Solar Radiation Data Validation

Eoghan James McKenna
Loughborough University
Loughborough, Leics. LE11 3TU

A paper for publication in *Solar Energy*

Abstract: This paper describes and applies a procedure for the quality control and validation of solar radiation data for two independent co-located measurement systems based at Loughborough University, United Kingdom. An assessment of the measurement error of simultaneous data from four pyranometers was undertaken over a range of averaging periods. A data filter of 0-1500W/m^2 was found to reduce measurement errors by a factor of between 2 and 4 with observed hourly, daily and monthly errors of approximately 9%, 5% and 3.5% respectively for all sensors. These errors were greater than those found in the literature, indicating the possible presence of a systematic component of error. Analysis of the temporal variation of measurement error and its relationship with incident irradiance indicated the presence of an inter-system discrepancy in sensor offset. The close proximity of the two systems indicated that this was due to differences in system calibration, maintenance or response rather than environment and the results can therefore be used for future system re-calibration and to improve data accuracy. This paper demonstrates that straightforward validation procedures can yield meaningful results and greater emphasis on data validation is recommended for the solar community as a whole.

Keywords: Solar radiation; irradiance; data validation; pyranometry; measurement uncertainty.

1. INTRODUCTION

Accurate estimates of solar energy conversion device performance indicators such as system efficiency or performance ratio are directly dependent on accurate and reliable measurements of incident irradiance for the system's location. Where detailed prediction and monitoring of system performance are required, incident irradiance levels can be measured on-site using one or more radiation measuring devices such as a pyranometer. An understanding and quantification of the measurement error associated with the measurement device in use is therefore important, as this error or uncertainty will have a consequent effect on the uncertainty associated with the system performance estimate, possibly with corresponding financial implications. The development of data validation and quality control techniques that can be used to ensure high quality irradiance data is therefore of interest, particularly with the rapid growth of solar energy systems and applications. Hay [1] provides an introduction to the subject of solar radiation data validation and quality control and proposes the distinction between technical control, which includes aspects such as instrument calibration and maintenance, and critical control, which involves assessing the accuracy of the data for example through the identification of erroneous data, equipment malfunction and lack of continuity in the data series. On the subject of technical control, Myers [2] states that:

> “The results of the uncertainty analysis for the instrument calibration and field data measurement process indicate that the total measurement uncertainty in pyranometry (i.e. the measurement of global solar irradiance) can approach 5%.”

A figure of +/-5% uncertainty is also associated with hourly pyranometer measurements by Hay [3]. Longer averaging times are found to decrease the measurement uncertainty, with +/-3% found to be more appropriate for monthly data. On the subject of critical control, Hay [1] identifies several types of procedure for solar radiation data validation and states that “the optimum approach is to use entirely independent sensing and recording systems,”. Alternative validation procedures can use: historic co-located data (e.g. archive data), historic calculated data (e.g. using Meteonorm or similar packages) or concurrent calculated data (e.g. through the extrapolation of local Meteorological Office data).

This paper provides an assessment of the measurement error associated with the pyranometers of two independent, co-located measurement systems: the COMS3 Meteo and COMS3 Stability systems, situated on the roof of the Sir David Davies Building, Loughborough University and which are owned and operated by the Applied Photovoltaics Group at the Centre for Renewable Energy Systems Technology (CREST). The objectives are to quantify the discrepancies between the irradiance measurements of the two systems and to investigate the origin of any error. The following sections will cover the data sources used in this investigation, the statistical analyses applied to the data and the results, which have been presented in terms of the root mean square error and mean bias error. Note that whilst the term 'error' is used throughout the paper, this is used as a comparative term; no claim is made as to the relative accuracy of one measuring device over the other nor of the accuracy of either system as a whole.

2. METHODOLOGY

2.1 Data Sources

This paper aims to provide an assessment of the measurement uncertainty in irradiance data for two independent measurement systems: the COMS3 Meteo and COMS3 Stability systems. These consist of an array of sensors which have been taking automated measurements of weather and outdoor weather data for several years. This paper will focus on the analysis of measurements of global irradiance in the 45° inclined south facing plane, as this allowed the comparison of a maximum number of available pyranometers; two per system, four in total (Table 1). Concurrent irradiance measurements were available from April 2007 until October 2008, these dates marking the maximum period for which all four sensors were operational. The COMS3 Meteo data was available at a sampling frequency of five seconds while the COMS3 Stability data was available as average irradiance values as measured over ten minute intervals. Both systems used National Instruments data acquisition cards, a PCI-6225 for the Meteo system and a PXI-6221 for the Stability system. Whilst the high sampling frequency allowed the possibility of error assessments of a similarly high averaging frequency, in practice time scales of an hour to several months are commonly used for solar energy applications [4], so a base data set was created using hourly data from each sensor matched up chronologically.

Table 1. Sensor data sets.

Sensor (all Kipp & Zonen)	System Name	Duration of Available Records
CM11	COMS3 Meteo	Jan 2007 - Oct 2008
CM22	COMS3 Meteo	Jan 2007 - July 2009
CMP11	COMS3 Stability	April 2007 - June 2009
SPLite	COMS3 Stability	Feb 2007 - June 2009

In order to provide a consistent comparison between sensor measurements, data was only included in the base data set where all four sensors were available. Given that night time measurements were seldom available for the Stability system and that occasionally one or more of the sensors were inactive for periods of up to several days due to system maintenance or repair, for these occasions data was not included for any of the sensors. Finally, as will be detailed in the Results section below, it was found that a number of obviously erroneous measurements had to be filtered out from the data in order to allow meaningful analyses. An arbitrary though reasonable filter range of 0-1500W/m^2 was therefore applied to the sensor irradiance measurements for comparison purposes and values outside of this range were rejected.

2.2 Statistical Analyses

In order to provide an indication of measurement uncertainty for each sensor, two components of measurement error were calculated. The root mean square error (RMSE) is a measure of the difference between two sets of variables and includes the systematic and non-systematic components of the error. The absolute value was calculated using:

$$\mathrm{RMSE} = \sqrt{\left(\sum_{i=1}^{N} (X-Y)^2 / N\right)} \qquad [1.]$$

whilst the relative or normalised root mean square error (NRMSE) was calculated using:

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{\bar{Y}} \qquad [2.]$$

where X = irradiance measurements of the sensor in question, Y = mean irradiance as measured by all the sensors, $\bar{X}$ = mean of X, $\bar{Y}$ = mean of Y and N = number of observations.

The second component of measurement error that was calculated was the mean bias error (MBE), which is a measure of the systematic component of the differences between two variables. The absolute value was calculated using:

$$\mathrm{MBE} = \bar{X} - \bar{Y} \qquad [3.]$$

and the relative or normalised mean bias error (NMBE) was calculated using:

$$\mathrm{NMBE} = \frac{\mathrm{MBE}}{\bar{Y}} \qquad [4.]$$

We note that whilst the measurement uncertainty for each sensor would have ideally been measured with respect to a reference measurement of high and known accuracy, such as measurements from a Baseline Solar Radiation Network site, data from such sources would have had to be altered in order to compensate for their distance from Loughborough (Table 2), thus incorporating an additional level of error brought on by the calculations used. It was for this reason and due to the fact that there were a relatively large number of co-located sensors, that it was felt that using the mean irradiance as measured by all of the sensors would be a sufficiently good reference.

Table 2. Possible reference data.

Type	Location	Approximate distance (miles)
Baseline Surface Radiation Network	Cambourne	65
UK Met Office	Cottesmore	20

In the following the RMSE and MBE statistics will generally be normalised by the appropriate mean irradiance in order to compensate for the seasonal variation in irradiance observed at the site, shown in Fig. 1.

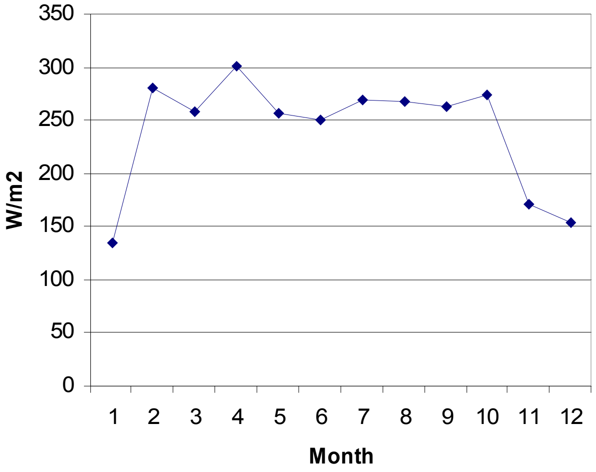

Figure 1. Measured monthly mean global 45° inclined south facing irradiances (mean of all sensors) for the period April 2007-October 2008.

3. RESULTS

3.1 Root mean square error

The RMSE, which provides a measure of the difference between a pair of variables, was evaluated for each sensor with reference to the mean irradiance as measured by all four sensors over a range of averaging time periods from one hour to up to one month. This was done for the two years worth of data that was available and the mean value of the two years is displayed in Fig. 2 for both unfiltered and filtered (0-1500W/m^2) data sets. The number of data points for each averaging period is given in Table 3. Fig. 2 demonstrates the effect that filtering had on measurement uncertainty; the RMSE was reduced by a factor of between 2 and 4 for all sensors. The effect was greatest for the CM11 and smallest for the CM22 indicating that the CM11 tended to produce the most erroneous measurements. The filtered data demonstrated consistency across all four sensors with RMSE values of approximately 9%, 5% and 3.5% for hourly, daily and monthly averaging time periods respectively, indicating that RMSE values are reduced as the averaging time period is increased. This is due to the fact that individual hourly differences that contribute to the hourly RMSE will be more likely to be offset over the course of a longer averaging period. The greatest reduction in RMSE values was observed when changing from hourly to daily time periods. Evidently the error associated with a particular irradiance estimate is dependent on the time period over which the estimate is measured and this time period should be stipulated when irradiance estimates are given in general. The observed RMSE were higher than those found in the literature [2, 3]. For example, Hay gives hourly RMSE values nearer to 5% and monthly values nearer +/-3%, indicating a possible contribution to the RMSE above and beyond the inherent uncertainty that is to be expected from even the most accurate of irradiance sensors.

Table 3. Number of data points for unfiltered and filtered data shown in Fig. 2.

Averaging Period	Number of Data Points	
	Unfiltered	Filterd
Month	19	19
Day	526	523
Hour	8262	6949

3.2 Mean Bias Error

The observed RMSE values for all four sensors indicated the possibility of a relatively large contribution of a systematic component of error, for example due to a discrepancy in the offsets of each pyranometer. Indeed, when the RMSE and MBE, a measure of the systematic component of the differences between two values, were analysed for each sensor over different averaging periods, it was found that the systematic component of error, given by the MBE, was dominant. Table 4 shows values for a typical day where it can be seen that the magnitude of the MBE is similar to the magnitude of the RMSE. Fig. 3 shows the variation in MBE throughout the year for the four sensors using unfiltered and filtered (0-1500W/m^2) data sets. The magnitudes of monthly MBE values for unfiltered data

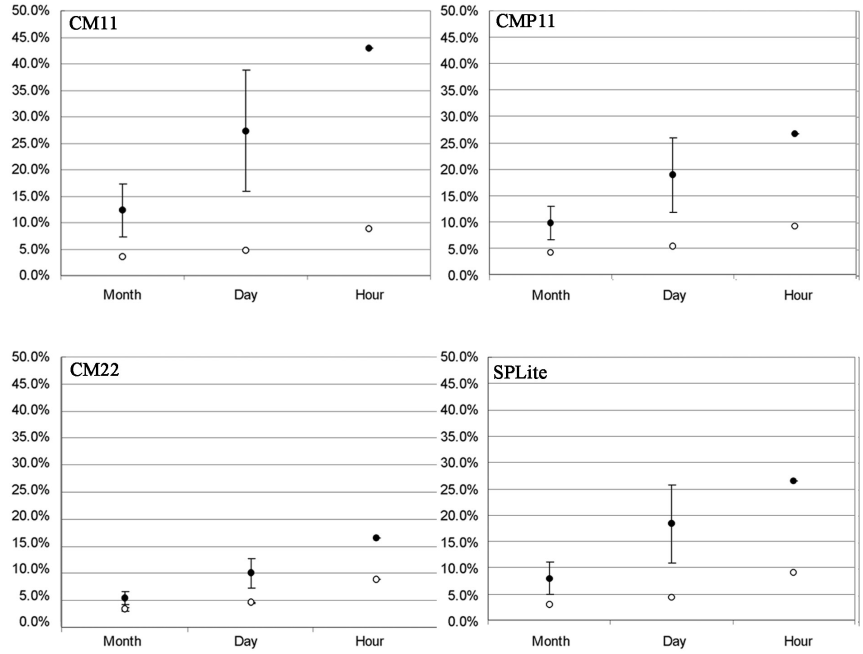

Figure 2. Root mean square error for each sensor over differing averaging periods. Unfiltered [●] and filtered [○] (0-1500W/m^2) data is shown and has been normalised by the appropriate measured mean incident irradiance. Error bars indicate maximum and minimum values that were observed.

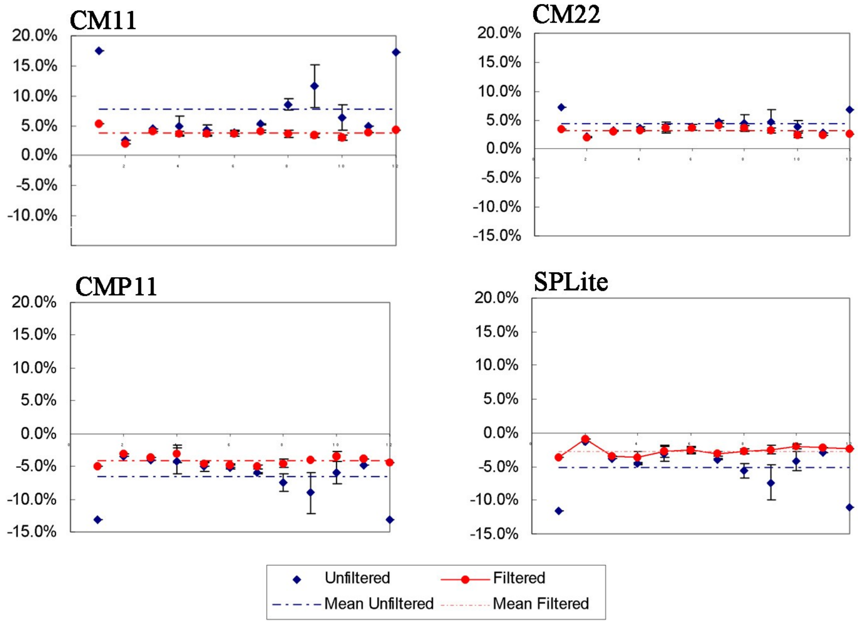

Figure 3. Mean bias (or average) error for each month and for each sensor. Unfiltered and filtered data are shown and have been normalised by the appropriate measured mean incident irradiance.

were greater during winter months than during summer months. By contrast, the magnitude of monthly MBE values for the filtered data remained relatively steady throughout the year. Interestingly, the sensor pairs from each measurement system displayed similar MBE values in both magnitude and sign; 3.7% and 3.1% for CM11 and CM22 (COMS3 Meteo) and -4.1% and -2.7% for CMP11 and SPLite (COMS3 Stability). This relative consistency between sensor pairs could be characteristic of a calibration issue rather than an environmental issue e.g. shading, as both systems are located in close proximity to each other on the same rooftop. For a complete data set the MBE is independent of averaging time [3] and a single value can be given for each sensor. This value should be consistent with the RMSE value approached asymptotically as the averaging period is increased and indeed this is the case when considering the asymptotic RMSE values implied from Fig. 2.

Table 4. Typical daily sensor errors. Values shown are for the 17th July 2008.

Sensor	Daily Error [W/m^2]	
	RMSE	MBE
CM11	7.54	7.32
CM22	5.37	4.87
CMP11	9.52	-8.96
SPLite	3.84	-3.23

3.3 Effect of incident irradiance on observed errors

In an effort to characterise the relationship of measurement error with respect to incident irradiance, hourly RMSE values (using filtered data) were calculated for days with the following distinct irradiance profiles: a cloudy winter day, sunny winter day, cloudy summer day and sunny summer day (Fig. 4). RMSE values have been normalised using observed hourly incident irradiance. For sunny days the observed hourly RMSE values fluctuated around the 5% level observed by Hay [3] when incident irradiances were high. Consistent with previous observations, CM11 showed the greatest errors in general, in particular during the cloudy winter day, whereas in general CM22 had the lowest errors. RMSE values were observed to be greater when the observed irradiance was low, therefore tending to be greatest during the early mornings and late afternoons as well as during cloudy days in particular during winter. This behaviour can be partly explained by lower zenith angles during the beginning and end of the day and the consequent effect on pyranometer accuracy [5]. However the large errors that were observed, in particular during the middle of the day, are indicative of the presence of a sensor offset error, small enough to become relatively insignificant when irradiances are high but a significant source of relative error during low irradiance conditions.

In order to get an indication of sensor offset error, daily RMSE and MBE values were calculated and plotted against the daily irradiance class intervals shown in Table 5 for each sensor using filtered data (Fig. 5). All sensors displayed a broad trend of increasing errors with increasing irradiance. As the random component of measurement error should increase proportionally with the magnitude of the observed variable, this is to be expected. The slight decrease in error observed for very high irradiances could be due to the small number of data points for these class intervals. All sensors also demonstrated characteristic offsets which were estimated from the y-intercept of a linear fit to the data. The resulting offsets were 7.11, 4.36, 8.81, 5.00 (W/m^2) for CM11, CM22, CMP11 and SPLite respectively. This evidence confirms previous observations, that the high levels of measurement error observed are in part due to the presence of inconsistent pyranometer zero-offsets, particularly in between the Meteo and Stability systems. It is apparent that whilst sensor calibration has been consistent within each measurement system, this has not been the case between systems and this has been the cause of the unexpectedly large measurement errors that were observed for this period.

Table 5. Daily irradiance class intervals and corresponding number of sensor data points. * Data for these intervals is not plotted below.

Class Intervals	Number of Data Points			
Daily Irradiance [kWh/m^2/day]	CM11	CM22	CMP11	SPLite
14-16	5	5	2*	2*
12-14	13	13	11	11
10-12	63	64	49	51
8-10	89	90	87	88
6-8	92	90	88	91
4-6	116	111	119	119
2-4	85	88	96	92
0-2	64	66	75	73

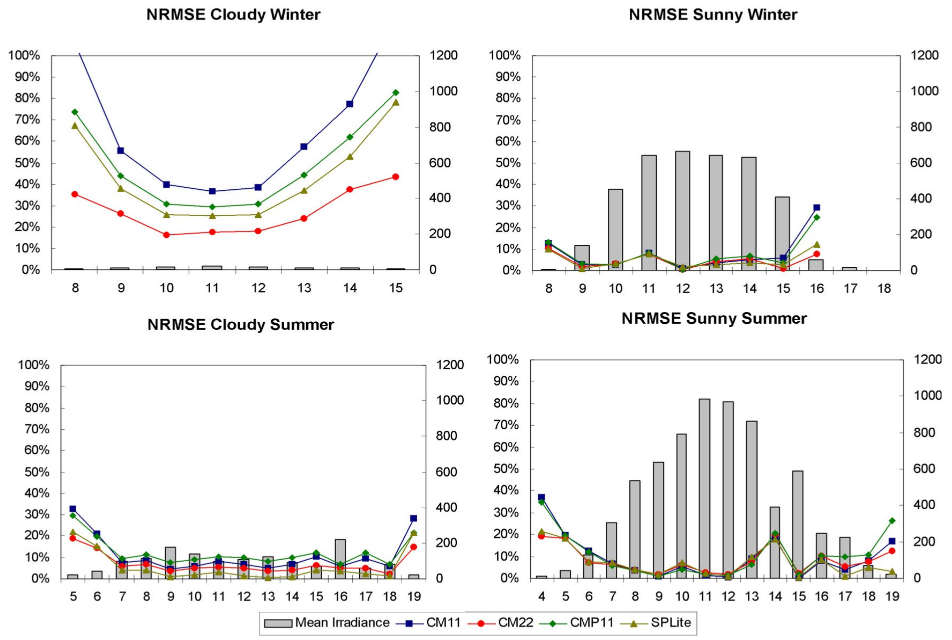

Figure 4. Hourly root mean square errors for each sensor for four types of day: cloudy winter (1st Jan 2008), sunny winter (30th Jan 2008), cloudy summer (17th July 2008) and sunny summer (26th July 2008). Data shown has been filtered and errors, which are plotted against the left hand axis, have been normalised by the hourly irradiances, plotted against the right hand axis as measured by the appropriate sensor.

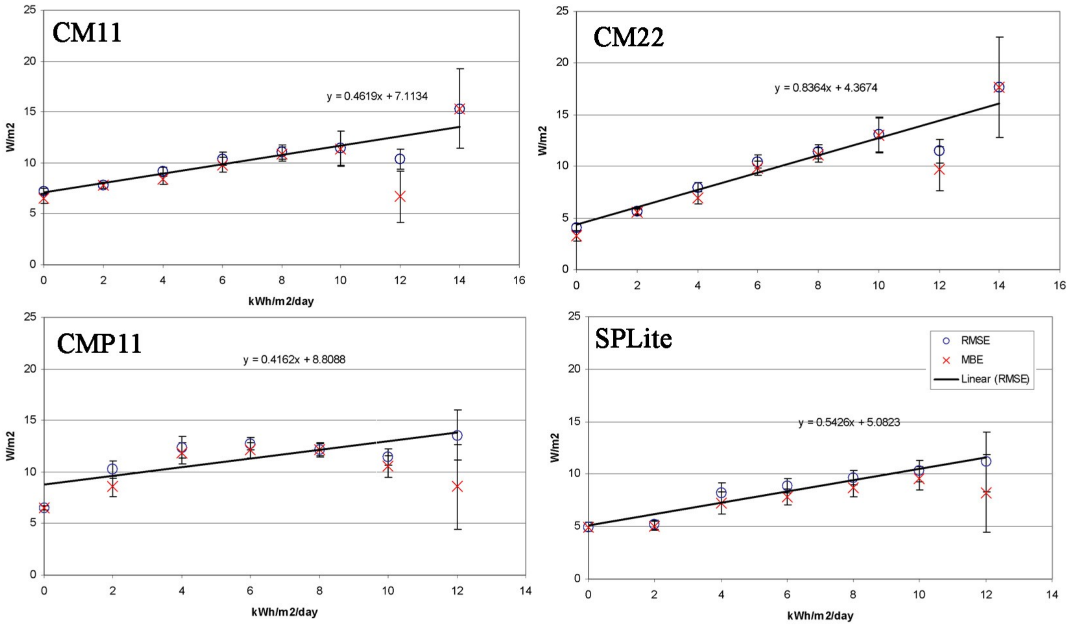

Figure 5. Root mean square and mean bias errors are shown as a function of measured daily irradiance by each sensor. Error statistics were calculated for each daily irradiance class interval and the mean error value is shown along with the standard error in the set. For ease of comparison the sign of the MBE for CMP11 and SPLite has been reversed. A linear fit has been applied to RMSE data.

The significance of the errors shown in Figs. 4 & 5 can be appreciated by considering a histogram plot of Loughborough's solar resource (Fig. 6). Irradiance levels above 12 kWh/m^2/day are rare, as are the high measurement errors associated with these irradiance levels. However it is apparent that significant low irradiance levels are experienced for a considerable proportion of the time, highlighting the fact that a small discrepancy in sensor offset between systems, as seen in Fig 5, can have a significant effect on measurement error in a low-irradiance site such as Loughborough.

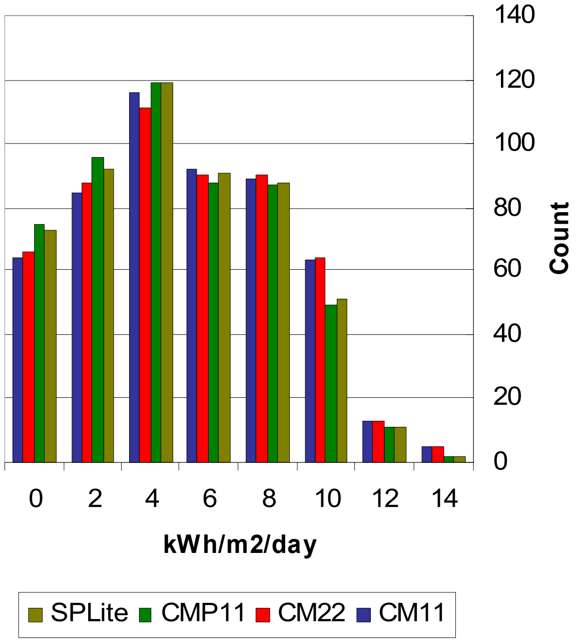

Figure 6. Frequency count for daily irradiance class intervals for each sensor at Loughborough University.

4. CONCLUSIONS

The co-location of two independent solar radiation measurement systems at Loughborough University allowed a detailed investigation of pyranometer measurement error using inter- and intra-system data validation procedures. The inclusion of a 0-1500W/m^2 data filter, which eliminated erroneous 'outlier' measurements, yielded similar measurement errors for all four sensors of +/-9%, +/-5% and +/-3.5% for hourly, daily and monthly averaging periods respectively. Whilst consistent, the errors were greater than those expected from the literature and subsequent analyses of the annual and daily variation of measurement error alluded to the presence of a systematic component of error, specifically on an inter-system basis, small enough to be relatively insignificant during high irradiance periods but a significant source of relative error during periods of low irradiance. Subsequent analysis of the relationship between daily measurement error and daily incident irradiance identified the presence of a zero-offset error for each sensor and the probable source of the unexpectedly large errors observed previously. The two COMS3 Meteo sensors (CM11 and CM22) were found to have consistently positive mean bias errors whereas the COMS3 Stability sensors (CMP11 and SPLite) were found to have consistently negative mean bias errors. This would indicate that the source of the systematic errors was due to technical or operational differences between the two systems rather than environmental differences, given their close proximity to each other.

Data validation can be a lengthy, laborious and inconclusive exercise and can often be forsaken for these reasons. However it is clear that high quality irradiance data and performance estimates cannot be obtained without it. This paper presents a data validation procedure using concurrent co-located independent measurement data, demonstrating how relatively straightforward statistical analysis techniques can be used to yield valuable insight into sources of pyranometer measurement error, in this case an inter-system zero-offset discrepancy. Further validation steps would be needed in order to assess the absolute accuracy of the systems, notably using concurrent calculated data from a source of known and high accuracy, for example the nearby Cottesmore Met Office. Validation steps naturally lend themselves to being developed in discrete stages and can be a useful addition to regular data acquisition procedures. A greater emphasis on data validation is therefore recommended for the solar community in general.

5. ACKNOWLEDGEMENTS

I would like to thank Dr. Thomas Betts for his supervision and guidance throughout my research project. Thanks also to all my fellow MSc students this year, in particular; Mark Harley, Paul Ross, John New, Jeremy Conn and Alison Cartwright. Finally, I would like to express my deepest gratitude to my ever-supportive partner Isabelle.

6. REFERENCES

[1] Hay, J.E., 1993. Solar Radiation Data: Validation and Quality Control. *Renewable Energy*, Vol. 3, No. 4/5, pp. 349-355.

[2] Myers, D.R., 1989. Uncertainty Estimates for Global Solar Irradiance Measurements Used to Evaluate PV Device Performance. *Solar Cells*, Vol. 27, pp. 455-464.

[3] Hay, J.E., 1982. An Assessment of the Uncertainty in Measurements of Solar Radiation. Solar *Energy*, Vol. 29, No. 4, pp.271-278.

[4] Hay, J.E., 1985. Evaluating the Solar Resource: a review of problems resulting from temporal, spatial and angular variations. *Solar Energy*, Vol. 34, No. 1, pp. 151-161.

[5] Kipp & Zonen, 2004. *CM22 Precision Pyranometer Instruction Manual*. Delft, Holland.