Assessment and development of diffuse irradiance models for application within the UK

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: SMALE, N., 2012. Assessment and development of diffuse irradiance models for application within the UK. Loughborough University MSc Project Report.

Additional Information:

- Submitted in partial fulfilment of MSc Renewable Energy Systems Technology.

Metadata Record: [https://dspace.lboro.ac.uk/2134/25371](https://dspace.lboro.ac.uk/2134/25371)

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: [https://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Please cite the published version.
Assessment and Development of Diffuse Irradiance Models for Application within the UK

Nicholas Smale, Loughborough University, Loughborough, LE11 3TU

1. ABSTRACT

The ratio of direct to diffuse radiation directly impacts on solar system design. Currently no diffuse fraction models have been specifically developed for the UK; development of existing global models could lead to increased accuracy within Loughborough and the UK. Data from Loughborough was used to test existing global models and linear square regression used to optimise model coefficients. Two successfully optimised models were combined with sensitivity testing to develop a much simplified Loughborough and UK specific diffuse fraction model; this model reduced error over global models by 75% at the Loughborough test site and 25% at a second test site in Lerwick. It has been demonstrated that the model developed within this study should be used within the UK for diffuse fraction prediction. The methodology used has been shown to be beneficial in characterising sites; this may enable enhanced short term prediction of the diffuse fraction at measurement sites, depending on local environmental factors such as humidity and solar elevation angle.

2. INTRODUCTION

Accurately predicting solar irradiation is critical in the development of solar energy systems across the UK. Commonly only global horizontal irradiation data is available for sites, necessitating the application of diffuse fraction and inclination models to account for irradiation composition and solar module angle. Knowing the ratio of diffuse to direct irradiation allows designers to tailor solar PV systems for increased performance. Existing diffuse fraction models have been developed globally; this study aims to develop an accurate model for application at Loughborough and sites throughout the UK.

The founding work of Liu and Jordan [1] forms the basis for the predominant relationship underpinning all existing diffuse fraction models. Fundamentally a relationship between the clearness index \((k_t)\) and diffuse fraction \((K_d)\) is identified and used to predict diffuse fraction values from clearness index values. Clearness index is defined as the ratio of global solar radiation to the extraterrestrial horizontal solar radiation and diffuse fraction is defined as the ratio of diffuse solar radiation to global solar radiation.

The work of Torres et al [2] showed the majority of existing models fit a standard polynomial style equation (Equation 1) with prediction of the diffuse fraction split between three distinct clearness index zones. Torres evaluated nine separate studies with differing coefficients and showed that these models produce a strong correlation for hourly data values.

\[
\begin{align*}
K_d &= a_1 + a_2 k_t \quad (\text{Zone 1} \sim 0 \leq k_t \leq 0.3) \\
K_d &= a_3 + a_4 k_t \quad (\text{Zone 2} \sim 0.3 \leq k_t \leq 0.8) \\
K_d &= a_5 \quad (\text{Zone 3} \sim 0.78 \leq k_t)
\end{align*}
\]

Equation 1 – Typical polynomial diffuse fraction model
(Where \(K_d = \text{Diffuse Fraction} \quad a_x = \text{Coefficient} \quad k_t = \text{Clearness Index}\)

Jacovides et al [3] conducted a study similar to Torres and found the majority of existing polynomial models performed similarly; from this it was concluded that diffuse fraction models must be location independent. However, comparison of the performance of identical models from the Jacovides (Cyprus) and Torres (Spain) (Table 1) studies contradict these findings showing significant differences in model performance, suggesting location dependant factors which are not accounted for within the clearness index may affect the diffuse fraction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Orgill and Hollands [4]</td>
<td>0.747</td>
<td>0.902</td>
</tr>
<tr>
<td>Reindl et al [5]</td>
<td>0.747</td>
<td>0.901</td>
</tr>
<tr>
<td>Hawlader [6]</td>
<td>0.724</td>
<td>0.902</td>
</tr>
<tr>
<td>Miguel et al [7]</td>
<td>0.748</td>
<td>0.901</td>
</tr>
<tr>
<td>Karatasou et al [8]</td>
<td>0.712</td>
<td>0.903</td>
</tr>
<tr>
<td>Erbs et al [9]</td>
<td>0.737</td>
<td>0.899</td>
</tr>
<tr>
<td>Oliveira et al [10]</td>
<td>0.728</td>
<td>0.898</td>
</tr>
<tr>
<td>BRL [11]</td>
<td>0.787</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Reindl et al [5] attempted to include predictors other than clearness index within ‘dynamic’ diffuse fraction estimation models to better account for location dependant characteristics such as humidity, temperature and solar elevation (Equation 2). This approach led to improved statistical performance based on a more complex version of the traditional polynomial model.

\[
0 \leq k_t \leq 0.3 \quad K_d = 1.00 - 0.232k_t + 0.0239\sin(\alpha) - 0.000682T_a + 0.0195\phi \\
0.3 \leq k_t \leq 0.78 \quad K_d = 1.329 - 1.716k_t + 0.267\sin(\alpha) - 0.00357T_a + 0.106\phi \\
0.78 \leq k_t \leq 0.78 \quad K_d = 0.426k_t - 0.256\sin(\alpha) + 0.00349T_a + 0.0734\phi
\]

Equation 2 – Reindl ‘Dynamic’ model [12]

Where \( k_t = \text{Clearness Index} \), \( K_d = \text{Diffuse Fraction} \), \( \alpha = \text{Solar Elevation (°)} \), \( T_a = \text{Ambient Temperature (°C)} \), \( \phi = \text{Humidity Fraction} \)

A continued major drawback of the polynomial style model is the increased complexity involved in splitting the clearness index into zones. Ridley et al [11] removed the requirement for clearness index zones by applying a logistic modelling approach (Equation 3). The model developed by Ridley known as the BRL model introduced the concept of ‘persistence’ into diffuse fraction modelling by incorporating daily average clearness index values in addition to leading and lagging hourly clearness index values. This approach outperformed all existing models for both the test site and, importantly, tests sites around the world at differing locations, suggesting the BRL model inclusive of a logistic function, elevation angle and ‘persistence’ was the most successful existing model for diffuse fraction prediction.

\[
K_d = \frac{1}{1 + e^{(-5.38+6.63k_t+0.006\text{AST}+0.007\alpha+1.75K_t+1.31\phi)}}
\]


Where \( K_d = \text{Diffuse Fraction} \), \( k_t = \text{Clearness Index} \), \( \text{AST} = \text{Apparent Solar Time (24 hour)} \), \( \alpha = \text{Solar Elevation (°)} \), \( K_t = \text{Daily Average Clearness Index} \), \( \phi = \text{Hourly Persistence Index} \)

Ridley, Reindl and Skartveit [12] all concluded that elevation angle was a significant parameter affecting the diffuse fraction in ‘dynamic’ diffuse fraction models. Skartveit concluded that elevation affects were greatest below 30 degrees whilst Reindl concluded, under high clearness index values, elevation angle became the predominant parameter ahead of clearness index; it is evident that elevation angle may affect the diffuse fraction but these contradicting findings make it difficult to predict the contribution of the elevation angle within Loughborough data.

Equations 1, 2 and 3 were selected for analysis of Loughborough data as this combination of equations offer a range of complexity, parameters and ideology to be investigated within this study. Prior to this study diffuse fraction modelling had not been investigated in detail for the UK and by developing a model specifically for Loughborough; diffuse fraction prediction should become more accurate. Improved diffuse fraction modelling will improve the accuracy of plane translation models used to estimate irradiance on inclined surfaces; this will in turn improve the validity of solar energy feasibility studies within the UK. By removing uncertainty within energy yield prediction the returns and costs of solar energy should become more predictable.

The central aims and objectives of this study were:

1. To accurately assess global diffuse fraction models for application within the UK
2. To model the Loughborough data set using the most successful global models
3. To develop a model with increased accuracy for Loughborough data
4. To compare the accuracy of a new model for data across the UK

3. METHODOLOGY

Literature Review

A literature review was conducted of existing diffuse fraction models in order to assess their application within the UK. It was evident that a wide range of statistical parameters, data points, exclusion rules and data averaging had been used for each study. Where possible a numerical comparison of each study was made and the results are detailed in Table 1. In scenarios where results were not directly comparable, it was necessary to match studies to the data, methods and data available at Loughborough. Generally each study was assessed for validity against the following criteria; number of data points, data exclusion criteria, averaging period for each data point and model performance.
Data Sorting

In order to apply diffuse fraction modelling to Loughborough data it was necessary to evaluate the full 2010-2011 data set. An initial plot of the diffuse fraction plotted against clearness index showed numerous errors with spurious diffuse fraction values and many null values as displayed in Figure 1.

A review of multiple diffuse fraction studies indicated that the data exclusion rules of Jacovides [3] provided the most comprehensive explanation of data validation. The following exclusion rules were applied to the Loughborough 2010-2011 data set. \( G_d > 1.1 G_h \), \( G_h > 1.2 G_{h,oa} \), \( G_d > 0.8 G_{h,oa} \), \( G_h < 20 \text{ Wm}^{-2} \), \( G_B > G_{h,oa} \), \( G_d > 0.9 \) (\( k_t > 0.6 \)). The following exclusion rule \( \frac{G_d}{G_h} < 0.9 \) (\( k_t < 0.2 \)) was omitted as it did not fit with the general trend of data; valid data at low clearness index values is excluded by the omitted rule, as observed when comparing Figure 2 and Figure 3.

Jacovides recommended a minimum global irradiance reading of 5 Wm\(^{-2}\); by amending to 20 Wm\(^{-2}\) inherent errors within the Pyranometer were reduced. Readings with an elevation angle of less than 10° were also eliminated. An original data set of 17000 hourly data values was reduced to 6193 points (Figure 3), the shape and distribution closely resembling plots in studies by Torres.

Statistical Analysis

To choose a set of statistical parameters for model evaluation, the work of Reindl and Ridley was analysed. Mean Absolute Percentage Error (MeAPE) and Residual Sum of Squares (RSS) were selected as these parameters allow comparison with the findings of this study. In addition, the coefficient of determination \( (R^2) \) was chosen to enable comparison with the work of Torres and Jacovides and Mean Percentage Error (MPE) was chosen to enable potential improvements in accuracy to be quantified in Watts.

\[
\text{MeAPE} = \frac{\sum (G_{dm} - G_{dp})}{G_{dm}} \times 100
\]
\( \text{Equation 4 – Mean Absolute Percentage Error} \)

\[
\text{MPE} = \frac{\sum (G_{dm} - G_{dp})}{G_{dm}} \times 100
\]
\( \text{Equation 5 – Mean Percentage Error} \)

\[
R^2 = 1 - \frac{\sum (G_{dmi} - G_{dp})^2}{\sum (G_{dmi} - G_{dm})^2}
\]
\( \text{Equation 6 – Coefficient of determination} \)

\[
\text{RSS} = \sum (G_{dm} - G_{dp})^2
\]
\( \text{Equation 7 – Residual Sum of Squares} \)

(Where \( G_{dm} = \text{Measured diffuse fraction} \) \( G_{dp} = \text{Predicted diffuse fraction} \))
Total annual error of each model in \( \text{W} \text{h}^{-2} \) was calculated to quantify model performance. An average of the annual total diffuse irradiance for each data set was multiplied by the Mean Percentage Error of each respective model. Total annual diffuse fraction for Loughborough was 625 \( \text{kHz} \text{h}^{-2} \) and Lerwick 459 \( \text{kHz} \text{h}^{-2} \).

### Application of Diffuse Fraction Models

The BRL model uses persistence (\( \psi \)) and daily average clearness index (\( K_d \)) parameters to allow for trends within data. Daily average clearness index values (\( K_d \)) were simply calculated as an average of all valid values in each day and the persistence index (\( \psi \)) calculated from hourly values (Equation 8):

\[
\psi = \left( \frac{k_{t-1} + k_{t+1}}{2} \right) \quad \text{sunrise} < t < \text{set sun}
\]

\[
\psi = k_{t+1} \quad t = \text{First daily value}
\]

\[
\psi = k_{t-1} \quad t = \text{Last daily value}
\]

**Equation 8 – Persistence index**

(Where \( \psi = \) Hourly Persistence index \( k_t = \) Clearness Index)

To determine radiation at the earth’s atmosphere (\( G_{h,oa} \)) calculations were completed using Equation 9 taken from [13].

\[
G_{h,oa} = E_o \cdot |\sin \alpha|
\]

**Equation 9 – Radiation at the earth’s atmosphere**

(Where \( E_o = \) Eccentricity factor \( k_o = \) Solar constant \( 1367 \text{W/m}^2 \) \( \alpha = \) Solar Elevation (\( \gamma \))

### Model Optimisation

In order to optimise existing diffuse fraction models for Loughborough, the parameter coefficients in the three studied equations were adjusted. The SOLVER function within Microsoft Excel was utilised to complete least squares regression and establish coefficients which minimised the total error of each predictive model.

To establish the contribution to accuracy of each parameter within the Ridley and Reindl equations, sensitivity testing was completed. Parameters were systematically removed from each respective equation and SOLVER used to re-optimise the simplified equation, the performance of the whole and ‘simplified’ equation was then compared to establish the change in accuracy attributable to each parameter.

### Uncertainty Analysis

It was identified that errors potentially existed within the measurement system. Data was collected via a Delta-T Instruments BF3 sensor array, which had up to \( \sim \pm 9\% \) error for hourly averaged data [14]. To mitigate the affect of measurement error a data filter was included to remove readings with a global horizontal reading of less than 20W/m², this was higher than any other study evaluated and deemed a reasonable level given the poor performance of array inverters below 50W. The potential compounded errors identified above could mask the trends identified within the analysis of this study; however the high number of data points and consistent trends observed should mitigate this risk.

All other previous studies have completed analysis with similar data and a similar methodology.

### Model Iteration

All the models developed within this study are derivatives of Equations 1, 2 & 3. Details of each model referenced within this study can be found in Table 2.

<table>
<thead>
<tr>
<th>Number</th>
<th>Model</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reindl Polynomial</td>
<td>Standard Reindl polynomial model [3] (Equation 1)</td>
</tr>
<tr>
<td>2</td>
<td>Reindl Dynamic</td>
<td>Standard Reindl ‘dynamic’ model [5] (Equation 2)</td>
</tr>
<tr>
<td>4</td>
<td>Reindl Polynomial Loughborough</td>
<td>Model 1 optimised for Loughborough data</td>
</tr>
<tr>
<td>5</td>
<td>Reindl Dynamic Loughborough</td>
<td>Model 2 optimised for Loughborough data</td>
</tr>
<tr>
<td>6</td>
<td>BRL Loughborough</td>
<td>Model 3 optimised for Loughborough data</td>
</tr>
<tr>
<td>7</td>
<td>BRL Loughborough no AST</td>
<td>Model 6 optimised with the Apparent Solar Time parameter removed</td>
</tr>
<tr>
<td>8</td>
<td>BRL Loughborough no ( \bar{K}_t )</td>
<td>Model 6 optimised with the daily average clearness index parameter removed</td>
</tr>
<tr>
<td>9</td>
<td>BRL Loughborough no ( \bar{a} )</td>
<td>Model 6 optimised with the elevation parameter removed</td>
</tr>
<tr>
<td>10</td>
<td>BRL Loughborough no ( \psi )</td>
<td>Model 6 optimised with the hourly persistence parameter removed</td>
</tr>
<tr>
<td>11</td>
<td>Reindl Dynamic Loughborough no ( \psi )</td>
<td>Model 5 optimised with the elevation parameter removed</td>
</tr>
<tr>
<td>12</td>
<td>Reindl Loughborough no ( \bar{a} )</td>
<td>Model 5 optimised with the humidity parameter removed</td>
</tr>
<tr>
<td>13</td>
<td>Reindl Loughborough no ( T_s )</td>
<td>Model 5 optimised with the temperature parameter removed</td>
</tr>
<tr>
<td>14</td>
<td>BRL Loughborough with ( \bar{a} )</td>
<td>Model 6 with a humidity parameter added (Equation 10)</td>
</tr>
<tr>
<td>15</td>
<td>BRL Loughborough simplified with ( \bar{a} )</td>
<td>Model 6 with humidity added and temperature, elevation and persistence parameters removed (Equation 11)</td>
</tr>
<tr>
<td>16</td>
<td>BRL Loughborough simplified</td>
<td>Model 6 with temperature, elevation and persistence parameters removed</td>
</tr>
</tbody>
</table>

Table 2 – List and definition of models used within this study
4. RESULTS AND DISCUSSION

Modelling of Loughborough diffuse fraction

To establish the performance of existing diffuse fraction models at Loughborough, the Reindl Polynomial (Model 1), Reindl Dynamic (Model 2) and BRL (Model 3) models were applied to the Loughborough data set. The initial data comparison in Figures 4-6 display predicted diffuse fraction values against corresponding measured values. All three models broadly fit within the measured data however, model differences are clear. The Reindl Dynamic and BRL models display a range of diffuse fraction values for equivalent clearness index values, this is a result of multiple input parameters. It is clear the Reindl Dynamic and BRL models produce a better fit than the basic polynomial model with a distribution of values broadly mirroring the trends within the measured data. It is also clear that both Reindl models are split into three separate zones, this produces abrupt changes in data prediction and does not mirror the data trends well. The BRL model produces a more continuous and improved data fit due to its logistic function.

Figure 7 shows the average of the predicted and measured values sorted into bins with 0.05Kt intervals. It is clear that all three predictive models track the average of the measured values up to clearness index values of 0.8. Above 0.8 the models diverge, a result of the low number of readings and the lack of consistency in the measured data. Variability in the diffuse fraction on clear days may also be attributed to increased albedo affects, as was concluded by Skartveit.

![Figure 4 – Reindl dynamic (Model 2) comparison](image)

![Figure 5 – BRL (Model 3) comparison](image)

![Figure 6 – Reindl polynomial (Model 1) comparison](image)

![Figure 7 – Averaged comparison](image)

Performance of the Reindl Dynamic, BRL and Reindl Polynomial models fitted to Loughborough data can be seen in Table 3 where it is broadly in line with that observed within the Torres and Jacovides studies. The basic polynomial model marginally outperforms the more complex dynamic models. The lower error Polynomial model is recommended for use when the possibility of modifying coefficients within more complex dynamic models is not available.

<table>
<thead>
<tr>
<th>Model</th>
<th>MeAPE</th>
<th>MePE</th>
<th>RSS</th>
<th>R^2</th>
<th>Error (kWh/m^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reindl Dynamic – Model 2</td>
<td>24.69</td>
<td>-14.82</td>
<td>118.72</td>
<td>0.791</td>
<td>92.63</td>
</tr>
<tr>
<td>BRL – Model 3</td>
<td>24.33</td>
<td>-13.35</td>
<td>121.56</td>
<td>0.794</td>
<td>83.44</td>
</tr>
<tr>
<td>Reindl Polynomial - Model 1</td>
<td>23.46</td>
<td>-11.73</td>
<td>117.49</td>
<td>0.790</td>
<td>73.31</td>
</tr>
</tbody>
</table>
Optimisation of existing models for Loughborough conditions

To optimise the Reindl Dynamic, BRL and Reindl Polynomial models for Loughborough data, Least Squares Regression was used to re-define the equation coefficients to form models 4-6. Original and ‘optimised’ coefficients are available in Tables 4, 5 and 6. During regression analysis the starting values for each coefficient was equal to its original value.

For the optimised models with location specific coefficients (4-6), performance increased markedly over global models. Comparison of Tables 3 and 7 reveals all three models have drastically increased performance with lower mean percentage errors reducing the predicted model error by at least 30%. The more basic polynomial style model displays the smallest increase in performance whilst the more complex ‘dynamic’ models, especially model 6 the BRL style model, reduces total prediction error by 80% from 83.44 to 17.63 kWh/m².

The increased performance of the optimised dynamic models can be attributed to the extra parameters present within the equations; setting coefficients for location dependant parameters such as humidity, temperature and persistence should enhance performance for individual locations. The biggest benefit of ‘optimising’ the coefficients in models 4-6 is the change in data distribution. Figures 8 to 11 show predicted values closer to the centre of the measured values illustrating a reduction in prediction error. The distribution of predicted data in model 6, the optimised BRL model (Figure 9), closely matches the distribution of measured data; an improvement of the global BRL model in Figure 5. The performance of model 6, the optimised BRL model is clearly the best of the six models tested so far. The performance of this BRL style model is indicative that the parameters of time, elevation and persistence are influential in the diffuse fraction component for the Loughborough climate.
Sensitivity testing of model parameters

To further increase the performance of model 6, sensitivity testing was undertaken to establish the contribution of each ‘dynamic’ parameter used in the Reindl and BRL style models. Each parameter was removed from the respective equation and the remaining parameter coefficients re-optimised in a manner identical to that used in the initial optimisation process, both dynamic models were interrogated to produce models 7-13.

It is clear from Tables 8 & 7 that removing elevation as an indicator from both dynamic models only marginally reduces performance; contrary to the conclusions of Skartveit, Reindl and numerous authors who found elevation to be an important indicator in diffuse fraction prediction. Similarly, temperature and hourly persistence parameters also only marginally affect the overall performance of the optimised dynamic models. Given the small contribution to accuracy and the added complexity in measuring and/or calculating these parameters, they could be removed.

<table>
<thead>
<tr>
<th>Model</th>
<th>MeAPE</th>
<th>MePE</th>
<th>RSS</th>
<th>R^2</th>
<th>Error (kWh/m^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRL Loughborough no AST - Model 7</td>
<td>17.66</td>
<td>-3.83</td>
<td>71.81</td>
<td>0.855</td>
<td>23.94</td>
</tr>
<tr>
<td>BRL Loughborough no K_t - Model 8</td>
<td>16.42</td>
<td>-3.56</td>
<td>60.56</td>
<td>0.878</td>
<td>22.25</td>
</tr>
<tr>
<td>BRL Loughborough no α - Model 9</td>
<td>15.38</td>
<td>-2.97</td>
<td>54.96</td>
<td>0.889</td>
<td>18.56</td>
</tr>
<tr>
<td>BRL Loughborough no ϕ - Model 10</td>
<td>15.28</td>
<td>-2.83</td>
<td>53.97</td>
<td>0.891</td>
<td>17.70</td>
</tr>
<tr>
<td>Reindl Dynamic Loughborough no α - Model 11</td>
<td>17.51</td>
<td>-5.61</td>
<td>69.52</td>
<td>0.859</td>
<td>35.06</td>
</tr>
<tr>
<td>Reindl Loughborough no ϕ - Model 12</td>
<td>19.82</td>
<td>-7.03</td>
<td>86.06</td>
<td>0.826</td>
<td>43.93</td>
</tr>
<tr>
<td>Reindl Loughborough no T_a - Model 13</td>
<td>17.31</td>
<td>-5.57</td>
<td>68.58</td>
<td>0.861</td>
<td>34.78</td>
</tr>
</tbody>
</table>

Elevation, temperature and hourly persistence parameters do not significantly affect the diffuse fraction at Loughborough; however they do affect the diffuse fraction at other sites. This observation mirrors the conflicting conclusions made regarding the diffuse fraction in many studies; suggesting parameters vary in significance depending on location.

The results in Table 8 reveal the high performance of model 6, the optimised BRL model, can be attributed to the Apparent Solar Time (AST) and daily Clearness Index parameters (K_t). Their inclusion reduces the model error by 25% and 20% respectively. It can be concluded that diffuse fraction values for any given day are likely to vary little from the daily average. There must exist a relationship between solar time and the diffuse fraction, a relationship revealed in Figure 12. The diffuse fraction tends to climb throughout the day perhaps reflecting regular weather patterns within Loughborough; an instance where a parameter is likely to be highly location dependant.

Figure 12 – Daily variation of average diffuse fraction (Kd)
Development by addition and removal of parameters

Table 8 reveals that the humidity parameter of the Reindl Dynamic model has the largest contribution to performance of all tested parameters. Humidity however, is not included within model 6, the optimised BRL model. By including humidity as an indicator it may be possible to further increase the accuracy of model 6; to test this theory Equation 10 (model 14) was developed and the coefficients re-optimised.

\[
K_d = \frac{1}{1 + e^{(-5.821905+4.972504k_t+0.114471AST+0.00564\alpha+2.039125K_t+1.418471\varphi-0.4109\emptyset)}}
\]

Equation 10 – Model 14

\( K_d = \text{Diffuse Fraction} \)
\( k_t = \text{Clearness Index} \)
\( AST = \text{Apparent Solar Time (24 hour)} \)
\( \alpha = \text{Solar Elevation (°)} \)
\( K_t = \text{Daily Average Clearness Index} \)
\( \varphi = \text{Hourly Persistence index} \)
\( \emptyset = \text{Humidity Fraction} \)

In addition to adding a humidity parameter to a BRL style equation (model 14) two further models were tested. Model 15, a BRL style equation, with humidity added and hourly persistence and elevation removed, to test the overall effect of removing parameters with little contribution and adding a new parameter thought to have a large contribution. Model 16, a BRL style model, simplified by removal of hourly persistence and elevation to test a simplified version of the standard BRL style equation. The optimised coefficients of models 14-16 can be seen in Table 9 and their respective performance in Table 10.

<table>
<thead>
<tr>
<th>BRL Coefficients</th>
<th>BRL Loughborough humidity (14)</th>
<th>BRL Loughborough simplified humidity (15)</th>
<th>BRL Loughborough simplified (16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-5.8219</td>
<td>-5.2638</td>
<td>-6.1627</td>
</tr>
<tr>
<td>( k_t )</td>
<td>4.9725</td>
<td>6.2313</td>
<td>6.3965</td>
</tr>
<tr>
<td>( AST )</td>
<td>0.1145</td>
<td>0.1068</td>
<td>0.1288</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>-0.0056</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>( K_t )</td>
<td>2.0391</td>
<td>2.0444</td>
<td>2.1383</td>
</tr>
<tr>
<td>( \varphi )</td>
<td>1.4185</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>( \emptyset )</td>
<td>-0.4109</td>
<td>-0.7720</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 10 - Comparison of models with parameters modified

<table>
<thead>
<tr>
<th>Model</th>
<th>MeAPE</th>
<th>MePE</th>
<th>RSS</th>
<th>R^2</th>
<th>Error (kWh/m^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRL Loughborough with ( \emptyset ) - Model 14</td>
<td>15.05</td>
<td>-2.84</td>
<td>52.70</td>
<td>0.894</td>
<td>17.76</td>
</tr>
<tr>
<td>BRL Loughborough simplified with ( \emptyset ) - Model 15</td>
<td>15.22</td>
<td>-2.94</td>
<td>54.32</td>
<td>0.890</td>
<td>18.37</td>
</tr>
<tr>
<td>BRL Loughborough simplified - Model 16</td>
<td>15.50</td>
<td>-2.95</td>
<td>55.70</td>
<td>0.888</td>
<td>18.46</td>
</tr>
</tbody>
</table>

Analysis of model 14 revealed that adding humidity improved correlation patterns and reduced the total error whilst adversely increasing the mean error in comparison to model 6, the optimised BRL model. This negative impact on mean error was however reversed in the analysis of the two simplified models where the simplified model with humidity (model 15) outperformed model 16, a simplified BRL model with no humidity parameter.

Overall the simplified equation with humidity (model 15) performed marginally worse than the more complex full BRL style model 14, however on balance the benefits of removing two parameters justified the small drop in performance. From these findings model 15 (Equation 11) was selected as the best model for diffuse fraction prediction at Loughborough. Figure 13 shows a comparison of model 15 against the measured diffuse fraction values.

\[
K_d = \frac{1}{1 + e^{(-5.26384+6.23133k_t+0.10676AST+2.044K_t-0.7720)}}
\]

Equation 11 – Proposed optimum model for Loughborough, model 15

\( K_d = \text{Diffuse Fraction} \)
\( k_t = \text{Clearness Index} \)
\( AST = \text{Apparent Solar Time (24 hour)} \)
\( K_t = \text{Daily Average Clearness Index} \)
\( \emptyset = \text{Humidity Fraction} \)
Fig 14 shows the average MeAPE of model 15 plotted against the diffuse fraction; the error is largest in clear sky conditions, a feature inherent within the BRL logistic function and not easily solved without detriment to the performance of the basic model.

Comparison of the chosen Loughborough model against UK data

The proposed model for Loughborough, model 15, was tested against data from Lerwick to evaluate its applicability across the UK. By treating the Lerwick data to identical exclusions rules as the Loughborough data, 2870 hourly values were used from 2002. Models 1, 3, 6 and 15 were applied to the Lerwick data and compared to the measured values, similar to the analysis of the Loughborough data.

It is clear from Table 10 the optimised BRL models developed for Loughborough (models 6 and 15), outperform the generic global models put forward by Ridley and Reindl (models 1 and 3), this may be because location dependant factors such as weather patterns in Lerwick are more similar to Loughborough than the sites used to develop the ‘global’ models.

It is likely that the model proposed by this study will predict the diffuse fraction within the UK more accurately than any existing global diffuse fraction models; this is because the model has been optimised for location dependant factors which are likely to be relatively similar throughout the UK.

Comparison of coefficients calculated on a yearly basis

By optimising equation coefficients it is possible to greatly improve diffuse fraction predication. If it were possible to group coefficients relative to climate classifications then it may be possible to move forward from a standard set of global coefficients used in existing models. For this method to be applicable however, it would be necessary to show coefficients are specific to location and not individual data sets. To test this hypothesis, data from separate years at Lerwick and Camborne was used to generate coefficients using SOLVER. The results in Table 11 show some trends with no overlap of values for the two separate sites.

The results of Table 11 show there may be merit in developing a table of coefficients based on climate classification; this would enable designers to select climate specific values for their site of interest.
5 CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER WORK

Existing global diffuse fraction models perform well within the UK with performance equal to existing European studies. Model performance however can be significantly improved by optimisation of parameter coefficients; this approach can reduce error for specific sites by up to 75%. Models optimised for Loughborough outperformed global models for additional UK sites, suggesting factors affecting the diffuse fraction are relatively consistent throughout the UK, from this it can be concluded that the models developed within this study are the best suited for further use within the UK. The findings of this study have shown that equation coefficients may be specific to climate types; a study matching coefficients to climate classifications may yield further improvements in diffuse fraction prediction.

Through sensitivity testing, it is possible to rank the ‘contribution’ of parameters for specific sites. Thus, characterising a site where data is recorded will allow more accurate short term prediction of the diffuse fraction given changes in high ‘contribution’ parameters such as humidity or clearness index; this capability could increase the confidence in power provision from solar farms.

6 ACKNOWLEDGEMENTS

The author would like to thank Dr Tom Betts for his guidance and advice whilst supervising this study and Charles Greenwood for maintenance of the measurement system that provided the research data.

7 REFERENCES

[2] J.L. Torres a,*, M. De Blas a, A. Garca’a a, A. de Francisco b Comparative study of various models in estimating hourly diffuse solar irradiance Department of Projects and Rural Engineering, Public University of Navarre, Edificio Los Olivos, 31006 Pamplona, Spain & Department of Agricultural and Forestry Engineering, Polytechnic University of Madrid, 28080 Madrid, Spain
[13] Study Notes 1 2010-2011, Crest MSc in Renewable Energy Systems technology, Loughborough University, United Kingdom
[14] McKenna, E, Solar Radiation data Validation, MSc Project, Loughborough University, 2008-2009

8 NOMENCLATURE

K_d = Diffuse Fraction  
K_t = Clearness Index  
a_x = Coefficient  
α = Solar Elevation (°)  
T_a = Ambient Temperature (°C)  
∅ = Humidity Fraction  
K_d=Daily average Clearness Index  
K_t=Daily clearness Index  
φ = Hourly persistence index  
G_d = Global horizontal diffuse irradiance (Wm^{-2})  
G_b = Global horizontal beam irradiance (Wm^{-2})  
G_m = Global horizontal irradiance (Wm^{-2})  
E_o=Eccentricity factor  
G_{t,a} = Irradiation at the earths atmosphere(Wm^{-2})  
G_{dm} = Measured diffuse fraction  
G_{dp} = Predicted diffuse fraction  
φ = Hourly persistence index  
I_o=Solar constant (1367 Wm^{-2})