 = \frac{\left[\sum (G_{d,est} - \overline{G_{d,est}}) \cdot (G_{d,meas} - \overline{G_{d,meas}})\right]^2}{\sum (G_{d,est} - \overline{G_{d,est}})^2 \cdot \sum (G_{d,meas} - \overline{G_{d,meas}})^2}$$

Equation 18

Where $\overline{G_{d,est,i}}$ and $\overline{G_{d,meas,i}}$ are the estimated and measured mean hourly diffuse solar irradiance, respectively. r^2 values range between 0 and 1, the latter indicates a perfect linear relationship.

3. H-based empirical model evaluation

The models introduced at the beginning and presented in Table 1 were evaluated to determine which model was the most suitable to estimate values of horizontal diffuse irradiance (G_d). A description of the process to evaluate the existing empirical models is shown in Figure 2.

Once the quality control was performed, values of clearness index (kt) were used to calculate the diffuse fraction (kd) using the correlations presented in Table 1 for each of the models. Next, using values of measured global horizontal solar irradiance (G_H) and Equation 1, the values of estimated horizontal diffuse solar irradiance (G_{dest}) were obtained and compared to real values of G_d. The statistical results for each model can be seen in Table 4.

The results in Table 4 are displayed for the three regions studied, South Yorkshire, Norfolk, and West Sussex. Overall, the correlations that used more than one predictor (Model 5 and Model 6) did not achieve better results when compared with correlations that only used kt as the predictor, for any of the studied locations. However, there is a significant difference in the results of the statistical errors between Model 5 and Model 6. These results show that the solar altitude (γ_s), used in Model 5 has a higher influence in the results than a combination of solar altitude (γ_s), SF and m (as used in Model 6) when these are introduced to the correlation. For the three locations, Model 5 achieved higher values of r^2 and lower values of MBE and RMSE when compared with Model 6. The values of MBE and RMSE show relatively small changes if compared with the differences displayed by r^2.

The results obtained by using Model 1, Model 2 and Model 3 demonstrate a consistency in r^2, MBE and RMSE for all sites studied. No model outperforms the others in all three statistical errors. The highest value of r^2 was achieved by Model 1 and the lowest values of MBE and RMSE corresponded to Model 3 and Model 2.

With Model 4, r^2 has a value equal to 0.77-0.79 depending on the location. This value closely agrees with the results obtained by [20]. It is worth noting that this correlation was designed to cover values of kt>0.2 however, if the correlation is applied to the full range of kt for each dataset, Model 4 performs the best for all the locations. In fact, Muneer's model (Model 4) returns the highest values of r^2 and the lowest values of MBE and RMSE.

As mentioned previously, the empirical models are based on specific datasets that correspond to different locations and so are not universal correlations. Thus, it is expected that Muneer's model correlation, which corresponds to a regressed equation for the United Kingdom, would fit better with measured values than the other models for the locations selected. In general, for South Yorkshire, Norfolk and West Sussex, the values of MBE and RMSE for each model display small variations. Nevertheless, the values of r^2 decrease in every model as latitude decreases.

Table 4. Statistical evaluation of performance of existing empirical models based on ground measured data for 3 locations, South-Yorkshire, Norfolk, and West-Sussex.

South-Yorkshire			
	r^2	MBE	RMSE
Predictors: kt			
Model 1	0.863	-0.022	0.045
Model 2	0.856	-0.013	0.040
Model 3	0.861	-0.015	0.041
Model 4	0.770	-0.0016	0.048
Predictors: kt, γ_s, SF and m			
Model 5	0.858	-0.0097	0.039
Model 6	0.778	0.027	0.065
Norfolk			
	r^2	MBE	RMSE
Predictors: kt			
Model 1	0.839	-0.023	0.048
Model 2	0.836	-0.015	0.044
Model 3	0.837	-0.017	0.044
Model 4	0.789	-0.001	0.039
Predictors: kt, γ_s, SF and m			
Model 5	0.826	-0.0098	0.043
Model 6	0.732	0.031	0.071
West-Sussex			
	r^2	MBE	RMSE
Predictors: kt			
Model 1	0.829	-0.026	0.051
Model 2	0.823	-0.017	0.047
Model 3	0.825	-0.018	0.047
Model 4	0.771	-0.001	0.041
Predictors: kt, γ_s, SF and m			
Model 5	0.819	-0.01	0.045
Model 6	0.734	0.035	0.075

4. Empirical correlations developed at different locations in the UK

After having compared the empirical models with measured data at the three selected locations, a new model has been developed to better predict the horizontal diffuse irradiance. The aim of this section is to develop a specific correlation for each location that best fits with measured data. In accordance with the results obtained in Table 4, from empirical models, the correlations including the clearness index as a predictor (i.e. kt) had the most impact on the accuracy of the results. Thus, the correlations developed here are dependent on one predictor, the clearness index, (kt).

As noted by [4], it is not recommended to use the same dataset to develop and validate the model. Therefore, a certain portion of the datasets, used to develop the correlation, will be referred to as "training datasets" and, the portion of the datasets used to validate the model will be called "validation datasets" [7]. Table 5 provides information on the years used for the training and validation datasets for each location.

Table 5. Detailed information of the dataset used to develop and validate the correlations at each location

	Training dataset	**Validation dataset**
South-Yorkshire	1982-1989	1990-1995
Norfolk	1982-1993	1994-1999
West-Sussex	1982-1987	1988-1992

The description of the process followed to develop a new correlation can be seen in Figure 4 as a diagram and will be described here.

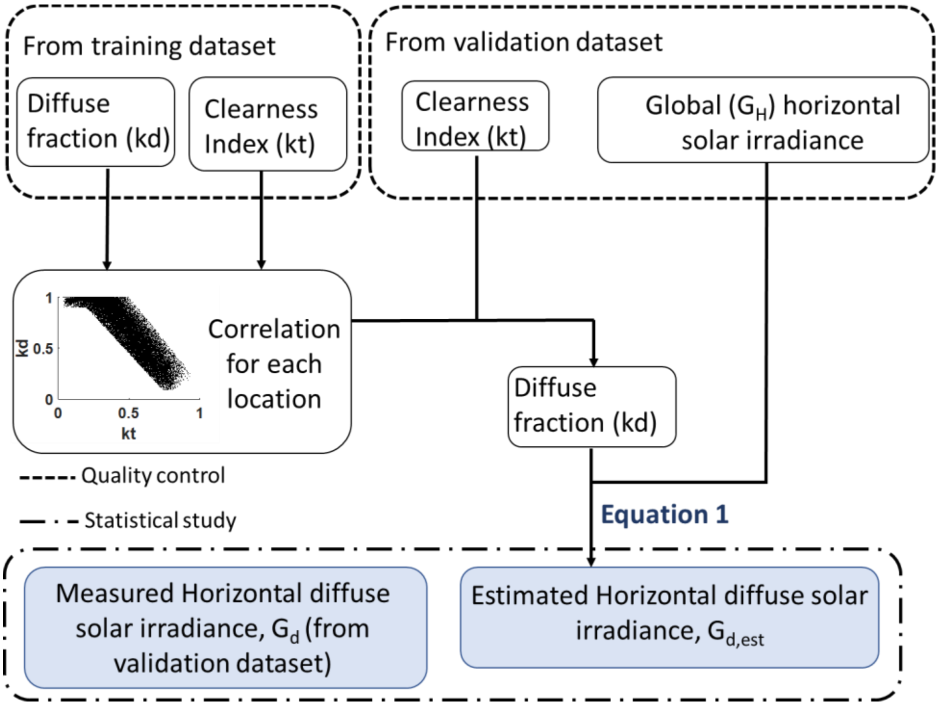

Figure 4. The complete process for the development of a new correlation

For each region, the quality control process was performed for both datasets. Figure 5 illustrates the hourly measurements of clearness index (kt) and diffuse fraction (kd) before and after the quality process for the region of South-Yorkshire.

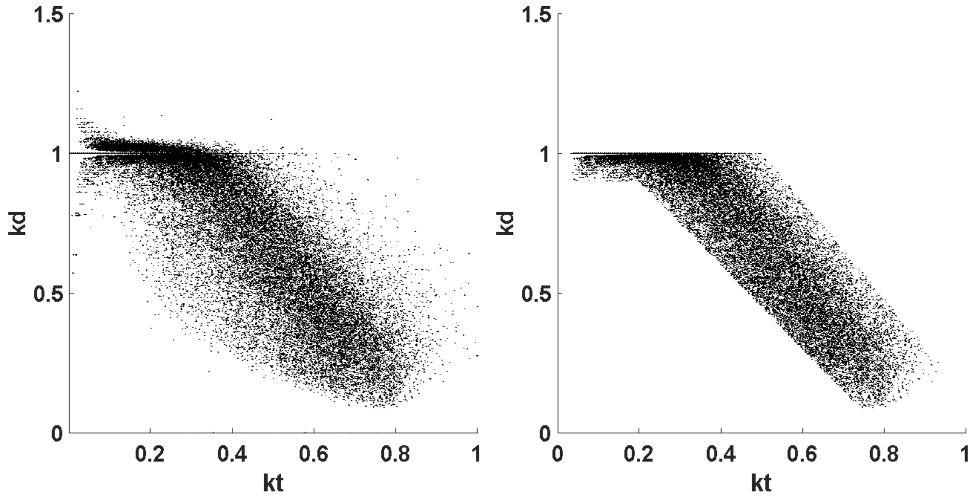

Figure 5. Hourly values of clearness index (kt) and diffuse fraction (kd) corresponding to raw data (left) and the resulting measurements after the quality process (right).

As it can be seen in Figure 6, values of kd and kt from the "training dataset" were used to create a correlation for each location.

Data was separated into sections, so the correlations could better fit each section. For the initial values of kt, the length of these intervals was decided following the approach used by [20]. For 0≤kt≤0.3, values of kd were averaged in increments of 0.05 and plotted. This section of data points showed a linear trend for values of kt between 0 and 0.2, as can be seen in Figure 6. Thus, a linear fit was used for values lower than 0.2

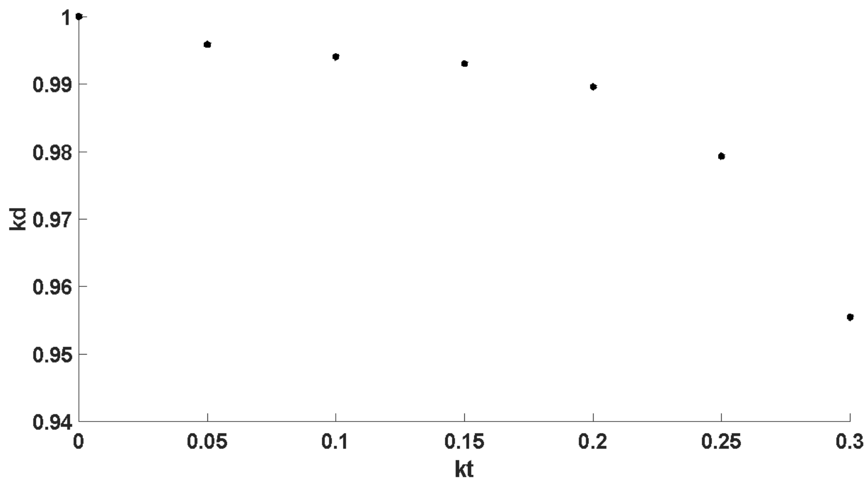

Figure 6. Approach to set intervals for kt<0.2

For values higher than 0.2 for kt, the correlation that reported better results was a 4th order polynomial for South-Yorkshire and Norfolk, and a cubic correlation for West-Sussex. Higher order polynomial equations and correlations with intervals between 0.2 and 0.8 or 0.9 were studied for all the locations but did not display significant improvement in the accuracy of the model, so they were discarded.

Once the correlations were developed, the values of the clearness index (kt) from the “validation dataset” were introduced in each correlation to calculate the diffuse fraction (kd). This diffuse fraction was used together with the global horizontal solar irradiance (G_H) to estimate the hourly values of horizontal diffuse solar irradiance ($G_{d,est}$) using Equation 1. These values were compared with G_d measurements from the “validation dataset” and the statistical evaluation was performed. The equations and the statistical error metrics for South-Yorkshire, Norfolk and West-Sussex are shown in Table 6.

Table 6. Best correlation for data measured in each location.

		r^2	MBE	RMSE
Correlation for South-Yorkshire (ground measurements 1990-1995)				
kt<0.2	kd=0.9982-0.0473kt	0.872	0.0015	0.027
kt>0.2	$kd=0.7392+2.428kt-6.739kt^2+2.626kt^3+1.366kt^4$			
Correlation for Norfolk (ground measurements 1994-1999)				
kt<0.2	kd=0.9996-0.0497kt	0.840	0.004	0.04
kt>0.2	$kd=0.6671+3.76kt-8.078kt^2+5.231kt^3+0.05306kt^4$			
Correlation for West-Sussex (ground measurements 1988-1992)				
kt<0.2	kd=1.0011-0.075kt	0.830	0.003	0.03
kt>0.2	$kd=0.01898+7.117kt-16.54kt^2+10.07kt^3$			

For the region of South-Yorkshire, when compared with the empirical models analysed in Table 4, the new correlation developed shows a clear improvement for the statistical errors. When compared to Model 4, which previously provided the most accurate predictions within the limits of the correlation, the r^2 increases significantly, the MBE also improves slightly, while the RMSE is the statistical error metric that showed the largest improvement.

The correlation developed for Norfolk also shows an improvement if compared to existing empirical models, particularly the value of MBE. Results obtained for r^2 and RMSE are slightly better although the difference is not as significant as it is for the correlation developed for the region of South-Yorkshire.

Finally, the new correlation model created for the region of West-Sussex displays remarkable changes if compared to existing empirical models. The RMSE has been reduced for an average of 0.051 to 0.03. MBE has also been improved when compared to Model 3, Model 1 and Model 2. The r^2 has improved greatly considering Model 4.

Although the overall improvement for Norfolk is not as significant as it is for South-Yorkshire and West-Sussex, the new correlation also improved the results obtained when compared to existing models (Table 4).

5. Model validation

After the new correlations were developed, the model had to be validated. The accuracy of each new correlation was tested by error histograms and by comparison with different datasets.

5.1 Error histograms

Each data point of $G_{d,est}$ calculated with the new correlation, was compared to the measured values of G_d. The errors were calculated keeping the same units as $G_{d,est}$ and G_d (kWh/m^2) as follows:

$$Error = G_{d,est} - G_d$$

Equation 19

The error histograms for South-Yorkshire, Norfolk and West-Sussex are presented in Figure 7, Figure 8, and Figure 9, respectively.

A correlation of horizontal diffuse solar irradiance is considered to be accurate when most of the estimates have no error, and the errors have a symmetrical distribution centred at 0. For the three correlations, the error distribution is completely symmetrical, proving a high accuracy of the results and therefore validating the correlations developed. Furthermore, 53% of the estimates lie in intervals of [-0.01 0.01] kWh/m^2 for the correlation corresponding to South-Yorkshire and 51% and 41% for the Norfolk and West-Sussex correlations, respectively.

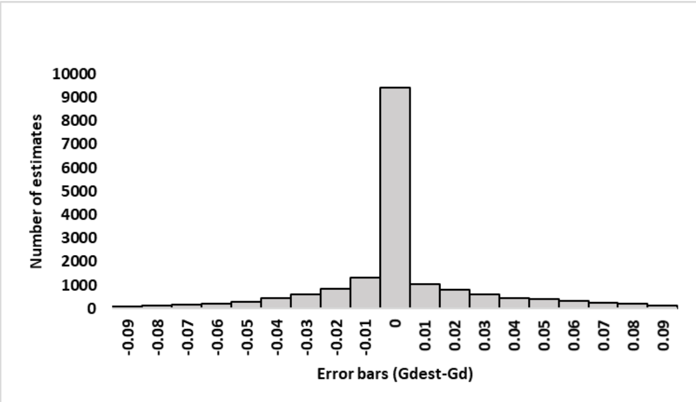

Figure 7. Error histogram for the South-Yorkshire correlation.

Figure 8. Error histogram for the Norfolk correlation.

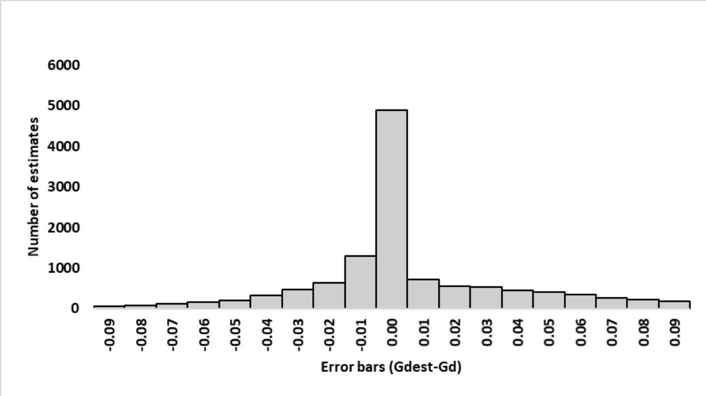

Figure 9. Error histogram for the West-Sussex correlation.

5.2 Comparison between different datasets

A complementary test has been carried out to evaluate the accuracy of the new correlations developed for South-Yorkshire, Norfolk, and West-Sussex. To verify the accuracy of the results, three different datasets have been used for each location. The expectation is that the statistical error metrics should not change significantly between datasets. This would imply that the correlations can be used to accurately predict the horizontal diffuse solar irradiance at the specific location.

To study the accuracy of the developed correlations, first, the existing empirical models have been evaluated with three different datasets. This makes it possible to compare the existing models and the new correlations.

Table 7 provides information on the years used for complete, training and validation datasets for each location.

Table 7. Detailed information of the dataset used as complete, training and validation datasets.

	Complete dataset	**Training dataset**	**Validation dataset**
South-Yorkshire	1982-1995	1982-1989	1990-1995
Norfolk	1982-1999	1982-1993	1994-1999
West-Sussex	1982-1992	1982-1987	1988-1992

The results of existing empirical models for the three locations can be seen in Table 8. It can be seen that the results obtained from the existing models exhibited small variations between datasets.

Table 8. Statistical error metrics obtained for the existing empirical models using different datasets

South Yorkshire									
	Complete dataset			Training dataset			Validation dataset		
	r^2	MBE	RMSE	r^2	MBE	RMSE	r^2	MBE	RMSE
Model 1	0.86	-0.022	0.045	0.862	-0.023	0.046	0.863	-0.0215	0.044
Model 2	0.856	-0.013	0.040	0.857	-0.014	0.041	0.859	-0.0129	0.039
Model 3	0.861	-0.015	0.041	0.859	-0.015	0.042	0.860	-0.0142	0.040
Model 4	0.770	-0.0016	0.048	0.769	-0.002	0.040	0.77	-0.0018	0.040

Norfolk									
	Complete dataset			Training dataset			Validation dataset		
	r^2	MBE	RMSE	r^2	MBE	RMSE	r^2	MBE	RMSE
Model 1	0.839	-0.023	0.048	0.842	-0.024	0.049	0.836	-0.022	0.048
Model 2	0.836	-0.015	0.044	0.838	-0.016	0.044	0.833	-0.014	0.044
Model 3	0.837	-0.017	0.044	0.839	-0.017	0.045	0.833	-0.016	0.044
Model 4	0.789	-0.001	0.039	0.797	-0.002	0.038	0.778	-0.0002	0.039

West Sussex									
	Complete dataset			Training dataset			Validation dataset		
	r^2	MBE	RMSE	r^2	MBE	RMSE	r^2	MBE	RMSE
Model 1	0.829	-0.026	0.051	0.826	-0.019	0.048	0.827	-0.0108	0.033
Model 2	0.823	-0.017	0.047	0.823	-0.018	0.048	0.822	-0.0148	0.045
Model 3	0.825	-0.018	0.047	0.826	-0.019	0.048	0.824	-0.0161	0.045
Model 4	0.771	-0.001	0.041	0.783	-0.003	0.045	0.757	0.00146	0.045

The differences between the training dataset and the validation dataset for the new correlations can be found in Table 9. In general, the performance indicators only show small variations between both datasets, which means that the correlations are accurate and could be used to determine values of horizontal diffuse solar irradiance for the locations studied. To conclude, the new correlations present an improvement if compared with existing models for the regions studied.

Table 9. Accuracy evaluation of developed correlation using different datasets

	Training dataset			**Validation dataset**		
	r^2	MBE	RMSE	r^2	MBE	RMSE
South-Yorkshire	0.877	0.001	0.025	0.872	0.0015	0.027
Norfolk	0.85	0.002	0.033	0.840	0.004	0.04
West-Sussex	0.844	0.002	0.04	0.830	0.003	0.03

6. Conclusions

Solar modelling aims to predict solar energy generation on an annual, monthly, daily, or hourly basis at any location. These models are created to assist engineers on photovoltaics, solar water systems, or in air conditioning applications and in the assessment of solar energy resources. Accurate values of horizontal solar irradiation are essential for these applications. However, the values of horizontal solar irradiation, particularly diffuse horizontal solar irradiation (H_d) are very scarce. Satellite measurements of horizontal solar irradiation are available for any part of the world, but these measurements are not as accurate as ground measured data. This has led to many authors developing H-based empirical correlations to calculate values of horizontal diffuse solar irradiance (G_d) from measured values of global horizontal solar irradiation (H).

In this study, a set of six different existing empirical models using one or more predictors have been compared. Due to the dependency that empirical models have with the location of the dataset used to develop the correlation, only models that were developed with datasets geographically close to the United Kingdom were selected. The models selected were developed in the northern hemisphere in different cities within Europe and the United Kingdom. Each model was evaluated using datasets documented in three regions within the United Kingdom: South-Yorkshire, Norfolk, and West Sussex using data available between 1982 and 1999.

From the comparison of six existing empirical models, it was found that adding more than one predictor to the correlation had diminishing returns in terms of accuracy in the solar irradiation predictions when compared to correlations that only used the clearness index (kt) as a predictor. From the Model 1, Model 2 and Model 3, neither outperformed the others for the three statistical error metrics. However, Model 4 reported the best results for all the studied locations when the full set of kt values were included.

Three new correlations were developed, one for each location studied. These correlations were based on one predictor parameters, the clearness index (kt). The correlations showed a clear improvement if compared with the existing models for all the locations. This demonstrates the necessity to develop empirical correlations for the estimation of horizontal diffuse solar irradiance for a specific location. The new correlations developed were validated using histogram bar errors. The results displayed indicate that the correlations are accurate based on the symmetrical distribution of the errors, with [-0.01 0.01] interval having the largest number of estimates. Finally, each model was also evaluated by comparison between datasets from different years at each location. The results showed minor differences between the training and the validation dataset, proving that the correlations are accurate to predict the horizontal diffuse solar irradiance at the specific location.

CRediT authorship contribution statement

Maria Nunez Munoz: Data curation, Formal analysis, Investigation, Writing—review and editing, Methodology, Validation and Writing—original draft. **Erica E.F. Ballantyne**: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Data curation, Formal analysis, Investigation, Writing—review and editing. **David A. Stone**: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Data curation, Formal analysis, Investigation, Writing—review and editing, Methodology, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC-University of Sheffield). We gratefully acknowledge this support.

Nomenclature

G_0	Horizontal extra-atmospheric irradiance (W/m^2) in time periods of one minute
G	Hourly horizontal extra-atmospheric irradiance (kW/m^2)
G_H	Hourly global horizontal irradiance (kW/m^2)
G_d	Hourly diffuse horizontal irradiance (kW/m^2)
$G_{d,c}$	Hourly diffuse horizontal irradiance on a clear day (kW/m^2)
$G_{d,oc}$	Hourly diffuse horizontal irradiance in an overcast day (kW/m^2)
G_{dest}	Estimated hourly diffuse horizontal irradiance (kW/m^2)
H	Hourly global horizontal irradiation (kJ/m^2)
H_d	Hourly diffuse horizontal irradiation (kJ/m^2)
γ_S	Solar altitude
SF	Sunshine fraction
m	Air mass

7. References

[1] Muneer T, Munawwar S. Improved accuracy models for hourly diffuse solar radiation. J Sol Energy Eng Trans ASME 2006;128:104–17. https://doi.org/10.1115/1.2148972.

[2] Berrizbeitia SE, Gago EJ, Muneer T. Empirical models for the estimation of solar sky-diffuse radiation. A review and experimental analysis. Energies 2020;13. https://doi.org/10.3390/en13030701.

[3] Burnett D, Barbour E, Harrison GP. The UK solar energy resource and the impact of climate change. Renew Energy 2014;71:333–43. https://doi.org/10.1016/j.renene.2014.05.034.

[4] Vignola F, Harlan P, Perez R, Kmiecik M. Analysis of satellite derived beam and global solar radiation data. Sol Energy 2007;81:768–72. https://doi.org/10.1016/j.solener.2006.10.003.

[5] Polo J, Wilbert S, Ruiz-Arias JA, Meyer R, Gueymard C, Súri M, et al. Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. Sol Energy 2016;132:25–37. https://doi.org/10.1016/j.solener.2016.03.001.

[6] Lave M, Weekley A. Comparison of high-frequency solar irradiance: Ground measured vs. satellite-derived. 2017 IEEE 44th Photovolt Spec Conf PVSC 2017 2017:1–6. https://doi.org/10.1109/PVSC.2017.8366211.

[7] Ruiz-Arias JA, Alsamamra H, Tovar-Pescador J, Pozo-Vázquez D. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. Energy Convers Manag 2010;51:881–93. https://doi.org/10.1016/j.enconman.2009.11.024.

[8] Muneer T, Etxebarria S, Gago EJ. Monthly averaged-hourly solar diffuse radiation model for the UK. Build Serv Eng Res Technol 2014;35:573–84. https://doi.org/10.1177/0143624414522639.

[9] Despotovic M, Nedic V, Despotovic D, Cvetanovic S. Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation. Renew Sustain Energy Rev 2016;56:246–60. https://doi.org/10.1016/j.rser.2015.11.058.

[10] Bashahu M. Statistical comparison of models for estimating the monthly average daily diffuse radiation at a subtropical African site. Sol Energy 2003;75:43–51. https://doi.org/10.1016/S0038-092X(03)00213-5.

[11] Gouda SG, Hussein Z, Luo S, Yuan Q. Review of empirical solar radiation models for estimating global solar radiation of various climate zones of China. Prog Phys Geogr 2020;44:168–88. https://doi.org/10.1177/0309133319867213.

[12] Li H, Bu X, Long Z, Zhao L, Ma W. Calculating the diffuse solar radiation in regions without solar radiation measurements. Energy 2012;44:611–5. https://doi.org/10.1016/j.energy.2012.05.033.

[13] Diez FJ, Martínez-Rodríguez A, Navas-Gracia LM, Chico-Santamarta L, Correa-Guimaraes A, Andara R. Estimation of the hourly global solar irradiation on the tilted and oriented plane of photovoltaic solar panels applied to greenhouse production. Agronomy 2021;11. https://doi.org/10.3390/agronomy11030495.

[14] CEDA. CEDA Archive 2017. https://www.ceda.ac.uk/services/ceda-archive/ (accessed March 16, 2021).

[15] Salhi H, Lazhar Belkhiri ·, Tiri · Ammar. Evaluation of diffuse fraction and diffusion coefficient using statistical analysis 2020;10:133. https://doi.org/10.1007/s13201-020-01216-0.

[16] Lou S, Li DHW, Lam JC, Chan WWH. Prediction of diffuse solar irradiance using machine learning and multivariable regression. Appl Energy 2016;181:367–74. https://doi.org/10.1016/j.apenergy.2016.08.093.

[17] Tapakis R, Michaelides S, Charalambides AG. Computations of diffuse fraction of global irradiance: Part 1 – Analytical modelling. Sol Energy 2016;139:723–32. https://doi.org/10.1016/j.solener.2015.12.042.

[18] Liu BYH, Jordan RC. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Sol Energy 1960;4:1–19. https://doi.org/10.1016/0038-092X(60)90062-1.

[19] Erbs DG, Klein SA, Duffie JA. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. Sol Energy 1982;28:293–302. https://doi.org/10.1016/0038-092X(82)90302-4.

[20] Muneer T, Saluja GS. Correlation between hourly diffuse and global solar irradiation for the UK. Build Serv Eng Res Technol 1986;7:37–43. https://doi.org/10.1177/014362448600700106.

[21] Reindl DT, Beckman WA, Duffie JA. Diffuse fraction correlations. Sol Energy 1990;45:1–7.

[22] De Miguel A, Bilbao J, Aguiar R, Kambezidis H, Negro E. Diffuse solar irradiation model evaluation in the North Mediterranean Belt area. Sol Energy 2001;70:143–53. https://doi.org/10.1016/S0038-092X(00)00135-3.

[23] Bailek N, Bouchouicha K, Al-Mostafa Z, El-Shimy M, Aoun N, Slimani A, et al. A new empirical model for forecasting the diffuse solar radiation over Sahara in the Algerian Big South. Renew Energy 2018;117:530–7. https://doi.org/10.1016/j.renene.2017.10.081.

[24] Torres JL, De Blas M, García A, de Francisco A. Comparative study of various models in estimating hourly diffuse solar irradiance. Renew Energy 2010;35:1325–32. https://doi.org/10.1016/j.renene.2009.11.025.

[25] Muneer T, Younes S, Munawwar S. Discourses on solar radiation modeling. Renew Sustain Energy Rev 2007;11:551–602. https://doi.org/10.1016/j.rser.2005.05.006.

[26] Met Office. MIDAS Open: UK hourly solar radiation data, v202007. Cent Environ Data Anal 2020. https://doi.org/10.5285/1dc8578eb7434a7d8a661744d53eedf9.

[27] Gueymard C. Prediction and performance assessment of mean hourly global radiation. Sol Energy 2000;68:285–303. https://doi.org/10.1016/S0038-092X(99)00070-5.

[28] Met Office. UK regional climates - Met Office. MetOffice 2016:1. https://www.metoffice.gov.uk/research/climate/maps-and-data/regional-climates/index (accessed September 21, 2021).

[29] Muneer T, Fairooz F. Quality control of solar radiation and sunshine measurements - lessons learnt from processing worldwide databases. Build Serv Eng Res Technol 2002;23:151–66. https://doi.org/10.1191/0143624402bt038oa.

[30] Duffie JA, Beckman WA. Solar Radiation. Sol. Eng. Therm. Process. 4th ed., John

Wiley & Sons, Incorporated; 2013.

[31] Kittler R, Darula S. Determination of time and sun position system. Sol Energy 2013;93:72–9. https://doi.org/10.1016/j.solener.2013.03.021.

[32] Cooper P. The Absorption of Solar Radiation in Solar Stills. Sol Energy 1969;12:333–46. https://doi.org/10.1016/0038-092X(69)90047-4.

[33] Sunter M. MIDAS Data User Guide for UK Land Observations. 2021.

[34] Brownson JRS. Sun-Earth Geometry. Sol. Energy Convers. Syst., Elsevier Science & Technology, Saint Louis; 2014.

[35] Younes S, Claywell R, Muneer T. Quality control of solar radiation data: Present status and proposed new approaches. Energy 2005;30:1533–49. https://doi.org/10.1016/j.energy.2004.04.031.

[36] Chaâbane M, Masmoudi M, Medhioub K. Determination of Linke turbidity factor from solar radiation measurement in northern Tunisia. Renew Energy 2004;29:2065–76. https://doi.org/10.1016/j.renene.2004.03.002.

[37] Remund J, Wald L, Lefèvre M, Ranchin T, Page J. Worldwide Linke Turbdidity Information. Ises 2003 2003;41.

[38] Huang KT. Identifying a suitable hourly solar diffuse fraction model to generate the typical meteorological year for building energy simulation application. Renew Energy 2020;157:1102–15. https://doi.org/10.1016/j.renene.2020.05.094.

[39] Khorasanizadeh H, Mohammadi K, Goudarzi N. Prediction of horizontal diffuse solar radiation using clearness index based empirical models; A case study. Int J Hydrogen Energy 2016;41:21888–98. https://doi.org/10.1016/j.ijhydene.2016.09.198.

[40] Hofmann M, Seckmeyer G. A new model for estimating the diffuse fraction of solar irradiance for photovoltaic system simulations. Energies 2017;10. https://doi.org/10.3390/en10020248.

[41] Yao W, Zhang C, Wang X, Sheng J, Zhu Y, Zhang S. The research of new daily diffuse solar radiation models modified by air quality index (AQI) in the region with heavy fog and haze. Energy Convers Manag 2017;139:140–50. https://doi.org/10.1016/j.enconman.2017.02.041.

[42] Pelland S, Remund J, Kleissl J, Oozeki T, De Brabandere K. Photovoltaic and Solar Forecasting : State of the Art. 2013.

[43] Mahajan B, Namrata K. Performance evaluation of developed empirical models for predicting global solar radiation in western region of India. Int J Renew Energy Res 2019;9:1135–43.

[44] Jiang Y. Estimation of monthly mean daily diffuse radiation in China. Appl Energy 2009;86:1458–64. https://doi.org/10.1016/j.apenergy.2009.01.002.

[45] David M, Diagne HM, Lauret P. Outputs and error indicators for solar forecasting models. World Renew Energy Forum, WREF 2012, Incl World Renew Energy Congr XII Color Renew Energy Soc Annu Conf 2012;2:831–6.

[46] Notton G, Voyant C. Forecasting of Intermittent Solar Energy Resource. In: Yahyaoui I, editor. Adv. Renew. Energies Power Technol., Elsevier; 2018, p. 77–114. https://doi.org/10.1016/B978-0-12-812959-3.00003-4.

[47] Theristis M, Venizelou V, Makrides G, Georghiou GE. Energy Yield in Photovoltaic

Systems. In: Kalogirou S., editor. McEvoy's Handb. Photovoltaics, Elsevier; 2018, p. 671–713. https://doi.org/10.1016/B978-0-12-809921-6.00017-3.

[48] Kambezidis HD. The Solar Resource. In: Sayigh A, editor. Compr. Renew. Energy, Elsevier; 2012, p. 27–84. https://doi.org/10.1016/B978-0-08-087872-0.00302-4.

[49] Hoff TE, Perez R, Kleissl J, Renne D, Stein J. Reporting of irradiance modeling relative prediction errors. Prog Photovoltaics Res Appl 2013;21:1514–9. https://doi.org/10.1002/pip.2225.