Developing the Proof of Concept for the SERI QC Flag Translation

Stephen Wilcox and Thomas Stoffel

Solar Resource Solutions, LLC

NREL Technical Monitor: Aron Habte

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Subcontract Report
NREL/SR-5D00-85589
June 2023

Developing the Proof of Concept for SERI QC Flag Translation

Stephen Wilcox and Thomas Stoffel

Solar Resource Solutions, LLC

NREL Technical Monitor: Aron Habte

Suggested Citation
Wilcox, Stephen, and Thomas Stoffel. 2023. *Developing the Proof of Concept for the SERI QC Flag Translation.* Golden, CO: National Renewable Energy Laboratory. NREL/SR-5D00-85589. https://www.nrel.gov/docs/fy23osti/85589.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Subcontract Report
NREL/SR-5D00-85589
June 2023

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

Preface

The Data Quality and Uncertainty Integration project is a 3-year effort to address stakeholder needs for assessing solar radiation resource data quality based on existing tools for estimating radiometer measurement uncertainties and assessing post-measurement data quality. The annual research objectives for the project address a logical progression of effort needed to achieve the project goal:

- Fiscal Year 2022—review and evaluation:
  - Evaluate existing data quality assessment methods as they relate to measurement uncertainty metrics.
  - Using existing data and simulated error conditions, develop a proof of concept for translating SERI QC flags or related information into a measure of uncertainty.
- FY 2023—conceptual development:
  - Develop a method for translating data quality assessment flags from SERI QC into estimated measurement uncertainty values.
  - Develop a method that incorporates the National Renewable Energy Laboratory's (NREL's) Solar Resource Uncertainty Application[1] and the data quality assessment uncertainty to quantify the overall uncertainty of an individual time-stamped solar radiation measurement.
- FY 2024—outreach and code development:
  - NREL will solicit industry partners for approaches to testing and applying the newly developed code/method.
  - Develop, verify, and validate a new software package consistent with the project goal.

This technical report addresses the second objective in FY 2022—developing the proof of concept for estimating the operational uncertainty from information derived by SERI QC.

The work presented in this report was performed under agreement number SUB-2022-10137 between the Alliance for Sustainable Energy, LLC and Solar Resource Solutions, LLC as part of the U.S. Department of Energy prime contract number DE-AC36-08GO28308.

[1] See https://midcdmz.nrel.gov/radiometer_uncert.xlsx.

Acknowledgments

We are grateful to the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office and to the Systems Integration and Photovoltaic subprograms for supporting this project. Specifically, we acknowledge Dr. Tassos Golnas, Dr. Guohui Yuan, and Dr. Lenny Tinker for their support and encouragement.

We also appreciate the administrative and technical support provided by Dr. Manajit Sengupta and Aron Habte in the Power Systems Engineering Center at the National Renewable Energy Laboratory.

List of Acronyms

DHI	diffuse (sky) horizontal irradiance
DNI	direct normal (beam) irradiance
ETR	extraterrestrial radiation
ETRn	direct normal extraterrestrial radiation
GHI	global (total hemispheric) horizontal irradiance
MIDC	Measurement and Instrumentation Data Center
NOAA	National Oceanic and Atmospheric Administration
NREL	National Renewable Energy Laboratory
QCFIT	software for determining boundaries of acceptable solar irradiance data
SERI QC	software function for post-measurement quality assessment of solar irradiance data
SERI	Solar Energy Research Institute
SOLPOS	Solar Position and Extraterrestrial Intensity
SRRL	Solar Radiation Research Laboratory
SURFRAD	Surface Radiation Budget
SZA	solar zenith angle
TSI	total solar irradiance
U_O	operational uncertainty

Executive Summary

Acquiring solar resource data with known uncertainty directly supports the goal of making solar energy conversion more cost-competitive with other forms of energy by improving the tools and methods to measure and model solar radiation, thereby reducing uncertainty in predicting solar-generated energy output and improving the bankability, efficiency, profitability, and compliance of solar energy conversion systems. This project seeks to develop a method for determining the uncertainty of high-resolution solar irradiance measurements by incorporating results from an existing data quality assessment process with estimates of measurement uncertainty for specific radiometer design performance. The method presumes that the data were collected according to best-practice protocols designed to minimize measurement errors.

This report, the second of six in the Data Quality and Uncertainty Integration project, describes a method for estimating the operational uncertainty (U_O) (uncertainty attributable to errors during field data acquisition) from three-component solar irradiance measurements using information available from SERI QC, a robust solar data quality assessment software tool. With minor modifications, SERI QC will provide an assessment of U_O for each measurement record. The U_O is used in conjunction with the existing National Renewable Energy Laboratory method for estimating the expanded measurement uncertainty (U_{95}) for each radiometer type to provide an integrated estimate of uncertainty for solar measurement data sets. These modifications will capitalize on SERI QC's long-standing capabilities for evaluating data quality, including data input validation and a variety of built-in solar routines, and will further develop the utility of the software.

The proposed concept for assigning U_O to solar resource data is derived from examining measured solar irradiance data to detect measurement errors due to substandard measurement conditions (e.g., improper maintenance, weather-induced optical contamination, data acquisition performance, improper equipment installation). The goal of this deliverable is to determine the effective limits of the approach outlined in Deliverable 6.1 and to enable the use of recommendations for its application in Deliverable 6.3 and Deliverable 6.4 as described in agreement number SUB-2022-10137.

This report provides a review of the U_O computations, a detailed description of the annual solar irradiance data sets from three monitoring stations in the United States that were used in developing the proof of concept,[2] an overview of the custom data processing software used to facilitate the analysis, a confirmation that the U_O results agree with our expectations from test station operations, and recommendations for future work.

[2] The measurement stations are also the basis for developing best practices as described in Sengupta et al. (2021).

Table of Contents

1 Introduction **1**
1.1 The Scope of This Report 1
1.2 The Goal of This Report 1
2 Background **2**
3 Calculating Operational Uncertainty **3**
3.1 K-Space 3
3.2 Operational Uncertainty Equations 4
3.3 Caveats to the U_O Equations 4
4 Data Sets **5**
4.1 Data Set Assembly 5
4.1.1 NREL 5
4.1.2 SURFRAD 6
4.1.3 NREL Data Preparation 6
4.1.4 SURFRAD Data Preparation 6
4.1.5 Simulated Error Data Set 6
4.1.6 Final Output Files 7
5 Analysis **8**
5.1 K-Space Scatterplots 8
5.2 Simulated Error Data Sets 10
5.3 Data Filtered by SERI QC Flags 11
5.4 Unfiltered Data Sets 13
6 Supplemental Approaches **14**
6.1 Using U_OKt to Represent a Unified Operational Uncertainty 14
6.2 Clear-Sky U_OKn 15
6.3 Cloudy-Sky U_OKd 17
6.4 Using SERI QC Flags to Determine U_O 20
7 Conclusions **21**
8 Next Steps **23**
References **24**

List of Figures

Figure 5-1. K-space scatterplots.. 8
Figure 5-2. QCFIT plots and K-space residual histograms for the SRRL 9
Figure 5-3. QCFIT plots and K-space residual histograms for Penn State 10
Figure 5-4. QCFIT plots and K-space residual histograms for Fort Peck.. 10
Figure 5-5. Scatterplots of 3% simulated errors.. 11
Figure 5-6. Scatterplots for the filtered SRRL data set 12
Figure 5-7. Histograms of U_O for the filtered SRRL data.. 12
Figure 5-8. Scatterplots for the unfiltered data sets 13
Figure 6-1. Distribution of the U_OKt data for errors introduced in DNI and DHI 14
Figure 6-2. U_OKn versus Kd with introduced errors.. 15
Figure 6-3. Clear-sky U_OKn processing for the SRRL 16
Figure 6-4. Clear-sky U_OKn processing for Penn State 16
Figure 6-5. Clear-sky U_OKn processing for Fort Peck... 16
Figure 6-6. U_OKn versus Kt adjusted +3% for clear-sky filtered data... 17
Figure 6-7. U_OKd versus Kn with introduced errors.. 18
Figure 6-8. Cloudy-sky U_OKd processing for the SRRL 18
Figure 6-9. Cloudy-sky U_OKd processing for Penn State 19
Figure 6-10. Cloudy-sky U_OKd processing for Fort Peck 19
Figure 6-11. U_OKd versus Kt adjusted +3% for cloudy-sky filtered data.. 20
Figure 6-12. U_OKt as a function of the SERI QC flags.. 20

List of Tables

Table 5-1. Statistics for the Filtered SRRL Data Set 12
Table 5-2. Statistics for the Unfiltered Data Sets... 13
Table 6-1. Clear-Sky U_OKn Summary Statistics 16
Table 6-2. Cloudy-Sky U_OKd Summary Statistics 19

1 Introduction

1.1 The Scope of This Report

This report is the second of six deliverables in the Data Quality and Uncertainty Integration project assigned to Solar Resource Solutions, LLC and presents an overview of developing a new concept for translating solar resource data quality assessment results into estimated uncertainty values. This evaluation is based on three-component solar irradiance measurements collected at 1-minute intervals for a 12-month period. Results of this effort will provide a better understanding of the new process prior to developing a more formalized approach for estimating data uncertainty in subsequent tasks.

1.2 The Goal of This Report

This report documents the development of a proof of concept for determining the estimated operational uncertainty (U_O) for measured solar irradiance data based on the existing attributes of the SERI QC data quality assessment software as previously described in Deliverable 6.1 to further refine the uncertainty estimates of archived solar resource data.

Specifically, this work seeks to further develop the uncertainty estimates for pyrheliometers and pyranometers used for solar resource measurements by quantifying and incorporating additional contributions to uncertainty derived from the SERI QC data quality assessment method to field measurements. Such an analysis method could allow solar resource data providers to assign a more comprehensive uncertainty that includes the effects of radiometer measurement performance and the operation and maintenance aspects of an associated solar radiation measurement station.

2 Background

When determining the viability of a proposed solar energy project, analysts require a good measure of the solar resource to accurately predict power generation. In addition to the resource magnitude and variability, the uncertainty of a data set is required to understand the validity of the data and impose limitations on the analysis. Without a stated measure of uncertainty, a data set cannot provide context to the values therein.

In the past, frequently the uncertainty of a data set has been solely represented by either the manufacturer's stated instrument uncertainty or the uncertainty assigned by the calibration process. This approach, though it provides some basis for data set uncertainty, fails to acknowledge many additional sources of error in a measurement (Habte 2014).

The uncertainty of a data set is determined by several factors identified in current best practices (Sengupta et al. 2021), including the:

- Design and manufacturing characteristics of a measuring instrument
- Configuration and installation of a measurement station
- Quality of the instrument calibration and uncertainty of the reference instruments
- Uncertainty of the data logging equipment and associated electronic infrastructure
- Errors introduced during ongoing measurement operations.

The last item, which we call *operational uncertainty*, is difficult to ascertain because many uncontrolled factors affect a measurement, such as the frequency of instrument cleaning and other maintenance, degradation, or uncorrected failure of supporting equipment, and multiple environmental and weather conditions. In this project, we derive an estimate of U_O by examining interrelated data from measurement station instruments to detect errors resulting from substandard measurement conditions (e.g., improper maintenance, weather-induced or environmental optical contamination, improper equipment installation). The goal of this deliverable is to determine the effective limits of the approach outlined in Deliverable 6.1 and enable us to make recommendations for its application in Deliverable 6.3 and Deliverable 6.4.

The U_O method under study requires the use of simultaneous measurements of global horizontal irradiance (GHI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DHI) and assumes that data were carefully collected with protocols and guidance detailed in the current best practices manual (Sengupta et al. 2021). The method is not intended to assign uncertainty to data collected under deficient or unknown conditions.

3 Calculating Operational Uncertainty

U_O is derived from three-component solar irradiance measurements: GHI, DNI, and DHI. These values are converted by SERI QC to K-space, which is a normalized representation of a measurement independent of the effects of the atmosphere and station location.

3.1 K-Space

For each irradiance parameter, the measurement is normalized (divided) by the like component as if observed at the top of the atmosphere without any atmospheric attenuation, which we refer to as *extraterrestrial irradiance* (ETR).

The direct normal extraterrestrial irradiance (ETRn) is computed from the date and time information as:

$$ETRn = TSI * (R/Ro)^2 \tag{1}$$

where:

- TSI = total solar irradiance (1360.8 ±0.5 W/m^2)
- R = sun-Earth distance at the time of interest
- Ro = annual mean sun-Earth distance.

and the global horizontal ETR is computed as:

$$ETR = ETRn * \cos(SZA) \tag{2}$$

where:

- SZA = solar zenith angle at the location, date, and time of interest.

Thus:

$$Kt = global/ETR \tag{3}$$

$$Kn = direct/ETRn \tag{4}$$

$$Kd = diffuse/ETR \tag{5}$$

These measurements are related by the coupling equation in K-space:

$$Kt = Kn + Kd \tag{6}$$

3.2 Operational Uncertainty Equations

In the absence of a recognized measurement reference at the monitoring station, Eq. 6 is used to determine U_O by establishing a field reference for each component through the other two components in the coupling equation. A U_O ratio can then be calculated, which with "perfect data" will equal one. Any measurement error would result in a ratio less than or greater than 1. With Eq. 7, Eq. 8, and Eq. 9, a percentage error is calculated to determine the operational uncertainties for the three components:

$$U_O\ Kt = \left(\frac{Kt}{Kn+Kd} - 1\right) \cdot 100 \tag{7}$$

$$U_O\ Kn = \left(\frac{Kn}{Kt-Kd} - 1\right) \cdot 100 \tag{8}$$

$$U_O\ Kd = \left(\frac{Kd}{Kt-Kn} - 1\right) \cdot 100 \tag{9}$$

3.3 Caveats to the U_O Equations

This approach results in Eq. 8 and Eq. 9 becoming undefined if Kt = Kd or Kt = Kn. Additionally, the process might suffer when ratios are determined from measurements of a similar magnitude that are affected by significant noise or error. In theory, this occurrence in Eq. 9 is impossible except at night, when Kt = 0, unless a measurement error condition exists. The application of Eq. 8 under overcast skies (when the DNI is zero) will be undefined when Kt = Kd. Further, when the difference between Kt and Kd is nonzero but small, the ratio can produce unrealistically large values that frustrate the goal of estimating uncertainty. Additionally, measurements that occur at high zenith angles (near sunrise and sunset) can result in similar unrealistic values from ratios between small numbers.

Thus, some effort will be required to provide protection from these effects when calculating and reporting uncertainties. This will almost certainly result in an incomplete assignment of uncertainty estimates among measurements in a data set. Although this is an undesirable limitation, in the broader scope, we recall that the general purpose of solar irradiance measurements in this context is to support power generation projects. Because the direct beam (as either the single component, DNI, or a significant constituent of global) is the primary contributor to the solar resource, clear-sky or other high-irradiance conditions are of the greatest interest, and this approach is well suited for such conditions.

4 Data Sets

For this deliverable, measurement data from the National Renewable Energy Laboratory (NREL) and the National Oceanic and Atmospheric Administration (NOAA) from 2021 were used in a variety of circumstances to test and understand the behavior of the Uo equations. A 1-year data set of daytime 1-minute values (approximately 262,000 records) was included for each scenario:

1. Data of known high quality from the NREL Measurement and Instrumentation Data Center (MIDC)—the Solar Radiation Research Laboratory (SRRL) baseline "best" three-component data set.[3] The instruments were chosen after consultation with SRRL personnel. NREL modified the GHI and DHI data at the time of data acquisition to correct infrared offsets common in thermopile instruments. For the chosen instruments, the modifications were typically a fraction of a percentage of reading and based on measurement performance characterizations of the radiometer as operated at the SRRL.
2. Data from Set 1 formed to a "perfect" three-component data set by calculating GHI from DNI and DHI and then modified with the introduction of systematic and quantifiable errors.
3. Data from two stations in the NOAA Surface Radiation Budget (SURFRAD)[4] network as an example from a long-term, high-quality field measurement campaign. These data were not corrected or modified, although NOAA removed egregious data.

NREL performs annual calibrations and daily (5 days per week) cleaning and inspection of the instruments. NOAA performs annual calibrations and instrument maintenance every 2 weeks.

4.1 Data Set Assembly

4.1.1 NREL

For the SRRL data set, these parameters were downloaded in a format provided by the MIDC system:

1. Date and time
2. SZA
3. Solar azimuth angle
4. ETR
5. ETRn
6. GHI (Kipp & Zonen CM22)
7. DNI (Kipp & Zonen CHP1)
8. DHI (Kipp & Zonen CM22)
9. Air temperature.

[3] See https://midcdmz.nrel.gov/.
[4] See https://gml.noaa.gov/grad/surfrad/.

4.1.2 SURFRAD

For the SURFRAD data, the stations at Pennsylvania State University (Penn State) and Fort Peck were selected to provide some degree of climate diversity. The data were extracted from an archive at NREL with these parameters:

1. Date and time
2. GHI (Spectrosun SR-75)
3. DNI (Eppley NIP)
4. DHI (Eppley 8-48).

4.1.3 NREL Data Preparation

For the NREL data set, a program was written to calculate and add these fields to the existing data records:

- Kt, Kn, and Kd
- U_OKt, U_OKn, and U_OKd
- Absolute values of U_OKt, U_OKn, and U_OKd
- Kt-Kn-Kd residual
- SERI QC flags for GHI, DNI, and DHI (using the original C version of SERI QC).

4.1.4 SURFRAD Data Preparation

For the SURFRAD data sets, a program was written to calculate and add these fields to the existing data records:

- ETR and ETRn (using the NREL SOLPOS algorithm)[5]
- Zenith and azimuth angles (using the NREL SOLPOS algorithm)
- Kt, Kn, and Kd
- U_OKt, U_OKn, and U_OKd
- Absolute values of U_OKt, U_OKn, and U_OKd
- Kt-Kn-Kd residual
- SERI QC flags for GHI, DNI, and DHI (using the original C version of SERI QC).

4.1.5 Simulated Error Data Set

The NREL data set was reprocessed to create data records with perfectly coupled three-component data by calculating Kt and subsequently GHI according to Eq. 6. In this configuration, all records would have $U_O = 0$ for each component.

To introduce controlled errors, all GHI (Kt) values in the file were biased by +3% while holding the other two parameters unchanged, and the results were added to the records in a new field. Likewise, the DNI and DHI values were similarly biased and added to the records. In this fashion, three new fields were added to the records. The fixed 3% was chosen as an arbitrary

[5] Solar Position and Extraterrestrial Intensity; see https://www.nrel.gov/grid/solar-resource/solpos.html.

value representing common operational errors that will impose a detectable increase in the overall uncertainty.

Subsequent processing applied Eq. 7, Eq. 8, and Eq. 9 to the K-space values of Kt + 3%, Kn + 0%, and Kd + 0%, resulting in U_OKt, U_OKn, and U_OKd using the biased Kt value to determine the effect of the Kt bias on U_OKn and U_OKd.

Similarly, further processing applied the U_O equations to the values of Kn + 3%, Kt + 0%, and Kd + 0%, resulting in U_OKt, U_OKn, and U_OKd using the biased Kn value to determine the effect of the Kn bias on U_OKt and U_OKd.

Finally, processing applied the U_O equations to the values of Kd + 3%, Kt + 0%, and Kn + 0%, resulting in U_OKt, U_OKn, and U_OKd using the biased Kd value to determine the effect of the Kd bias on U_OKt and U_OKn.

These last steps added nine new fields to the records, representing the related U_OKt, U_OKn, and U_OKd values for each adjustment of the three K-space parameters.

4.1.6 Final Output Files

For both the NREL and SURFRAD data sets, the final output files were reduced in size to facilitate further analysis by randomly removing 9 of 10 records. This step is not thought to significantly change the statistical character of the data sets by retaining approximately 26,600 records of approximately 266,000.[6] The programs also created files in the QCFIT format to build the required K-space boundary files for SERI QC and to allow for inspection of the data in QCFIT (see Section 5.1).

[6] Comparisons of the full and 10% data sets revealed differences for the average irradiance values to be less than 0.3% of the full data set statistics. The SZA results were within 0.03%, suggesting no diurnal bias in the reduced data set.

5 Analysis

Spreadsheets were created for the final analysis and to plot the data. Except as noted, the results of all U_O calculations were converted to their absolute values for the final analysis. Although this might not be the ultimate use of the data in the determination of uncertainty for a single record, it is necessary in this analysis to better understand the magnitude of the U_O values in an annual data set.

For the constructed data sets, the files for each station were imported into separate spreadsheets. The imported records were filtered for various threshold levels of DNI (to avoid division by zero or small numbers) and for the desired range of SERI QC flags (to filter egregious data), then the results were plotted to illustrate the desired analysis.

5.1 K-Space Scatterplots

To provide an overview of each station, we plotted the year of data *without filtering* as Kn versus Kt (Figure 5-1).

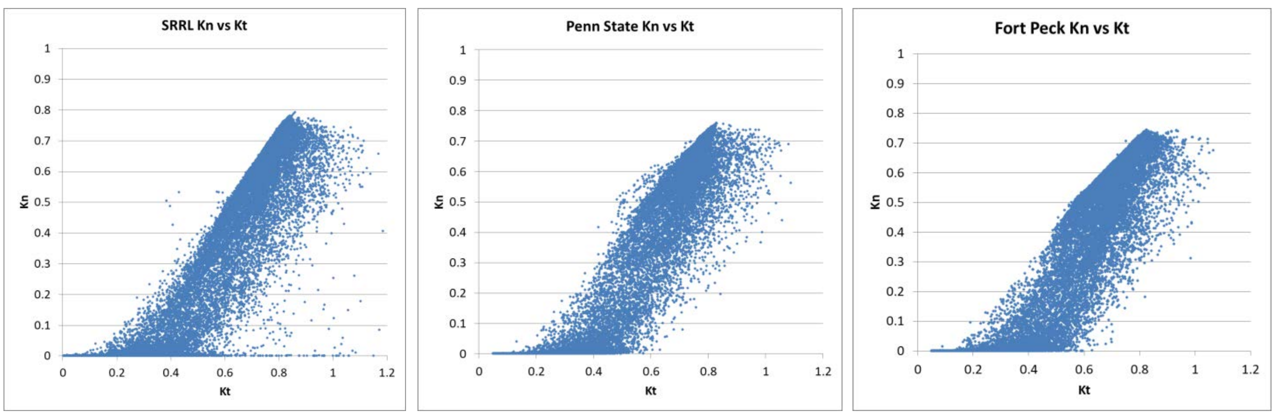

Figure 5-1. K-space scatterplots

An analysis of this figure shows:

- All three data sets exhibit a well-behaved data relationship between Kt and Kn, indicating a preponderance of good to excellent data with very few obvious anomalies.
- SRRL experiences higher irradiance than the other two sites, likely due to its higher elevation, lower values of total precipitable water vapor and lower aerosol optical depth, and more clear-sky periods. Clear-sky data are located in the upper right portion of the scatter; however, the clearer sky at SRRL is indicated by the narrower and longer diagonal region of high-density data at the upper left portion of the scatterplot envelope.
- All stations exhibit some data either near or over the 1-to-1 diagonal, which represents a boundary of impossible data where Kn > Kt.
- Some data in the SRRL plot show high Kt values for Kn at or near zero. These unrealistic data can be caused by tracker failure or obscured pyrheliometer optics from, for example, snow, ice, or cleaning. It is likely that if the NOAA stations recorded such data, they were removed from their published archives.

Data from the three stations were further inspected using QCFIT (Maxwell et al. 1993) to examine seasonally representative subsets of the data (figures 5-2, 5-3, and 5-4). This analysis provides seasonally representative subsets of the data in Kn-versus-Kt scatterplots according to the three airmass regions used by SERI QC and, separately, histogram plots of the K-space residual. The residual is calculated by rearranging Eq. 6 as:

$$Kt - Kn - Kd = 0 \tag{10}$$

Any nonzero result (residual) indicates a failure of the three-component coupling, Eq. 6, and represents a measurement error. By creating a histogram of all residuals in a data set, some measure of error among the three irradiances is portrayed. Ideally, all residuals would be zero, but the residual histograms show the distribution of data near zero. The wider the envelope of the histogram data, the greater the departure from ideal irradiance measurements during the year.

Note the narrower shape of the SRRL residual histograms relative to those of the SURFRAD stations. This indicates that more SURFRAD data points have a larger residual. The histograms also designate mean values as colored bars for the SERI QC realms of low, medium, and high airmass. These mean bars indicate that much of the larger residuals for the SURFRAD stations occur at high airmass associated with measurements near sunrise and sunset. The Kn-versus-Kt plots reveal some data errors for Fort Peck in October, and these anomalies should be apparent in the U_O calculations.

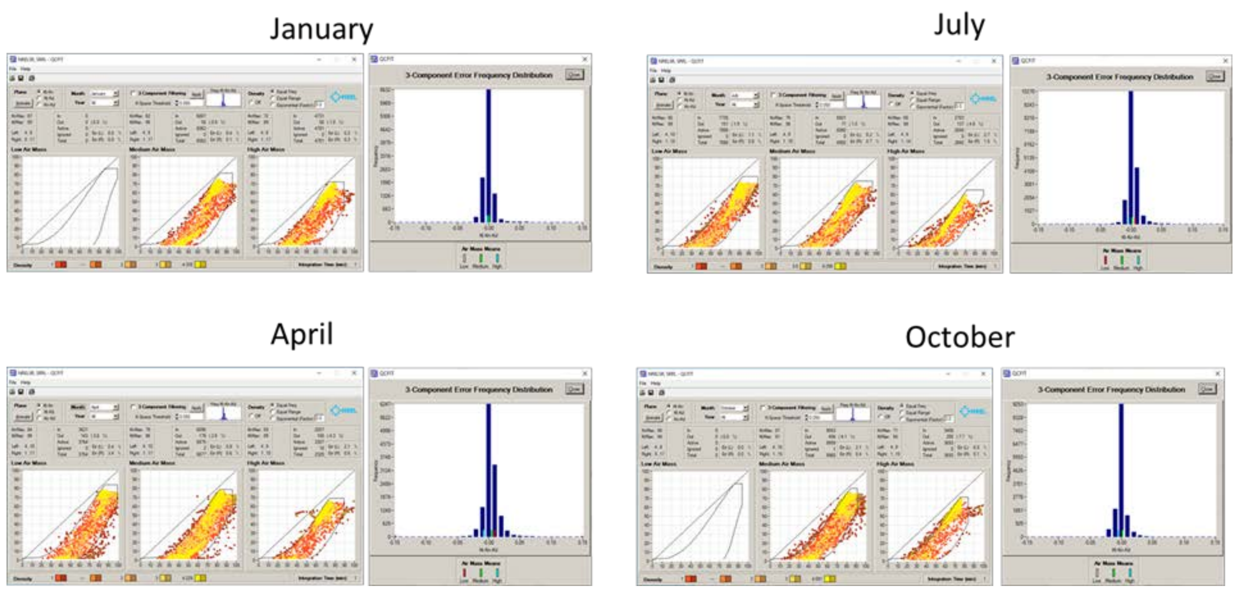

Figure 5-2. QCFIT plots and K-space residual histograms for the SRRL

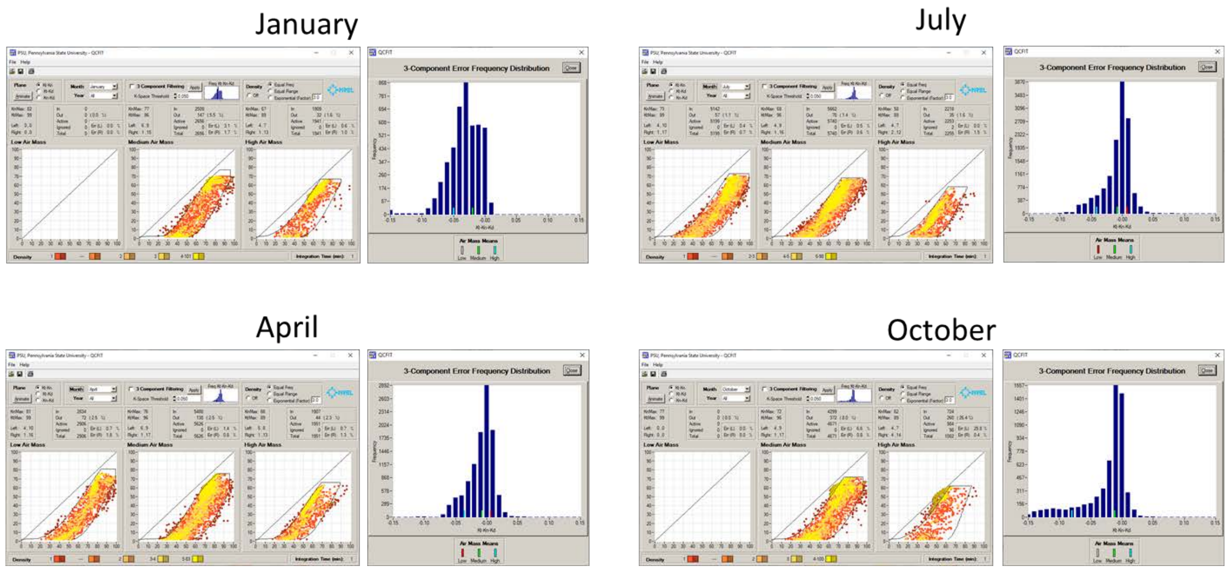

Figure 5-3. QCFIT plots and K-space residual histograms for Penn State

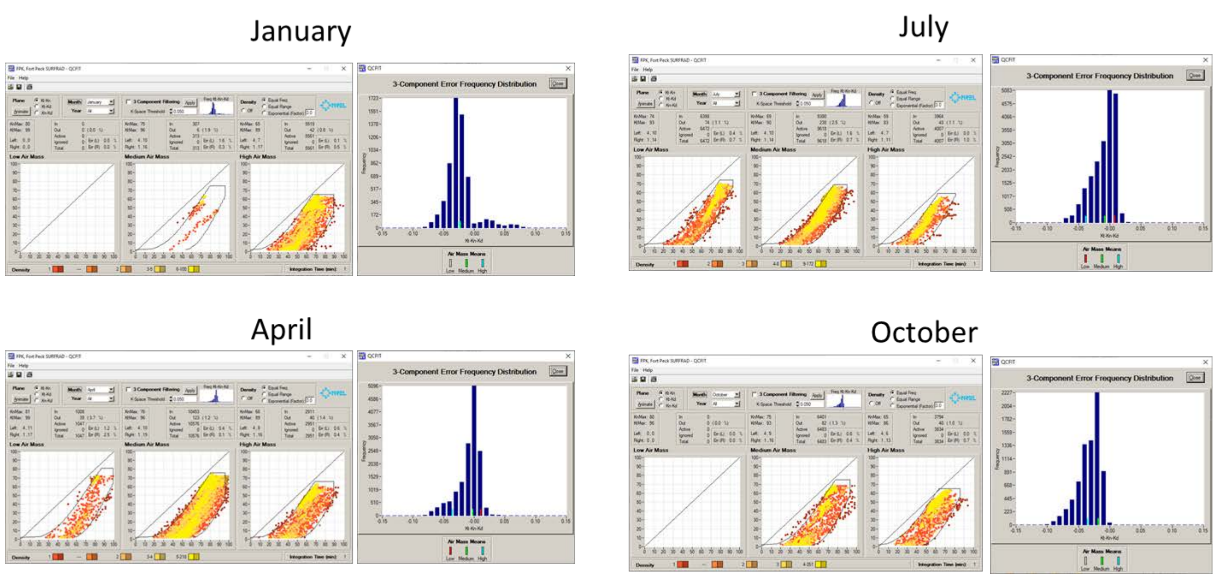

Figure 5-4. QCFIT plots and K-space residual histograms for Fort Peck

5.2 Simulated Error Data Sets

Because each of the three U_O equations use all three K-space components in the calculation, an error in any one component will affect the results of the other two. Using the simulated error data set described in Section 4.1.5, scatterplots of the U_O values versus the K-space parameters were created (Figure 5-5). Each of the three rows represents the introduced errors for Kt, Kn, and Kd, respectively. Each of the three columns shows the effect of the introduced errors of U_OKt, U_OKn, and U_OKd, respectively. Note that the plots on the diagonal from the upper left to the lower right show the resulting 3% introduced errors in U_O for that parameter. The other plots show the effect of that U_O of the unadjusted K-space value. For each column, the plot scales are held constant, though they change from one column to the next.

The plots in the U_OKn column show that the 3% error introduced in Kt and Kd is carried over to this DNI irradiance component (which we call an indicated *crossover error*), even when no error is present in Kn. U_OKn becomes even greater at low DNI with high Kd due to the subtraction in the denominator when Kd (with its error) nears the value of Kt:

$$\frac{Kn}{Kt - Kd}$$

A similar effect occurs for U_OKd when Kn is low.

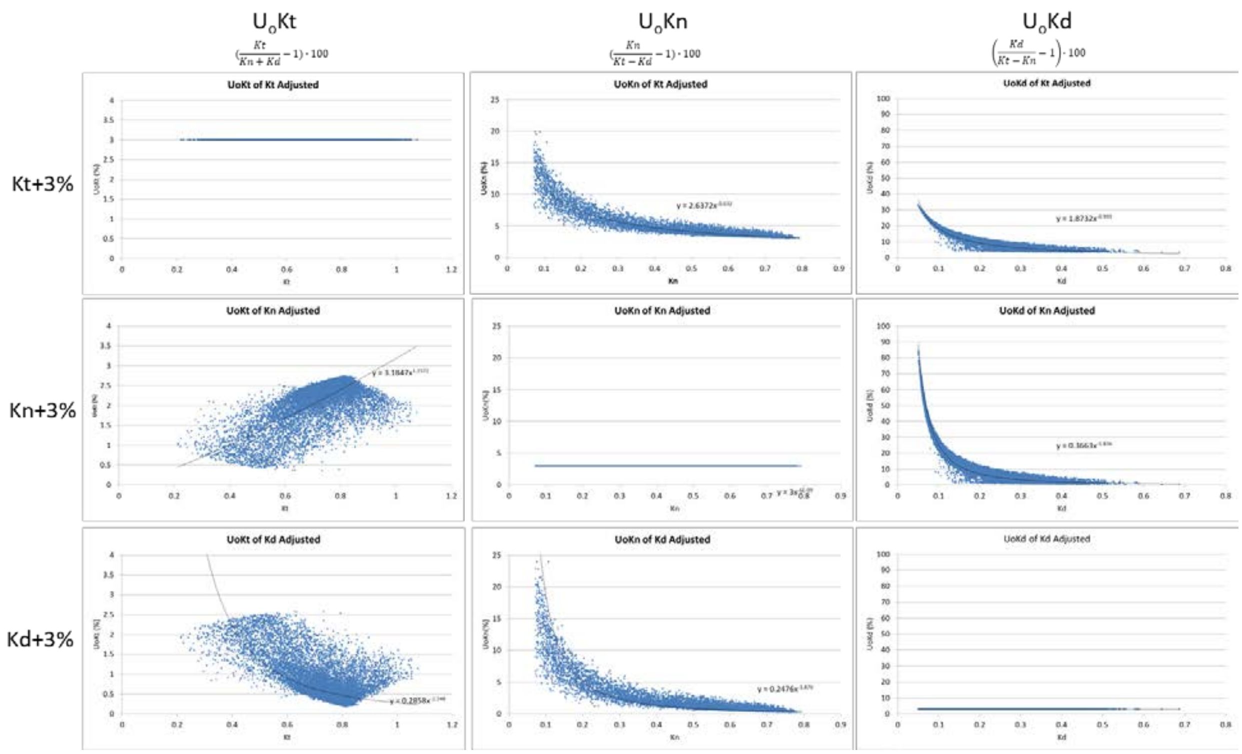

Figure 5-5. Scatterplots of 3% simulated errors

Rows indicate introduced errors; columns indicate effects on U_O.

The crossover error for U_OKt is less pronounced. The plots in the U_OKt column for the introduced errors in Kn and Kd are limited to less than the original error, whereas U_OKn and U_OKd can exhibit exaggerated errors many times the original.

5.3 Data Filtered by SERI QC Flags

Using the SERI QC flags produced during the formation of the evaluation data sets, data records for the SRRL were filtered to include only those that passed the most stringent of the SERI QC 3-component tests, i.e., Flag 03 and Flag 09. Additionally, data records from the three variables with DNI <25W/m^2 were eliminated to limit the inclusion of low levels of solar irradiance in the U_O computations.

With this SRRL data set, most spurious and anomalous data have been removed, allowing for a better analysis of the effects of the U_O process. Figure 5-6 shows the results for U_OKt, U_OKn, and U_OKd plotted against the corresponding K-space parameters, and Figure 5-7 shows histograms of the U_O distributions. Table 5-1 shows the statistics for each U_O for the filtered data.

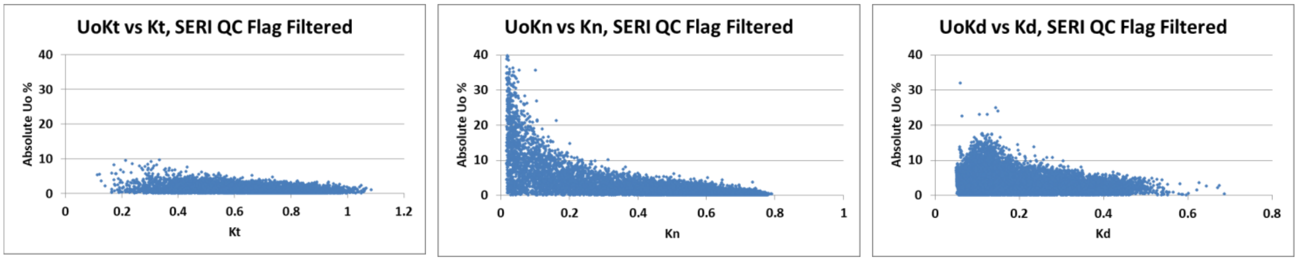

Figure 5-6. Scatterplots for the filtered SRRL data set

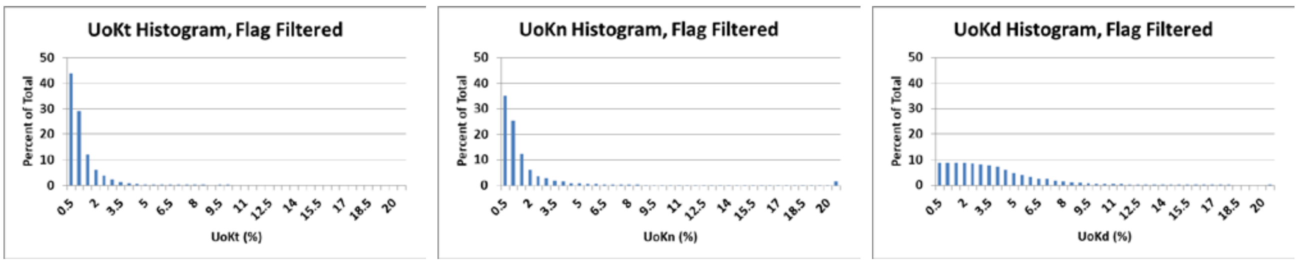

Figure 5-7. Histograms of U_O for the filtered SRRL data

Table 5-1. Statistics for the Filtered SRRL Data Set

	U_OKt (%)	U_OKn (%)	U_OKd (%)
Average	0.83	2.35	3.42
Median	0.58	0.75	2.88
P95	2.53	9.25	8.57
Aggregate	0.16	0.21	0.63

Note: The statistics in Table 5-1 are based on the absolute value of U_O for the 1-minute K-space values except for the aggregate statistic, which is the overall annual sum of the Kt, Kn, and Kd values in the data set applied to the U_O formulas.

These plots indicate an increasing U_O at lower values of Kn and Kd, illustrating the mathematical effect of ratios between small or similar numbers, as outlined in Section 3.3. Even small errors, as we expect in this data set, can be amplified under such circumstances because subtraction in the denominator skews the ratio. Nonetheless, approximately 70% of the U_OKt values and 60% of the U_OKn values in the SRRL data set are less than 1%; however, given that the data set was cleaned by filtering based on the SERI QC flag results, the values in the U_OKn and possibly U_OKd columns might not faithfully reflect the overall quality of the data. This effect is further discussed in Section 6.

5.4 Unfiltered Data Sets

In this section, data for all three stations are similarly analyzed but without the benefit of stringent data filtering and instead removing only records with egregious data. This treatment represents data sets that could be expected from well-run stations that occasionally suffer operational problems. Figure 5-8 and Table 5-2 show the results of this analysis.

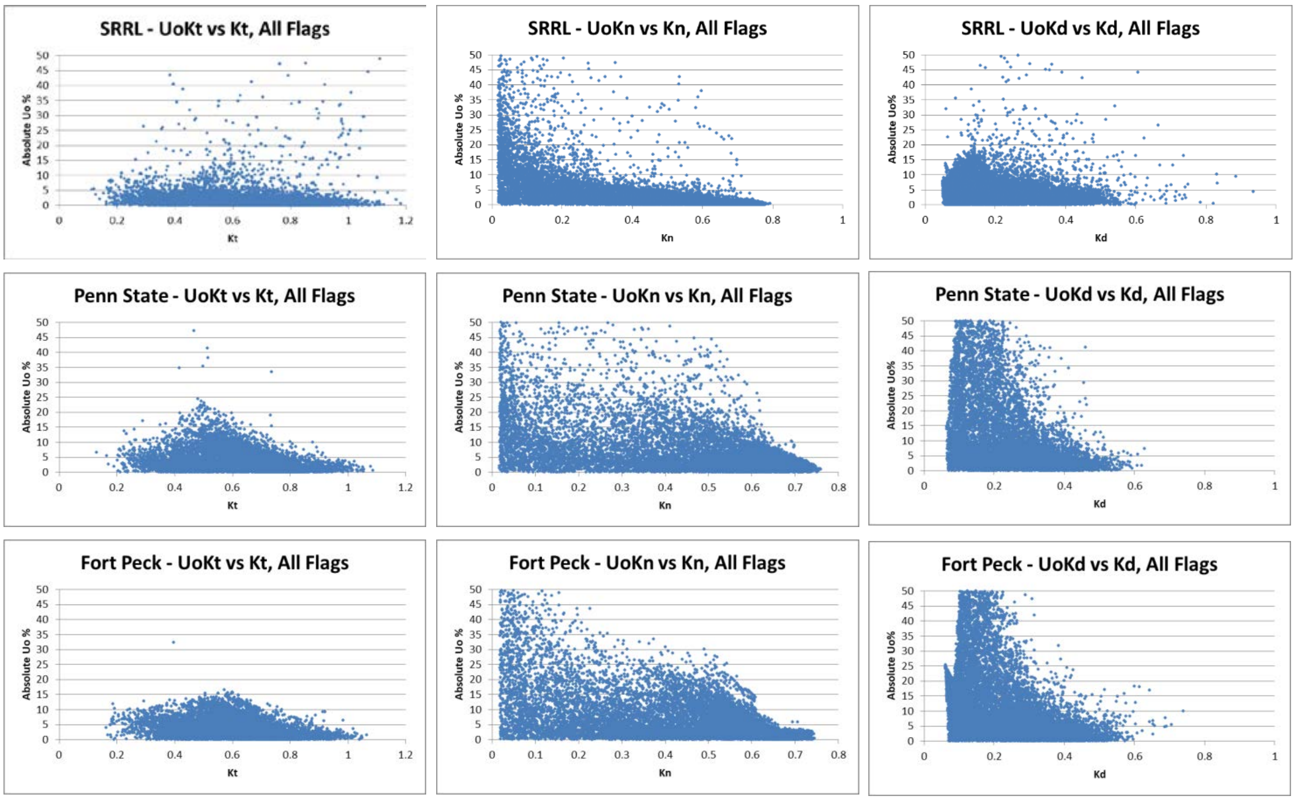

Figure 5-8. Scatterplots for the unfiltered data sets

Table 5-2. Statistics for the Unfiltered Data Sets

	NREL—All Flags, DNI >25			Penn State—All Flags, DNI >25			Fort Peck—All Flags, DNI >25		
	$U_O Kt$	$U_O Kn$	$U_O Kd$	$U_O Kt$	$U_O Kn$	$U_O Kd$	$U_O Kt$	$U_O Kn$	$U_O Kd$
Average	1.30	4.33	4.19	3.17	24142	46.75	2.96	453962	14.54
Median	0.66	0.89	3.02	1.98	3.96	5.88	1.92	3.13	7.20
P95	3.72	15.23	10.57	9.99	30.95	60.42	9.04	30.09	56.35
Aggregate	0.34	-0.47	-1.23	-2.24	3.53	7.00	-2.00	3.05	6.63

Note the anomalously large average values in the $U_O Kn$ column for Penn State and Fort Peck, which indicate that even with reasonable data, the process can yield wildly unrealistic values, and the results might not be suitable for a significant subset of measurements in a data set. In contrast, the values in the $U_O Kt$ columns are well within reason for solar measurement stations designed and operated according to current best practices.

6 Supplemental Approaches

Previous analysis in Section 5 shows that the proposed U_O equations for U_OKn and U_OKd might not be suitable under a wide range of common measurement conditions because of crossover errors resulting from the uncertainty formulation and irradiance coupling. In this section, we consider additional concepts for deriving the U_O from in situ field measurements.

6.1 Using U_OKt to Represent a Unified Operational Uncertainty

Because of the crossover effect in the U_O equations, the U_OKn and U_OKd values can be markedly unrepresentative; however, a closer look at U_OKt and the physical aspects of the coupling in Eq. 6 introduces the notion that U_OKt can be used to represent error conditions in the other two parameters. Because the ratio in Eq. 7 contains the sum of Kn and Kd in the denominator, the resulting calculation contains information from all three components in a stable ratio configuration, without the denominator trending to zero.

Referring to the analysis of the simulated data in Section 5-2, Figure 6-1 shows the distribution of the U_OKt data from Figure 5-2 in that discussion.

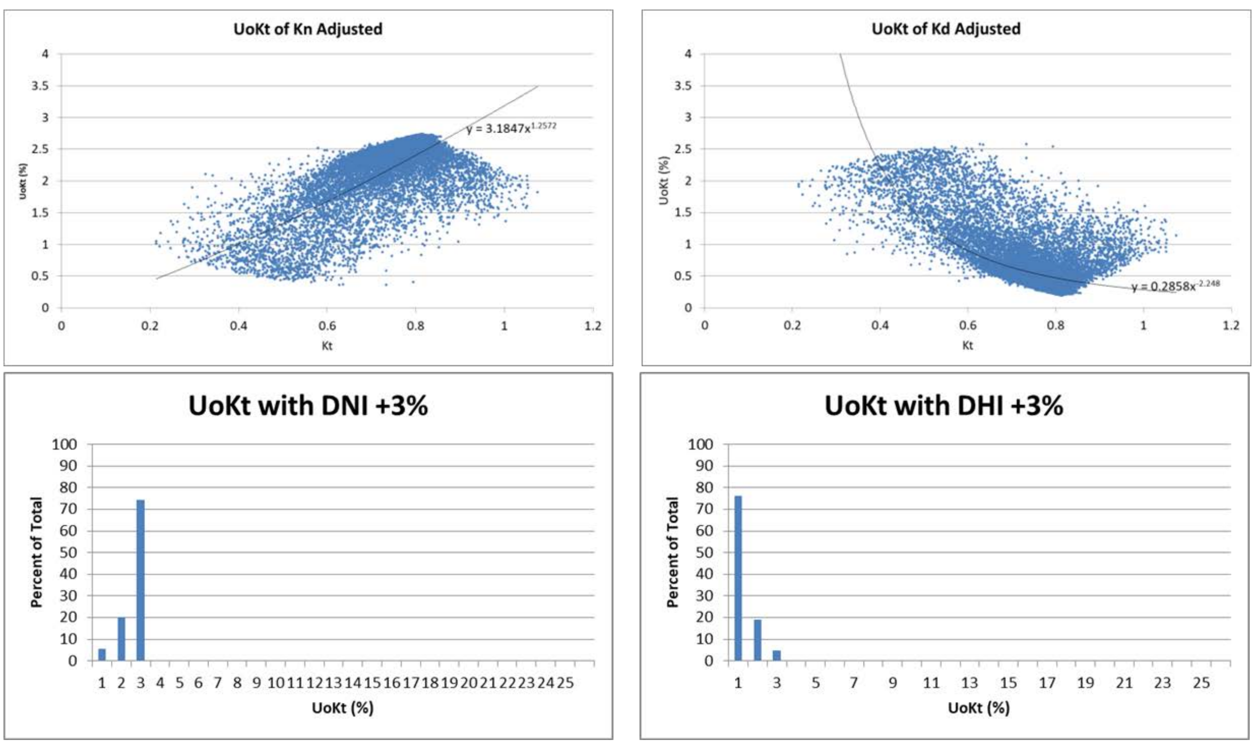

Figure 6-1. Distribution of the U_OKt data for errors introduced in DNI and DHI

The histogram for DNI+3% in Figure 6-1 shows that the U_OKt values are predominantly between 2% and 3%; thus, by itself, U_OKt provides a reasonable representation of the DNI error. The histogram for DHI+3% shows that the 3% does not carry over as well, and the DHI errors have an attenuated, though nonzero, representation in U_OKt.

Although no data set will have isolated errors such as those found in the simulated data set, this analysis indicates that $U_O Kt$ captures much of the errors in all three components and can be a stable and representative measure of the overall U_O for three-component irradiance data from a solar measurement station.

6.2 Clear-Sky $U_O Kn$

The analysis in Section 5 shows that $U_O Kn$ can be exaggerated by crossover errors from Kd during periods of low DNI and high DHI (generally overcast skies). By limiting the analysis to clear-sky data records, much of the problem data can be eliminated from the estimate of U_O. The Kt-Kn scatterplots (Figure 5-1) show clear-sky data in the upper right, with dense clustering near the diagonal portion in the upper left of the data envelope in each plot. Although it is not necessarily intuitive in these plots, the closer a Kt-Kn data point is to the graph's 1-to-1 diagonal, the lower the value of Kd (diffuse). By plotting $U_O Kn$ as a function of Kd, it becomes obvious that regions of low Kd result in low Kd crossover to $U_O Kn$. This is illustrated in Figure 6-2, which shows data from the simulated error data set (Section 4.1.5). Here, the crossover to $U_O Kn$ approaches zero as Kd (with the 3% introduced DHI error) approaches zero. The data in the red circle represent the clearest skies.

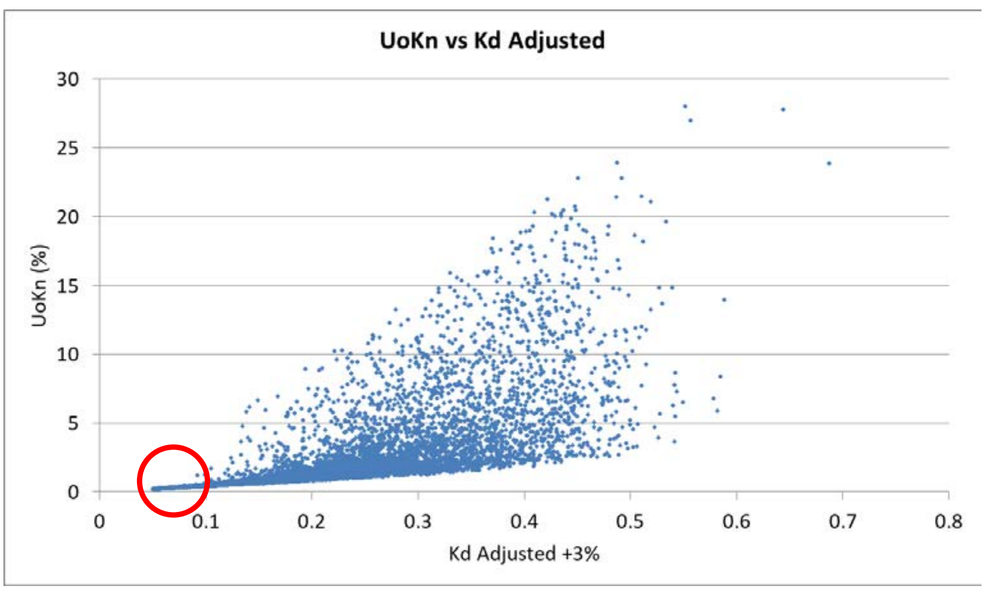

Figure 6-2. $U_O Kn$ versus Kd with introduced errors

The red circle indicates the area of interest.

A clear-sky filter can be applied by limiting Kn and Kd in the processing data set:

$$Kn > 0.5 \text{ and } Kd < 0.1 \text{ (These limits could be site specific.)}$$

The resulting data are processed as usual by Eq. 8. Figures 6-3, 6-4, and 6-5 show the results of the processing for the three test stations. Each left-hand figure plots the Kn compared to the Kt for the clear-sky subset of data (compare to Figure 5-1), and each right-hand figure plots the resulting $U_O Kn$ values as a function of Kn.

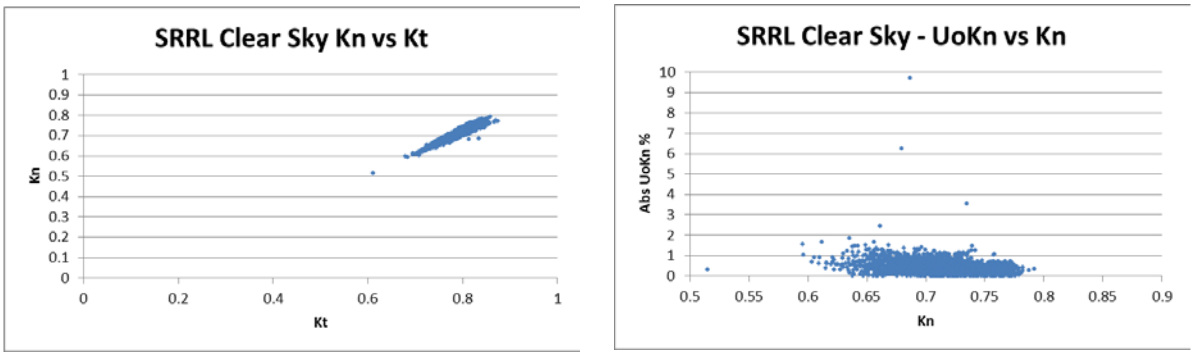

Figure 6-3. Clear-sky U_OKn processing for the SRRL

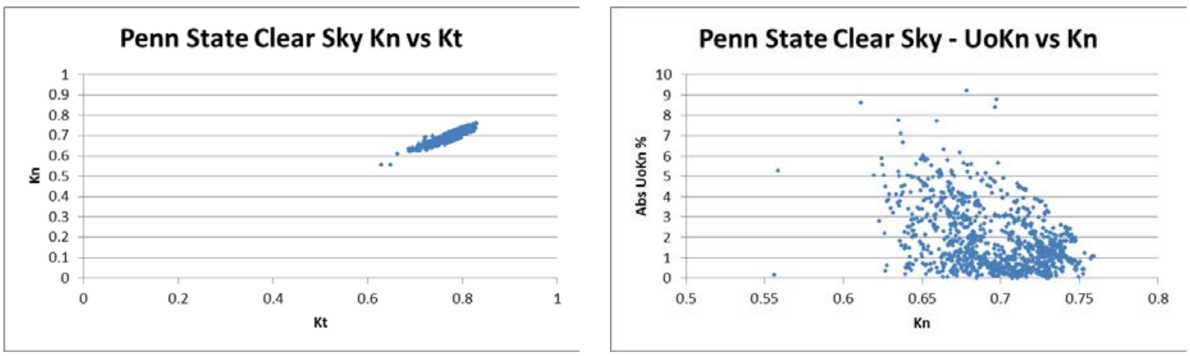

Figure 6-4. Clear-sky U_OKn processing for Penn State

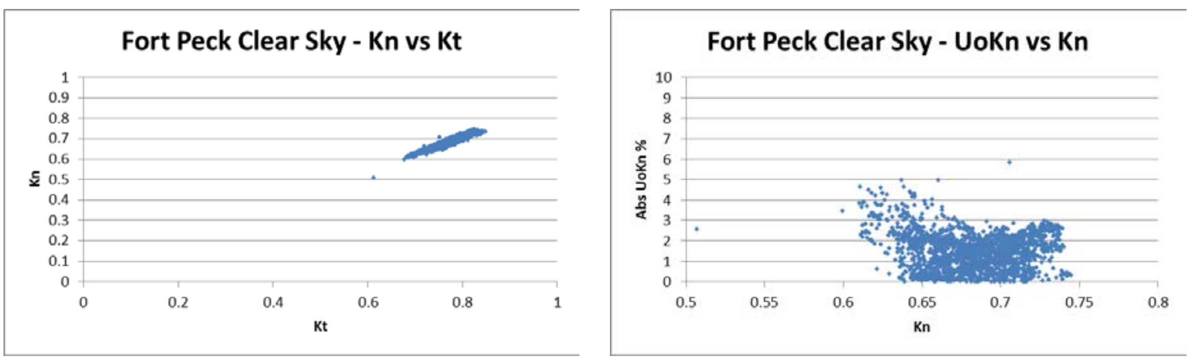

Figure 6-5. Clear-sky U_OKn processing for Fort Peck

Table 6-1 shows summary statistics for the clear-sky U_OKn for each station. For comparison, the U_OKt values for each clear-sky data set are included. The table also shows the percentage of U_OKn data points that exceed the U_{95} base uncertainty for the DNI instrument (instrument base uncertainty estimated from the NREL spreadsheet) (Habte 2014).

Table 6-1. Clear-Sky U_OKn Summary Statistics

	SRRL	Penn State	Fort Peck
Clear-sky U_OKn mean	0.44%	1.8%	1.5%
Clear-sky U_OKt mean	0.40%	1.6%	1.3%
Base U95	2.3%	2.5%	2.5%
Percentage exceeding base	0.1%	27.3%	12.4%

Although carryover errors from Kd to U_OKn will be nearly eliminated in this clear-sky data subset, carryover errors in Kt will still be evident. Figure 6-6 shows U_OKn compared to Kt adjusted +3% data that have been filtered for clear-sky values. The exaggerated carryover of the +3% errors shown in the top middle plot of Figure 5-5 (some greater than 20%) are limited to no more than 3.5% in the clear-sky filtered data set.

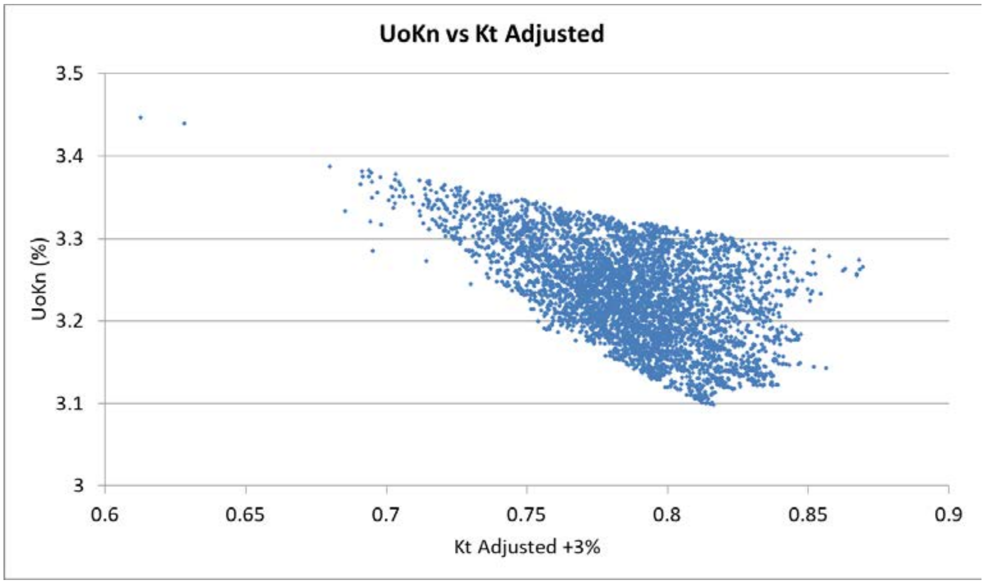

Figure 6-6. U_OKn versus Kt adjusted +3% for clear-sky filtered data

This analysis shows that a clear-sky subset of data can be used to eliminate most carryover errors from Kd to U_OKn; however, some carryover might be present from errors in Kt, which could be the cause of the higher statistics at Penn State and Fort Peck.

6.3 Cloudy-Sky U_OKd

The analysis in Section 5 shows that U_OKd can be exaggerated by crossover errors from Kn during periods of high DNI and low DHI (generally very clear skies). By limiting the analysis to cloudy-sky data records, much of the problem data can be eliminated from the estimate of U_O. The Kt-Kn scatterplots (Figure 5-1) place cloudy-sky data in the cluster of points near the bottom of the plot.

By plotting U_OKd as a function of Kn, it becomes obvious that regions of low Kn result in low Kn crossover to U_OKd. This is illustrated in Figure 6-7, which shows data from the simulated error data set (Section 4.1.5). Here, the crossover to U_OKd approaches zero as Kn (with the 3% introduced DHI error) approaches zero.

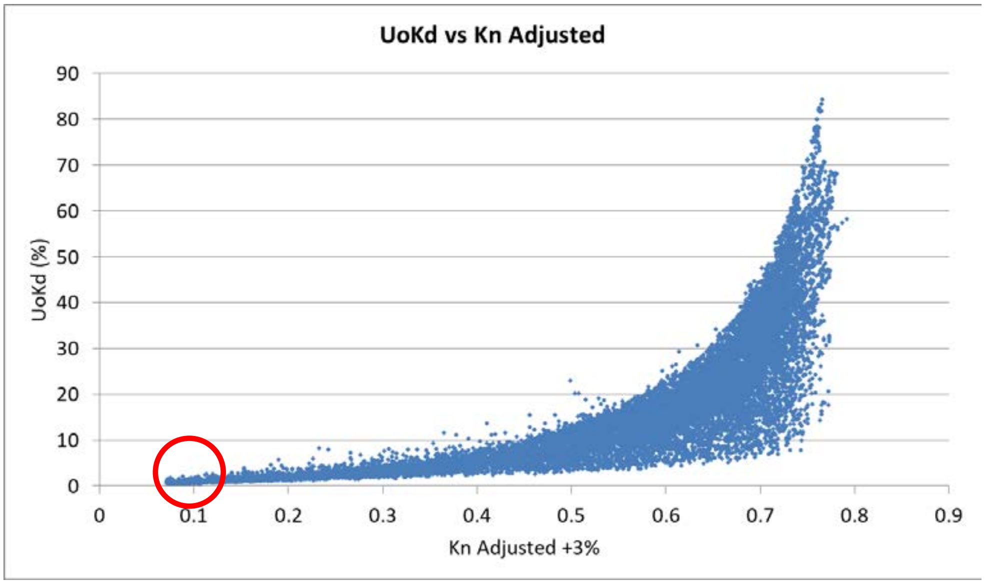

Figure 6-7. U_OKd versus Kn with introduced errors

The red circle indicates the area of interest.

A cloudy-sky filter can be applied by limiting Kn and Kd in the processing data set:

$$Kd > 0.2 \text{ and } Kn < 0.1 \text{ (These limits could be site specific.)}$$

The resulting data are processed as usual by Eq. 9. Figures 6-8, 6-9, and 6-10 show the results of the processing for the three test stations. Each left-hand figure plots the Kn compared to the Kt for the cloudy-sky subset of data (compare to Figure 5-1), and each right-hand figure plots the resulting U_OKd values as a function of Kd.

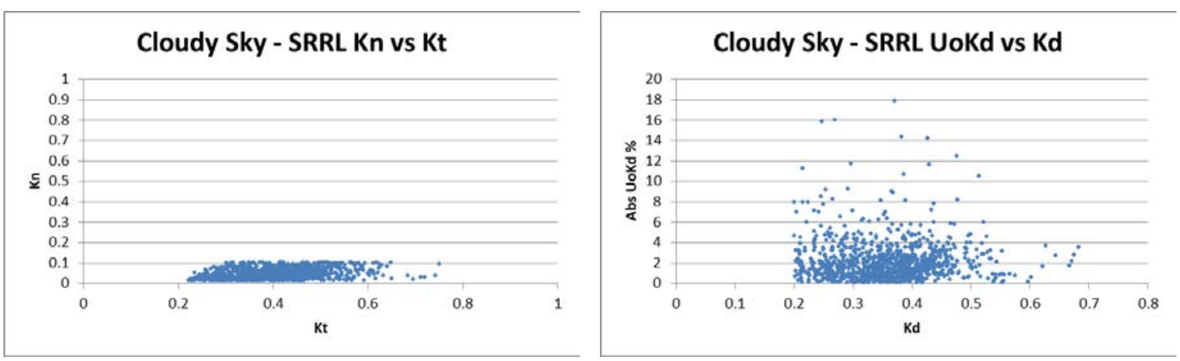

Figure 6-8. Cloudy-sky U_OKd processing for the SRRL

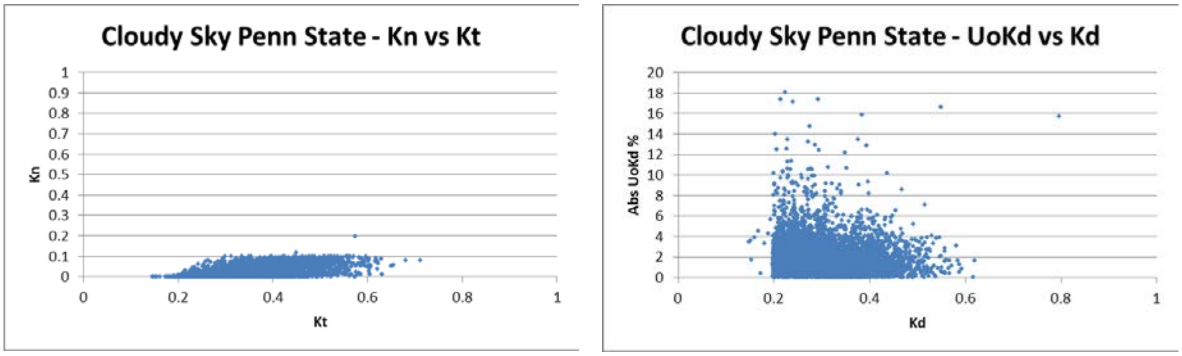

Figure 6-9. Cloudy-sky U_OKd processing for Penn State

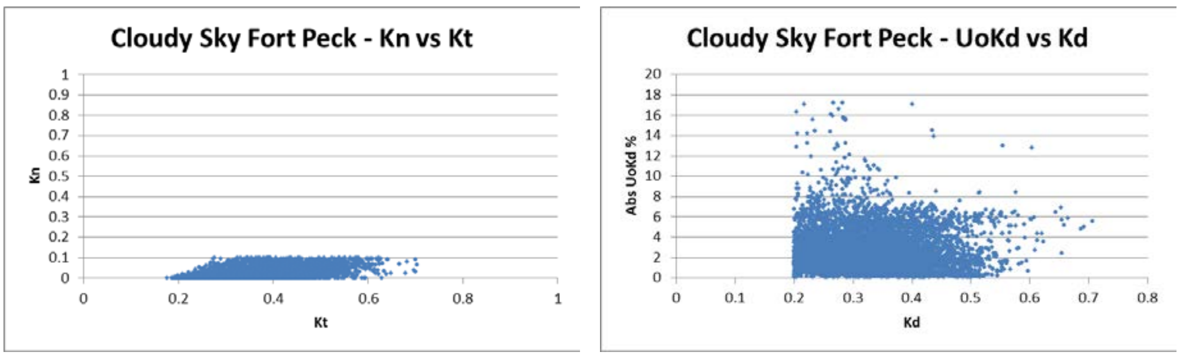

Figure 6-10. Cloudy-sky U_OKd processing for Fort Peck

Table 6-2 shows summary statistics for the clear-sky U_OKn for each station. For comparison, the U_OKt for each clear-sky data set is included.

Table 6-2. Cloudy-Sky U_OKd Summary Statistics

	SRRL	Penn State	Fort Peck
Cloudy-sky U_OKn mean	2.0%	1.8%	2.8%
Cloudy-sky U_OKt mean	1.8%	1.6%	2.6%

Although carryover errors from Kn to U_OKd will be nearly eliminated in this cloudy-sky data subset, carryover errors in Kt will still be evident. Figure 6-11 shows U_OKd versus the Kt adjusted +3% data that have been filtered for cloudy-sky values. The exaggerated carryover of the +3% errors shown in the upper right plot of Figure 5-5 (some greater than 30%) are limited to no more than 4.2% in the cloudy-sky filtered data set.

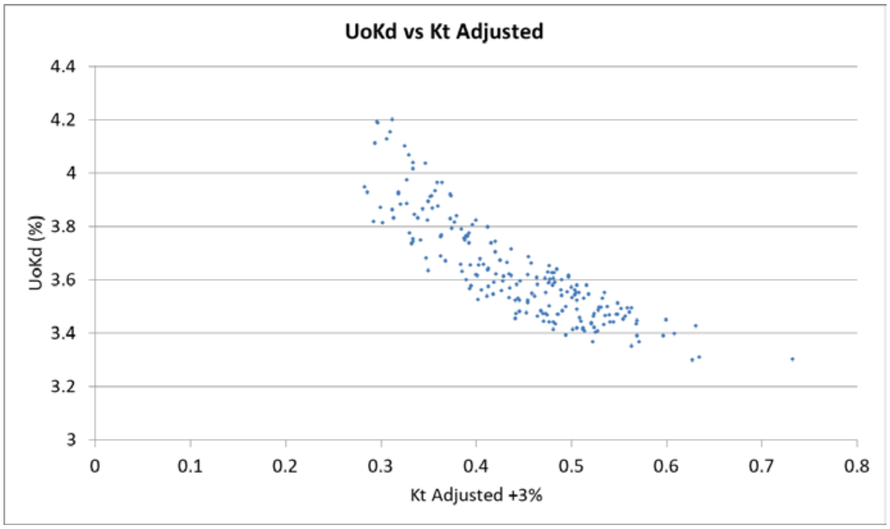

Figure 6-11. U_OKd versus Kt adjusted +3% for cloudy-sky filtered data

This analysis shows that a cloudy-sky subset of data can be used to eliminate most carryover errors from Kn to U_OKd; however, some carryover might be present from errors in Kt because of irradiance coupling.

6.4 Using SERI QC Flags to Determine U_O

Given the previous analysis, there is an opportunity to investigate the use of the SERI QC flags to estimate the U_O. For each station, the U_OKt was plotted as a function of the assigned SERI QC flags. Although these plots show a strong correlation between the SERI QC flags and U_OKt, they also reveal a range of U_OKt values for each flag. This indicates that the flags could not provide U_O estimates that are as precise as the proposed uncertainty equations. A similar analysis of U_OKn and U_OKd would show an even greater range of values for each SERI QC flag. Note that the original SERI QC flagging method was developed to account for an estimated ±3% radiometer measurement uncertainty associated with the better-performing instruments used at the time. This measurement uncertainty tolerance produces a range of estimated operational uncertainties for a specific SERI QC flag assignment; therefore, using the precise K-space values to determine the U_O results in the broad flag-to-uncertainty behaviors shown in Figure 6-12.

The SERI QC flags still play a valuable role in the determination of U_O by translating the irradiance data into K-space and establishing a logical data quality assessment process that specifically identifies data unsuitable for the uncertainty equations.

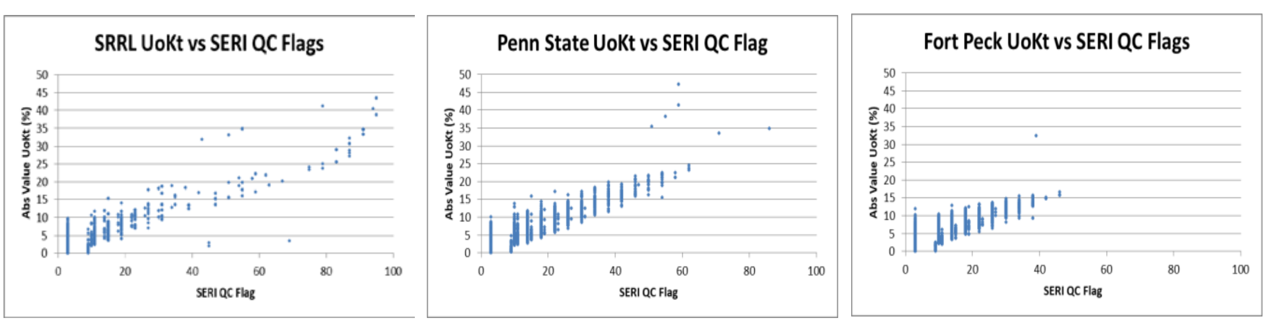

Figure 6-12. U_OKt as a function of the SERI QC flags

7 Conclusions

The bankability, efficiency, profitability, and compliance of solar energy conversion systems can be improved by quantifying and reducing the uncertainty in predicting solar-generated energy output using an integrated approach for estimating solar resource data quality and radiometer measurement uncertainties.

We have proposed a method for estimating U_O of the key solar irradiance measurements (global, direct, and diffuse) based on the existing SERI QC software to quantify errors resulting from substandard measurement conditions (e.g., improper maintenance, weather-induced or environmentally caused optical contamination, improper equipment installation). This method is designed to operate within the context of best practices for solar irradiance measurements and is not intended for measurement campaigns without strict operational protocols.

A proof of concept was developed, including programming in C and a collection of spreadsheet tools. In the absence of a measurement reference for radiometers operationally deployed, a field reference for determining U_O was established for each irradiance component based on the SERI QC coupling in Eq. 6–Eq. 9. After examining the behavior of U_O for each of the three irradiance components, including a fixed 3% bias, we determined that the U_O estimates based on global irradiance (U_OKt), Eq. 7, best represented the collective results for all sky conditions.

Further examination of the U_O concepts included separate analysis of clear-sky and cloudy-sky conditions to explore the crossover effects due to the intrinsic interdependence of the three irradiance components. Determining U_OKn for clear-sky and U_OKd for cloudy-sky conditions reduces the crossover effects on U_O based on U_OKt and offers additional options for assessing solar resource data uncertainty, particularly for clear-sky power production conditions.

The proposed method for estimating U_O has been applied to data from three measurement stations in the United States operated by NREL and NOAA during 2021. Examination of the results indicates U_OKt provides a representative measure of U_O consistent with the operation and maintenance practices applied to the stations, ranging from approximately a fraction of a percentage to 3%. These methods, though based on the fluid reference of field data, will provide analysts with valuable insight into additional uncertainty attributable to operational errors.

In developing the proof of concept for the SERI QC flag translation, we observed the following regarding the estimation of U_O:

- Advantages:
  - Measurement data provide U_O estimates without additional information or external references.
  - Estimates of U_O quantify data quality with more precision than the SERI QC method that produces data quality flags representing varying ranges of data uncertainty.
  - The concept provides alternative estimates for data subsets under clear-sky conditions for direct irradiance (DNI) and cloudy-sky conditions for diffuse irradiance (DHI) measurements.

- A single $U_O Kt$ for a data record can capture and represent much of the error in DNI and some of the error in DHI.

- Disadvantages:
  - No independent measurement reference was used to estimate U_O.
  - The ambiguity of fault among the three components was not resolved.
  - The concept cannot identify and account for measurement error cancellation or a shared bias among the radiometers (e.g., uniform soiling or error in the calibration reference).
  - Estimating U_O from global measurements ($U_O Kt$) might slightly underestimate the uncertainties in direct (DNI) and significantly underestimate errors in diffuse (DHI) irradiances, depending on the sky conditions and the relative levels of solar irradiance.

8 Next Steps

We have identified the following areas of special operational concern and suggestions for discussion topics with NREL staff in preparation for accomplishing the next task of developing the translation method:

1. Identify possible means of reporting uncertainties when:
    A. Kt = Kd or Kt = Kn
    B. Irradiance levels are affected by significant noise or operational errors
    C. Irradiance measurements at high SZAs produce unrealistic values from ratios of small numbers.
2. Address any need for selecting clear-sky and cloudy-sky conditions when determining U_O.
3. Select an appropriate DNI threshold for determining U_O if it is other than the current value of 25 W/m^2.
4. Further develop the U_O methodology and integration with radiometer measurement uncertainty estimates.

References

Habte, A. 2014. Spreadsheet for Estimating Radiometer Uncertainties Available from the NREL Measurement & Instrumentation Data Center. https://midcdmz.nrel.gov/radiometer_uncert.xlsx.

Maxwell, E., S. Wilcox, and M. Rymes. 1993. *Users Manual for SERI QC Software—Assessing the Quality of Solar Radiation Data*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-463-5608. http://www.nrel.gov/docs/legosti/old/5608.pdf.

Sengupta, Manajit, Aron Habte, Stefan Wilbert, Christian Gueymard, and Jan Remund, 2021. *Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications: Third Edition*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-77635. https://www.nrel.gov/docs/fy21osti/77635.pdf.