Evaluation and Suitability of Using SERI QC Software for Estimating Measurement Uncertainty

Stephen Wilcox and Thomas Stoffel

Solar Resource Solutions, LLC

NREL Technical Monitor: Aron Habte
Evaluation and Suitability of Using SERI QC Software for Estimating Measurement Uncertainty

Stephen Wilcox and Thomas Stoffel

Solar Resource Solutions, LLC

NREL Technical Monitor: Aron Habte

Suggested Citation
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 to 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 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 the application of the newly developed code/method.
  - Develop, verify, and validate a new software package consistent with the project goal.

This technical report addresses the first objective in FY 2022.

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 No. DE-AC36-08GO28308.

---

\(^1\) See [https://midcdmz.nrel.gov/radiometer_uncert.xlsx](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

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARM</td>
<td>Atmospheric Radiation Measurement</td>
</tr>
<tr>
<td>DQMS</td>
<td>Data Quality Management System</td>
</tr>
<tr>
<td>ETR</td>
<td>extraterrestrial radiation</td>
</tr>
<tr>
<td>ETRn</td>
<td>direct normal extraterrestrial radiation</td>
</tr>
<tr>
<td>GNDRAD</td>
<td>Ground Radiation</td>
</tr>
<tr>
<td>ISIS</td>
<td>Integrated Station Information System</td>
</tr>
<tr>
<td>K.A.CARE</td>
<td>King Abdullah City for Atomic and Renewable Energy</td>
</tr>
<tr>
<td>KACST</td>
<td>King Abdulaziz City for Science and Technology</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmospheric Administration</td>
</tr>
<tr>
<td>NREL</td>
<td>National Renewable Energy Laboratory</td>
</tr>
<tr>
<td>QCFIT</td>
<td>software for determining boundaries of acceptable solar irradiance data</td>
</tr>
<tr>
<td>SERI QC</td>
<td>software function for post-measurement quality assessment of solar irradiance data</td>
</tr>
<tr>
<td>SERI</td>
<td>Solar Energy Research Institute</td>
</tr>
<tr>
<td>SIRS</td>
<td>Solar Infrared Radiation Station</td>
</tr>
<tr>
<td>SKYRAD</td>
<td>Sky Radiation</td>
</tr>
<tr>
<td>SOLRMAP</td>
<td>Solar Resource and Meteorological Assessment Project</td>
</tr>
<tr>
<td>SURFRAD</td>
<td>Surface Radiation Budget</td>
</tr>
<tr>
<td>SZA</td>
<td>solar zenith angle</td>
</tr>
<tr>
<td>TSI</td>
<td>total solar irradiance</td>
</tr>
<tr>
<td>WRR</td>
<td>World Radiometric Reference</td>
</tr>
</tbody>
</table>
Executive Summary

SERI QC is a robust solar data quality assessment software tool that has been in continuous use for more than three decades. This report (the first of six in the Data Quality and Uncertainty Integration project) reviews and evaluates the software to determine its suitability for determining the estimated expanded measurement uncertainty of solar measurement data by including operational factors. Using redundancy incorporated in three-component (global, direct, diffuse) solar measurements, SERI QC can quantify the standard uncertainty estimates for estimating uncertainty.

With minor modifications, SERI QC will provide an operational uncertainty to be used in conjunction with the National Renewable Energy Laboratory method for instrument uncertainty to provide an integrated estimate of solar measurement data sets. These modifications will capitalize on SERI QC’s evaluation of data quality, including data input validation and a variety of built-in solar routines, and will further develop the utility of the software.
# Table of Contents

1 Introduction.................................................................................................................. 1  
1.1 The Scope of This Report............................................................................................... 1  
1.2 The Goal of This Report............................................................................................... 1  
1.3 Difficulties of Solar Measurements............................................................................ 2  
2 SERI QC Background.................................................................................................... 4  
2.1 The Purpose of SERI QC........................................................................................... 4  
2.2 Authorship................................................................................................................ 4  
2.3 SERI QC Development Goals.................................................................................. 4  
3 SERI QC Modes of Operation....................................................................................... 6  
3.1 K-Space Representation of the Data......................................................................... 6  
3.2 Narrowing the Expected Values............................................................................. 7  
3.3 The One-Component Test....................................................................................... 7  
3.4 The Two-Component Test...................................................................................... 8  
3.5 The Three-Component Test.................................................................................... 9  
4 The Gompertz Boundaries............................................................................................ 11  
4.1 The Gompertz Function......................................................................................... 11  
4.2 Determining the Gompertz Coefficients............................................................... 12  
4.3 Maximum Diffuse Values..................................................................................... 14  
5 QCFIT............................................................................................................................ 16  
5.1 The Original QCFIT................................................................................................. 16  
5.2 An Enhanced QCFIT............................................................................................... 16  
6 Historical Uses of SERI QC ....................................................................................... 18  
7 Pros and Cons of Continued Use of SERI QC............................................................. 20  
7.1 The Pros............................................................................................................... 20  
7.2 The Cons............................................................................................................. 20  
8 Suitability of SERI QC for Uncertainty Estimates......................................................... 21  
8.1 Description of the Flags....................................................................................... 21  
8.2 Determining a Standard for Operational Uncertainty........................................... 21  
8.3 Assumptions and Concerns................................................................................... 22  
8.4 Application of the Operational Uncertainty......................................................... 24  
8.5 A Foundation of Best Practices......................................................................... 24  
9 Modifications to SERI QC for Deriving Uncertainty.................................................. 25  
9.1 Required Modifications to SERI QC................................................................. 25  
9.2 Additional Changes to SERI QC to Consider.................................................... 26  
10 Conclusions and Recommendations...................................................................... 27  
10.1 Suitability of SERI QC for Supporting Operational Uncertainty.................... 27  
10.2 Recommendations for Implementation.............................................................. 27  
References....................................................................................................................... 28  
Appendix A. SERI QC Flagging Convention................................................................. 30  
Appendix B. SERI QC Flowchart.................................................................................. 31  
Appendix C. QC0 File.................................................................................................... 32  
Appendix D. Gompertz Curve Definitions.................................................................... 33
List of Figures

Figure 3-1. Scatterplot data........................................................................................................................... 8
Figure 4-1. Gompertz plot in K-space ........................................................................................................ 13
Figure 4-2. Gompertz plot in K-space ........................................................................................................ 14
Figure 4-3. Plot of diffuse maximum values............................................................................................... 15
Figure 5-1. QCFIT user interface.............................................................................................................. 17
Figure 6-1. XYZ plot of quality flags and K-space measurements............................................................. 19
Figure 8-1. Atmospheric mapping of Kt-Kn space..................................................................................... 23
Figure AA-1. SERI QC flagging convention.............................................................................................. 30
Figure AB-1. SERI QC flowchart............................................................................................................... 31

List of Tables

Table 3-1. Air Mass Ranges Used by SERI QC ........................................................................................... 7
Table 4-1. Derived Gompertz Coefficients for the Existing SERI QC Catalog of Boundaries ............... 12
1 Introduction

1.1 The Scope of This Report
This report is the first of six deliverables in Data Quality and Uncertainty Integration project assigned to Solar Resource Solutions, LLC. The focus is to review the existing SERI QC software package and to provide insight into how it can be used to further refine the uncertainty of solar resource data.

1.2 The Goal of This Report
This report documents major attributes of the SERI QC data quality assessment software package and evaluates its suitability as a foundation for determining an operational uncertainty of solar measurement data that can contribute to the overall uncertainty of a solar radiation resource data set.

1.2.1 Determining Uncertainty from Quality Assessment
This work seeks to further develop the uncertainty estimates of individual solar measurement instruments by quantifying and incorporating additional uncertainty derived from applying the SERI QC data quality assessment method to field measurement campaigns. A subsequent task will develop a software method that allows station operators to more easily and transparently assign a comprehensive measurement uncertainty to data streams emanating from solar radiation measurement setup and operations.

1.2.2 Data Quality Assessment Versus Data Quality Control
This method will build on quality assessment and quality control procedures developed and enhanced by the National Renewable Energy Laboratory (NREL) during the past several decades. Both data quality control and data quality assessment are components of data quality assurance, which ensures the overall integrity of the monitoring program or station. We define data quality control as:

A supervisory process by which routine and specialized operations are subjected to standards of operation that ensure the best outcome of a measurement.

We define data quality assessment as:

A critical component of quality control that subjects a measurement to scrutiny to produce a quantifiable judgement of quality. This assessment of quality becomes feedback to the quality control process to allow operators the ability to correct equipment malfunctions or measurement errors and to refine the quality control process as appropriate.

The SERI QC data quality assessment software was developed to automate the assessment of large quantities of solar irradiance measurements, to provide the means for further summarizing data quality, and to identify periods of quantifiable measurement errors for further investigation.
1.3 Difficulties of Solar Measurements

1.3.1 Fundamental Uncertainties of Solar Measurements

The measurement of solar irradiance is a difficult endeavor involving multiple physical properties (e.g., optical, electrical, thermal) for instrument design, proper installation methods (including capable data acquisition), and specialized operation-and-maintenance procedures (including equipment recalibration). The propagation of measurement uncertainty starts with calibrations traceable to an accepted reference.

The present internationally recognized measurement reference for direct normal solar irradiance is the World Radiometric Reference (WRR) maintained by the World Radiation Center, in Davos, Switzerland, for the World Meteorological Organization. The WRR has an accepted measurement uncertainty of ±0.3% with a 99% confidence interval, thus defining the present lower limit to the uncertainty of solar irradiance measurement (World Meteorological Organization 2018, 248). In the most fundamental realm, when the WRR is transferred to other instruments, the resulting estimated measurement uncertainty is no better than ±0.4% with 99% confidence (Reda et al. 2019), which is much greater than uncertainties in other areas of metrology, such as temperature, length, voltage, and the like. Much of this is because of the necessity of measuring solar irradiance through indirect means with a thermoelectric or photoelectric detector and the many atmospheric and environmental effects that introduce a variety of instrument responses not easily accounted for in the measurement itself.

And, finally, field instruments are largely unattended throughout a day, week, or longer, and are thus subject to environmental effects or contaminants that are otherwise better controlled in the operation of reference instruments. Through strict adherence to best practices, many of these effects can be minimized; however, relative to the uncertainty of reference instruments, these additional operational uncertainties are significant—perhaps by a factor of two or three under the best operational procedures—and must be accounted for to provide a critical uncertainty for a final analysis of the data quality.

1.3.2 Reliance on Best Practices

Because field measurements are often accomplished in remote areas and in some cases in harsh environments, the instruments can be subjected to a wide range of meteorological conditions that challenge the instrument design specifications and significantly differ from the calibration conditions, thus increasing measurement uncertainties of the operational data. Recognizing the importance of proper instrument selection, installation, operation, and maintenance, a standard of care, or best practices, has been established by NREL and International Energy Agency partners (Sengupta et al. 2021).

The Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications goes into depth helping users develop procedures to ensure high-quality data acquisition for solar measurements. The team of international authors have formed a

---

2 The World Radiation Center is also developing an improved reference based the Cryogenic Solar Absolute Radiometer to reduce the estimated measurement uncertainty of the WRR.
3 As applied to pyrheliometers for measuring direct normal irradiance.
consensus on the need for such practices and have shared their expertise to help readers establish quality control and quality assessment procedures prior to commencing a measurement campaign.

Because of the additional factors contributing to the uncertainties imposed by typical field measurements, determining an accurate and reasonable final uncertainty of a measurement will heavily rely on best practices to ensure that the instruments are maintained in the best form. Performing data quality assessment on poorly maintained instruments results in a high uncertainty that cannot be quantified other than with a general “quality unknown” designation.

Although many field instruments are assigned a calibration uncertainty ranging from ±1 to ±3%, it is unrealistic to carry that uncertainty forward to measurement values without also considering the effects of operating conditions; thus, a measurement value is expected to have a higher uncertainty than the assigned calibration uncertainty of the measuring instrument, but strict adherence to best practices will help keep the additional uncertainty to a minimum.
2  SERI QC Background

2.1  The Purpose of SERI QC

SERI QC was developed around 1989 by the Solar Energy Research Institute (SERI), which was later renamed the National Renewable Energy Laboratory. Development was funded by the U.S. Department of Energy in anticipation of an increasing need for high-quality solar irradiance measurements and the necessity of monitoring the quality of solar resource data. Several large solar irradiance data sets existed at the time, and with the expectation that many more field campaigns would be set in motion, the need for automated data quality assessment was identified. SERI QC was envisioned to meet that need to provide the capability of quickly assessing the quality of available data from a variety of measurement paradigms, accepting data from different time integrations and combinations of commercially available instruments.

2.2  Authorship

Eugene Maxwell led the development of SERI QC. He brought several innovations to solar irradiance data quality assessment, including the use of Gompertz curves to define digital bounds to the Kt-Kn scatterplots of expected values. Maxwell was assisted by Martin Rymes, who provided programming and mathematical support. Rymes translated the algorithms with expert programming knowledge and devised a method of encoding a variety of evaluation results into simple two-digit flags. He also provided programming for the original QCFIT program and developed a curve-fitting algorithm to automatically select boundary curves, allowing users to easily establish expected value boundaries from historical data at a site.

The robust SERI QC design and implementation has endured for several decades and provided routine data quality assessment for numerous data sets around the world.

Upon completion of development, the software was documented in the Users Manual for SERI QC Software—Assessing the Quality of Solar Radiation Data (Maxwell et al. 1993), hereafter referred to as the user’s manual. The user’s manual includes in-depth explanations of how the software was developed and the theory behind its operation. This report summarizes pertinent parts of the user’s manual.

2.3  SERI QC Development Goals

2.3.1  Designed for Versatile Implementations

With the knowledge that measurement campaigns can be designed for a wide variety of specialized purposes, and that no data acquisition standards existed, SERI QC was envisioned with the capacity to be incorporated into data acquisition software; thus, SERI QC was implemented as a Fortran function rather than a stand-alone program with a permanent purpose and fixed inputs and outputs. It processes only one data record for a single time stamp at a time and returns the results to the calling program.

With this design, a data processing program specific to a particular measurement campaign could be easily modified to direct the data stream via a function call to SERI QC, which would evaluate the data and return the appropriate flags associated with the data; hence, SERI QC does not interact with a user or even a data acquisition system. Rather, a calling program takes care of
all user interfaces and data acquisition as required by a particular project, then submits the data to SERI QC for evaluation. The calling program would then use the flag information returned by the SERI QC function for further processing, documentation, or evaluation. SERI QC does not delete or filter any data but rather provides information for external programs or analyses to carry out those details depending on the purpose of the analysis.

As the user base for SERI QC expanded, the Fortran code was translated by Augustyn+Company to the C programming language for incorporation in their proprietary Data Quality Management System (DQMS). NREL retained ownership of the C code. Most recently, the C code was translated by NREL into Python for use in this project. This section describes SERI QC as a function as originally developed. The Python version includes a user interface that can read, process, and output data files.

2.3.2 Function Inputs
The calling program passes several parameters to the SERI QC function:

- Global horizontal irradiance (single measurement)
- Direct normal irradiance (single measurement)
- Diffuse horizontal irradiance (single measurement)
- Measurement time stamp (year, month, day, hour, minute, integration time)
- Station latitude, longitude, and elevation
- A K-space boundary definition file for that station.

2.3.3 Function Outputs
After data evaluation, the SERI QC function passes flags back to the calling program for each measurement input. These two-digit flags (00–99) indicate either a successful evaluation, or in the event of errors, they indicate the mode, magnitude, and direction of a discrepancy.

As the flags increase in magnitude, they generally relate to increasing discrepancies in the data. The SERI QC flag scheme encodes the measurement discrepancy, the type of test performed, and the direction of the discrepancy; thus, the calling program can decode a flag, if desired, to extract precise information about the measurement evaluation. Appendix A documents all attributes of SERI QC flags.

---

4 See https://github.com/NREL/SolarResourceTools/tree/master/SERI-QC_with_QCFIT.
3 SERI QC Modes of Operation

SERI QC operates in three data realms depending on the number of solar irradiance parameters of the three possible (global, direct, diffuse) being sent to the function. Note that SERI QC can work on a single component, any combination of two of the three, or all three. In the first two cases (one- and two-component data), SERI QC relies on a significant amount of historical data for a site to establish limits by which expected values can be defined. These limits vary from one site to another and represent the effects of varying atmospheric effects on the solar energy, such as from clouds, water vapor, aerosols, and the like. This section describes the modes of testing as shown in a flowchart in Appendix B.

3.1 K-Space Representation of the Data

SERI QC converts irradiance measurements from W/m² to K-space, which is a normalized representation that is independent of the effects of the atmosphere and station location. For each parameter, the measurement is normalized (divided) by the like measurement as if observed at the top of the atmosphere without any atmospheric attenuation, here referred to as extraterrestrial.

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

\[ \text{ETRn} = \text{TSI} \times \left(\frac{R}{R_o}\right)^2 \]  

where:

- \( \text{TSI} \) = total solar irradiance (1360.8 ±0.5 W/m²)
- \( R \) = sun-Earth distance at the time of interest
- \( R_o \) = annual mean sun-Earth distance

and the global horizontal extraterrestrial irradiance (ETR) is computed:

\[ \text{ETR} = \text{ETRn} \times \cos (\text{SZA}) \]

where:

- \( \text{SZA} \) = solar zenith angle at the location, date, and time of interest.

Thus:

\[ K_t = \frac{\text{global}}{\text{ETR}} \]

\[ K_n = \frac{\text{direct}}{\text{ETRn}} \]

\[ K_d = \frac{\text{diffuse}}{\text{ETR}} \]

Note that both \( K_t \) and \( K_d \) use the global horizontal extraterrestrial. These three parameters are related in this closure equation:

\[ K_t = K_n + K_d \]

5 SERI QC was developed with a TSI value of 1367 W/m², and the operational difference is negligible.
3.2 Narrowing the Expected Values

Because of significant variations in atmospheric attenuation related to the weather conditions at a particular site throughout the year and even throughout a day, SERI QC establishes expected values for a particular measurement based on the position of the sun in the sky, the month of the year, and the integration interval of the measurement.

3.2.1 The Effects of Air Mass

Knowing that any effects of atmospheric attenuation can be amplified by the amount of atmosphere the radiation passes through, SERI QC establishes three air mass\(^6\) regimes: low, medium, and high. The high air mass occurs when the sun is near the horizon and the path through the atmosphere is longest. Low air mass occurs when the sun is high in the sky and the path length through the atmosphere is relatively less. Medium air mass occurs between the high and low air masses. The three air mass ranges used by SERI QC are presented in Table 3-1.

During winter months, many locations produce no low air mass data because the sun path never passes high enough in the sky.

<table>
<thead>
<tr>
<th>Ranges</th>
<th>Air Mass</th>
<th>Solar Zenith Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>1.00–1.25</td>
<td>0–36.96°</td>
</tr>
<tr>
<td>Medium</td>
<td>1.25–2.50</td>
<td>36.96–66.57°</td>
</tr>
<tr>
<td>High</td>
<td>2.50–5.76</td>
<td>66.57–80.00°</td>
</tr>
</tbody>
</table>

3.2.2 The Effects of Season

Throughout the year, seasonal changes in weather patterns and the path of the sun across the sky affect solar radiation intensity. A cloudy or humid season at a particular site might limit the range of solar intensity, information that is very useful in setting expected values for a solar irradiance measurement. To capitalize on this effect, SERI QC establishes bounds for the data in monthly increments, which is generally a suitable resolution for discerning seasonal effects.

3.2.3 The Effects of Integration Time

As higher resolution data are integrated (or averaged) over a time interval to a lower resolution realm, extremes in the measurement period are lost. As a result, characteristics of an integrated data set will typically have a narrower range of irradiance and thus a smaller range of expected values.

3.3 The One-Component Test

For both \(K_t\) (global) and \(K_n\) (direct), an analysis of historical data establishes the maximum value of each parameter that is likely to occur. These are referred to \(K_t\)-max and \(K_n\)-max, and they are the highest values encountered in the historical data (Figure 3-2). The maximum value for \(K_d\) (diffuse) is determined according to another analysis, which is described in Section 4.3.

---

\(^6\) Air mass describes the optical path length of the direct normal irradiance through the atmosphere. Air mass 1.0 is when the sun is directly above a sea-level location under ASTM standard atmospheric conditions for a cloudless sky.
A minimum value for each parameter was derived from an analysis of instrument behavior. A simple understanding of solar irradiance easily implies that the lowest value of irradiance must be 0.0 W/m²; however, many instruments in common use will actually report a negative irradiance under many measurement scenarios, much of it as a result of instrument sensitivity to longer wavelengths of irradiance in the infrared during night. This effect can also occur during periods of partly cloudy skies that produce rapidly changing irradiance values.

Although these negative irradiances are incorrect for visible light, they are nonetheless a well understood instrument artifact, and the developers of SERI QC allowed for a Kt- and Kn-minimum of -10 W/m².

Using these maximum and minimum values, an incoming measurement is compared to that range, and if it is out of bounds, it receives a flag indicating the value is too high (flag 08) or too low (flag 07). Otherwise, it is given a flag of 01 (passing the one-component test). No further tests of the single parameter are performed; however, this test is also performed on measurements arriving as part of a two- or three-component data set before progressing to subsequent tests. In these cases, if the one-component test fails, no further testing is performed.

3.4 The Two-Component Test

Any paired combination of a three-component measurement is subjected to the two-component test if both parameters pass the one-component test. To accomplish this test, historical data are first used to establish the bounds of the relationship between global and diffuse by examining the range of Kn as a function of Kt. This is best visualized as a scatterplot of Kn by Kt (Figure 3-1). Note the hand-drawn boundaries, which help define the scatter envelope.

![Figure 3-1. Scatterplot data](Typical scatterplot of Kn vs. Kt data along with Kt and Kn max designations (Maxwell et al. 1993))

7 Radiometers are now designed to minimize or eliminate this thermal offset.
The width, height, and overall shape of the scatterplot are driven by the effects of a site’s climate and typical atmospheric attenuation, and the envelope for each site will generally be unique. Also note the dotted lines that indicate the Kt and Kn maximums, as described in the previous section. The scatter of data, if bounded, forms a predictable and characteristic data envelope. If created with historical data of sufficient quantity and quality, it is within these bounds that future data can, with high confidence, be expected to appear.

One challenge of SERI QC development was to create digital representations of envelope boundaries such that an automated process could compare the position of a Kt-Kn pair to determine if it falls within the envelope boundaries. To accomplish this goal, the Gompertz function was employed to establish boundaries. In practice, the Gompertz function was used to create a catalog of boundary lines suitable for fitting to scatter envelopes. The Gompertz function as it relates to SERI QC is examined in more depth in Section 4.

A digital reference to the selected Gompertz curves as well as the Kt- and Kn-max are stored in a station-specific data file called the QC-Zero file (Appendix C). The file contains these descriptions for the three air masses, each month of the year, and four integration interval ranges. This file is created by a companion software package called QCFIT (see Section 5), which allows analysis and curve selection based on a site’s historic data.

Operationally, a set of input measurements is evaluated in Kt-Kn space with the Gompertz boundaries providing expected values by which the data can be evaluated. If found within the boundaries, the data pass the two-component test (flag 02). If a data pair falls outside the boundaries, it is given a flag encoded with the distance it resides beyond the boundary, indicating the severity and direction of the discrepancy.

The two-component test suffers from substantial ambiguity in the expected values. From the scatterplot (Figure 3-1), it is apparent that for a particular value of Kt, a range of Kn values is acceptable; thus, a measurement could be in error, and as long as it fell within the width or height of the boundaries, it would pass the evaluation. In common circumstances, this error could be 0.2 K-space, which makes the two-component test unsuitable for the uncertainty process under consideration.

The two-component test need not use both Kt (global) and Kn (direct). If the two-component data set includes diffuse along with either global or direct, SERI QC converts the existing diffuse to Kt-Kn space using the closure equation (4) by solving for either Kt or Kn.

### 3.5 The Three-Component Test

If all three components are included in the incoming data set, both the one- and two-component tests are performed first. If each parameter passes the one-component test, the data triplet is evaluated by the two-component test, and the results are retained. Next, the data are evaluated by the three-component test, as described next.
The relationship between the three components of global, direct, and diffuse hold redundant information, which is revealed by the closure equation (4). Any of the three components can be calculated from the other two, thus allowing a comparison of a measurement to the calculated value from the other two measurements:

\[ Kn = Kt - Kd \]  
\[ Kd = Kt - Kn \]  

In the three-component test, SERI QC evaluates a measurement by subtracting its K-space value from the value calculated from the other two components. If the determined discrepancy is within a K-space value of 0.03 (an arbitrary number thought to represent the best performance of instruments of the day\(^8\)), the data pass the three-component test (flag 03). If the discrepancy is beyond that threshold, the magnitude of the discrepancy becomes encoded in the flag.

The three-component test also looks at the results of the preceding two-component test. If the data fail the two-component test (outside the boundaries) but pass the three-component test (satisfies the closure equation), a special flag is assigned (09) to alert the analyst that a contradiction has occurred that might be of special interest. An excessive number of these flags for the time period under consideration indicate incorrect Gompertz boundaries or that a measurement condition exists that causes a cancellation of errors. Flag 09 also occurs in high-temporal-resolution data under intense cloud enhancement conditions that augment diffuse irradiance.

The three-component test provides the most robust evaluation of data quality because the only ambiguity in the test comes from the measurements themselves. Unlike the one- and two-component tests, the three-component test does not rely on historical data to predict the realm of good data. Although the possibility exists that all three measurements could be in error, and in the worst case, errors could cancel, this possibility is minimized with strict adherence to best practices in the measurement protocol.

\(^8\) Advances in radiometry since the SERI QC development suggest this value be revised downward in future versions.
4 The Gompertz Boundaries

4.1 The Gompertz Function

The Gompertz function is commonly used to model exponential population growth, and among its various forms is this basic formula:

\[ f(x, a, b, c) = ab^{-e^{-cx}} \]  
(9)

(where \( e \) is Euler’s number.)

The user’s manual references and uses a generalized Gompertz function (Parton and Innis 1972), where Euler’s number is replaced by variables \( b \) and \( c \):

\[ f(x, a, b, c, d) = ab^{-cb-\alpha} \]  
(10)

The user’s manual uses this form for K space with \( Kn \) as a function of \( Kt \):

\[ Kn = AB^{CB^{dKt}} \]  
(11)

Note the omission of the negative signs in (11). This modification was not mentioned or explained in the user’s manual, but with proper parameter formulation, it is of no consequence.

The SERI QC software uses the Gompertz function only indirectly to establish a catalog of expected value boundaries rather than calculating boundaries at run time. The function includes an array definition of x-y coordinates defining curves that establish the boundaries of the acceptable data envelopes. These coordinates, which are defined in the software and grouped by right and left boundary curves, are included in Appendix D. Each group of 100 values defines a Gompertz function curve by \( Kt-Kn \) coordinates, where the \( Kt \) values are the numbers in the table, and the \( Kn \) values are implicit in the position within the table (i.e., the first value represents the \( Kt \) number at \( Kn=0 \), the next position at \( Kn=0.01 \), etc.) Note that the values in the table and the implicit position represent K-space coordinates multiplied by 100; thus, contrary to the Gompertz function (11), where \( Kn \) is a function of \( Kt \), in these tables, the \( Kt \) values are essentially a function of \( Kn \). This configuration is well hidden from the user and was likely chosen as the best operational method for actually evaluating the position of a K-space data pair by the software.

Several of the left curve definitions in Appendix D start with -60.0, and one of the right curve definitions ends with 999.9. These values are not documented in the software or user’s manual, but they likely indicate array positions that are not used.

The user’s manual makes no mention of how these tables were generated, although it is likely that the Gompertz function was evaluated for high-resolution values of \( Kt \) with resolution much greater than the 0.01 increments used by SERI QC. In this approach, higher resolution \( Kn \) values could have been matched or interpolated to the evenly spaced \( Kn \) increments implicit in the tables.
4.2 Determining the Gompertz Coefficients

The user’s manual does not document or reference the actual Gompertz coefficients used to generate the boundary curves. This knowledge could be useful during any effort to modify the SERI QC algorithm; thus, to determine the coefficients for this report, an iterative approach of trial and error was employed to recreate the values found in the SERI QC table definitions. For this purpose, the Gompertz (11) function from the user’s manual was used. Note that more than one curve that fit an original curve could be generated using wildly different coefficients, indicating that the input coefficients are somewhat interdependent.

The quality of fit was determined using the Finkelstein-Shafer (1971) statistic, which quantifies the average degree of agreement (in K-space units) between two curves. The operating resolution of SERI QC is 0.01 K-space increments, and the F-S scores on most generated curves are approximately 0.003 or less, which should suffice for this purpose. The inability to exactly reproduce the function output is likely a result of several factors, including: (1) the original curves are defined only to three positions past the decimal point, and (2) the values defined in the SERI QC code had to be interpolated to the realm of Kd as a function of Kt for this evaluation; thus, an exact reproduction is not expected.

Table 4-1 shows the coefficients used to reproduce the curves (L1, R1, etc., refer to Left Boundary 1, Right Boundary 1, etc.)

<table>
<thead>
<tr>
<th>Curve</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Offset</th>
<th>Target Inflection (X,Y)</th>
<th>Curve Inflection (X,Y)</th>
<th>F-S Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>1.2360</td>
<td>9.2810</td>
<td>-3.5725</td>
<td>-2.4740</td>
<td>0.00</td>
<td>(.37, .36)</td>
<td>(.37, .43)</td>
<td>0.0030</td>
</tr>
<tr>
<td>L2</td>
<td>1.2430</td>
<td>8.1120</td>
<td>-3.0581</td>
<td>-2.3390</td>
<td>0.00</td>
<td>(.37, .44)</td>
<td>(.37, .43)</td>
<td>0.0015</td>
</tr>
<tr>
<td>L3</td>
<td>1.3860</td>
<td>8.6820</td>
<td>-2.1280</td>
<td>-1.7040</td>
<td>0.00</td>
<td>(.42, .52)</td>
<td>(.41, .50)</td>
<td>0.0028</td>
</tr>
<tr>
<td>L4</td>
<td>1.4950</td>
<td>24.000</td>
<td>-1.1426</td>
<td>-0.8950</td>
<td>0.00</td>
<td>(.47, .57)</td>
<td>(.45, .54)</td>
<td>0.0014</td>
</tr>
<tr>
<td>L5</td>
<td>1.1520</td>
<td>4.9960</td>
<td>-2.4436</td>
<td>-2.2900</td>
<td>0.00</td>
<td>(.38, .44)</td>
<td>(.37, .41)</td>
<td>0.0007</td>
</tr>
<tr>
<td>L6</td>
<td>0.9450</td>
<td>4.9450</td>
<td>-2.9548</td>
<td>-2.9760</td>
<td>0.00</td>
<td>(.31, .33)</td>
<td>(.32, .33)</td>
<td>0.0015</td>
</tr>
<tr>
<td>R1</td>
<td>1.3130</td>
<td>25.000</td>
<td>-2.6721</td>
<td>-1.2970</td>
<td>0.01</td>
<td>(.53, .50)</td>
<td>(.51, .47)</td>
<td>0.0019</td>
</tr>
<tr>
<td>R2</td>
<td>1.1720</td>
<td>23.890</td>
<td>-3.8448</td>
<td>-1.6140</td>
<td>0.01</td>
<td>(.50, .45)</td>
<td>(.48, .41)</td>
<td>0.0019</td>
</tr>
<tr>
<td>R3</td>
<td>1.0510</td>
<td>7.2300</td>
<td>-10.307</td>
<td>-3.2470</td>
<td>0.01</td>
<td>(.48, .39)</td>
<td>(.46, .36)</td>
<td>0.0018</td>
</tr>
<tr>
<td>R4</td>
<td>1.0060</td>
<td>5.5650</td>
<td>-27.415</td>
<td>-4.8200</td>
<td>0.01</td>
<td>(.46, .35)</td>
<td>(.46, .35)</td>
<td>0.0015</td>
</tr>
<tr>
<td>R5</td>
<td>0.9840</td>
<td>4.0820</td>
<td>-100.16</td>
<td>-7.5690</td>
<td>0.01</td>
<td>(.46, .35)</td>
<td>(.46, .34)</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

The values in the Table 4-1 Offset column are used to impose a lower bias on the right curves for a realistic fit. Such an offset is not mentioned or documented in the user’s manual, but an explanation is offered here: As shown in Appendix D for the right curves—for example, the Right Curve 1 table—the Kt value for the first Kn value (implicitly Kn=0) is 13.4. This indicates that f(kt at 0.134) is Kn=0, which the Gompertz function is not capable of producing with
nonzero values for coefficients A and B. This indicates that during the SERI QC development, the curve was likely moved toward zero by some amount. This approach would create a clean intercept of the curve with the Kt axis rather than decaying toward zero. This might have been done to prevent Kn values at or near zero from falling outside (below) the right boundary. Using an offset produces an otherwise impossible fit to the original curves and is applied by subtracting it from the Gompertz function value for each instance of Kt:

\[ Kn = f(Kt) - \text{offset} \]

The user’s manual notes that coefficient B (11) is used to move the curve inflection point vertically along the y axis; however, the reference for (10) is somewhat less direct, stating that coefficient B is the “control parameter that changes value of f(x) where the inflection point is located \([f(x) \approx a/b \text{ at the inflection point for values of } b \text{ between 2 and 6}]\).” Using the iterative process, changing parameter B resulted in some movement of the inflection point along the y axis, but not always predictably, and usually with other unwanted changes in the curve. In Table 1, the inflection point of the existing curves (target inflection) were estimated and used to position the inflection point of the generated curves, to the extent possible.

Figures 4-1 and 4-2 show examples of the curve-fitting process for the Right Curve 1 and the Left Curve 4. The magenta line is the curve currently used by SERI QC, and the superimposed thin green line is the Gompertz curve generated by the function using the coefficients from Table 4-1.

![Figure 4-1. Gompertz plot in K-space](image)
4.3 Maximum Diffuse Values
SERI QC requires maximum values for evaluating one-component data. The Kt and Kd maximums generally decrease with increasing data integration time that limits the variability of irradiance caused by rapid changes in cloud conditions. The Kn maximum remains fairly independent of the data integration time interval because the maximum direct normal irradiance occurs under generally stable clear-sky conditions.

The Kt and Kn max values are determined empirically from maximums encountered in historical data. The Kd (diffuse) max is not directly determined from data because SERI QC operates only in two-dimensional Kt-Kn space; however, according to the user’s manual, a Kd max can be determined from the Kt-Kn pairs of the chosen right Gompertz boundary by subtracting the Kn values from the Kt values. Further, the maximum Kd value is then identified by evaluating where the Gompertz curve has a slope of 1. In Figure 4-3, each of the five right-hand Gompertz curves were plotted as Kd = Kt - Kn, and the maximum Kd (square marker) corresponds exactly at the point where the slope of the curve equals 1. This can be explained intuitively given that Kt is the sum of Kn and Kd, and Kd initially dominates. As Kn increases from zero, it contributes an increasing proportion to Kt, where at some point it equals Kd (slope of 1), after which Kd begins to decrease, and Kn dominates.
SERI QC does not do a run-time evaluation of the right-hand curves in this manner, but rather the maximum Kd values are explicitly defined by this array in the code:

\[
\text{double XDm[5]} = \{ 0.19, 0.22, 0.24, 0.28, 0.32 \}
\]

These points are shown in Figure 4-3 with circles, which are clearly not indicative of the Kd max described in the user’s manual (square markers). There is no documentation on this discrepancy, but one could speculate that those values were possibly determined during development from other curves and were not updated, or possibly they were reduced for some other reason. Although those points specify a Kd max that is too low, they are modified upward during run time. The net result is an increase of 0.1 in the maximum Kd (triangle markers in Figure 4-3) for each curve. This rationale is described in the user’s manual (Section 6.4.1), where it is stated that both Kt max and Kd max are increased by 0.1 to account for brief cloud reflections that increase diffuse (and, consequently, global).

![Kd (Diffuse) Maximum](image)

**Figure 1-3. Plot of diffuse maximum values**

Comparing different methods of deriving Kd max from the right-hand Gompertz boundaries

Any modifications to SERI QC that include additional right-hand curves should carefully consider this discussion and the effect on Kd max.
5 QCFIT

A stand-alone companion program to SERI QC was created to facilitate the creation of the Gompertz boundaries. Because the one- and two-component tests require historical data to establish areas of expected values for quality assessment, the QCFIT program was developed to visualize and speed the selection of boundaries from historical data.

5.1 The Original QCFIT

The QCFIT program ingests multiple years of data for a site and presents the user with a visual display of the Kt-Kn scatterplot. Each data set is examined by month of year and by air mass to fine-tune the fitting of data.

The program will determine boundaries on the right and left of the plot as well as the Kt and Kn max limits. The user has the opportunity to adjust the curves based on analysis of the scatter, the boundary lines, and expert knowledge of including or excluding data points to balance Type 1 and Type 2 errors. The selected curves and maximums, along with geographic location, are written to a station-specific QC-Zero file (Appendix C). This file is accessed during a call to SERI QC to provide boundary information required for the one- and two-component tests.

5.2 An Enhanced QCFIT

During the examination of data from many solar measurement networks, it became apparent that the visualization provided by QCFIT could be used as a valuable data analytic tool. The program was enhanced by:

- Improving the interface in the Windows environment
- Adding color to illuminate the scatterplot density
- Adding the ability to view any pair of the three-component data on the scatterplot
- Animating the three combinations of three-component pairs
- Data filtering to better fit the boundaries
- Isolating and viewing individual years among aggregate
- One-click boundary changes
- Selection of scatterplot points to view in the context of the original data file
- Display of mean error of each air mass.

Figure 5-1 shows a sample interface screen from the enhanced QCFIT.
Figure 2-1. QCFIT user interface

In 2019, QCFIT was ported to the Python programming language.
6 Historical Uses of SERI QC

SERI QC and the companion program QCFIT have been used extensively to assess the data quality of numerous solar monitoring networks, including:

- Atmospheric Radiation Measurement (ARM) Solar Infrared Radiation Station (SIRS), Sky Radiation (SKYRAD), and Ground Radiation (GNDRAD) networks (Habte et al. 2014)
- Data Quality Management System (DQMS) (Maxwell et al. 1999)
- Historically Black Colleges and Universities Solar Measurements Network (Marion 1994)
- King Abdullah City for Atomic and Renewable Energy (K.A.CARE) (Zell et al. 2015)
- King Abdulaziz City for Science and Technology (KACST) (Al-Abbadi et al. 2002; Myers et al. 2002)
- National Oceanic and Atmospheric Administration (NOAA)/Integrated Station Information System (ISIS) and Surface Radiation Budget (SURFRAD) networks (Anderberg and Sengupta 2014)
- NREL Solar Radiation Research Laboratory, Baseline Measurement System (Sengupta et al. 2022)
- Solar Resource and Meteorological Assessment Project (SOLRMAP) (Wilcox and Myers 2009)

As part of data acquisition programs, SERI QC can be used daily as a first look at the data to detect error conditions. Although thorough station on-site inspections and regular maintenance are the fundamental steps in data quality control, many errors can be detected only via examination of the measured data. SERI QC provides a fast and automated means for inspecting large volumes of solar irradiance data, and the available summary reports of the quality assessment flags will alert station operators of conditions that require prompt attention.

Other tools that ingest and analyze the SERI QC flags provide immediate feedback to operators. One such tool is a three-dimensional plot that summarizes a period of data to show flag magnitude and the corresponding solar measurements in K-space for the three components. Figure 6-1 is an example of such a plot displaying monthly records of 1-minute data and the associated data quality flags at a glance, allowing an analyst to quickly find areas of concern.
**Figure 3-1. XYZ plot of quality flags and K-space measurements**

A monthly data set for the NREL Baseline Measurement Station

7 Pros and Cons of Continued Use of SERI QC

7.1 The Pros

- **Easily configured**—SERI QC can fully characterize expected values for a site through the use of historical data to use the two-component test. Further, for the three-component test, historical data are not critical (although desirable), making it possible to begin testing data at a new station with no prior measurement history.

- **Adaptable to existing processing software**—Assuming that a network has existing software for data acquisition, it is possible to insert the SERI QC function to adapt the processing to data quality tests.

- **Climate specific**—SERI QC accounts for local climate and typical weather conditions when configured with historical data. Such configurations could be adaptable with minor adjustments to nearby sites or other areas with a similar climate.

7.2 The Cons

- **Requires programming expertise**—Users must have access to programming capabilities to incorporate the function into existing software. Commercial data acquisition software without source code would not be adaptable, although an executable SERI QC package could be written to capture a data stream or process archived data.

- **Cannot identify component in error**—With the two- or three-component test, SERI QC can only indicate that one or more of the components are causing a discrepancy, but it cannot identify the component(s) in error.

- **Existing curves do not always work for high-resolution data**—The current catalog of Gompertz boundaries was developed primarily using hourly data. Characteristics of 1-minute data mean that many points can fall outside the left boundary as a result of the short-term effects of rapidly changing atmospheric conditions.

- **Loose expected values**—The one- and two-component tests have a wide range of expected values. Although this is valuable for spotting error conditions in a large data set, it cannot assign a true error value to an individual data point. This is not true for the three-component test, which has very tight bounds.
8 Suitability of SERI QC for Uncertainty Estimates

8.1 Description of the Flags
Measurement uncertainty analysis is a numerical, objective method for defining the potential that error exists in all data (Dieck 1992). Uncertainty estimates require a reference standard, or true value, by which a measurement can be compared to estimate the magnitude of the error. The SERI QC flags provide a continuum of well-defined error representations, but the discrepancies that occur are referenced to extraterrestrial solar irradiance values rather than a valid absolute reference measurement of the solar irradiance in units of flux density. The extraterrestrial values are known and accepted to within a fraction of a percentage, and being mathematically defined across K-space, they provide a consistent comparison to judge the quality of data, at least for the three-component test. Flags (10-93) described in Appendix A show how a discrepancy greater than 0.03 K-space is encoded to provide an unambiguous measure of its magnitude and direction. Discrepancies greater than or equal to 0.15 are encoded with flags 90–93, because errors outside that limit are considered beyond the need to be quantified. The flags, as noted in Appendix A, can be decoded, which will yield the K-space discrepancy of the flagged data point.

Note that without the ability to isolate which instrument is causing the discrepancy, all three components receive flags of the same magnitude. Although circumstances at the measurement station can provide the means for an intelligent conjecture of fault, those means are beyond the scope of SERI QC.

8.2 Determining a Standard for Operational Uncertainty
In this document, we use the term operational uncertainty to describe errors estimated from an examination of field measurements without regard to more fundamental instrument uncertainties. In other words, this approach looks at characteristics of the data after measurements are completed and attempts to ascertain errors apparent in an operational environment. Such estimates would include fundamental instrument uncertainties, but they are not immediately identified.

Using the closure equation (6) and its rearrangements (7), (8) that produce the flags, we can derive a measurement standard in K-space for any of the three global, direct, or diffuse measurements. Here, then, are the proposed standards for determining the operational uncertainty (U_o) in percentage for any measurement in a three-component triad:

\[ U_o K_t = \left( \frac{K_t}{K_n + K_d} - 1 \right) \cdot 100 \] (12)
\[ U_o K_n = \left( \frac{K_n}{K_t - K_d} - 1 \right) \cdot 100 \] (13)
\[ U_o K_d = \left( \frac{K_d}{K_t - K_n} - 1 \right) \cdot 100 \] (14)

---

9 There has never been and never will be a case when a person measured a variable and obtained the true value (Dieck 1992, 9).
10 The presently accepted value of the TSI is 1361.08 ±0.5 W/m² (Kopp and Lean 2011).
This approach does not hold for one- and two-component data sets, which, as stated in Section 3.4, have a large range of expected values for any particular measurement; thus, any attempt to derive a two-component reference measurement would result in an unacceptably large uncertainty for the reference itself.

8.3 Assumptions and Concerns

8.3.1 Computational Limitations

The application of (13) under overcast skies (when the direct beam is zero) will be undefined when \( K_t = K_d \). Further, when the difference between \( K_t \) and \( K_d \) 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 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 purpose of solar irradiance measurements is generally to support power generation projects. Because the direct beam (as either the single component or a 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.

Figure 8-1 is taken from page 27 of the user’s manual and gives some insight into this issue. We expect the most troublesome data to fall in the cloudy region close to the x axis (8–10 tenths cloud cover). The labeled regions in the figure are summarized here:

- Region A—short-lived phenomenon, such as cloud lensing and cloud edge illumination, captured during short integration times
- Region B—similar to Region A, short-lived events, possibly from high-intensity reflections from clouds near but not occluding the sun
- Region C—steady-state data from all integration times
- Region D—data from medium to cloudy conditions, more prevalent in shorter integration times
- Region E—data resulting from cloud cover in high-albedo conditions.

Another consideration is the decoupling of three-component measurements under rapidly changing irradiance conditions, an occurrence frequently encountered during passing clouds. Because the ability to correctly track highly irregular irradiance can vary among instruments, large transient errors can occur if instruments respond to changes at a different rate. Although this is a known and accepted instrument behavior, it nonetheless results in measurement error, and the operational uncertainty should capture its magnitude.
8.3.2 The Ambiguity of Fault

The approach proposed in Section 8.2 is not ironclad. As mentioned in Section 3.5 (three-component test), there is no assurance that only one of the measurements causes an anomalous error and the subsequent failure of a test. Further, we usually do not have information at hand to determine which instrument is at fault, particularly in an automated process.

As with any reference measurement, there is no doubt at some level that measurement error exists, but the process to determine uncertainty attempts to understand, quantify, document, and minimize the source(s) of the error.

In field measurements, we encounter many sources of instrument and environmental error that have the potential to adversely affect a measurement, but with robust and strict protocols for station operations, those errors can be controlled and minimized; however, this still does not necessarily provide sure guidance toward identifying the faulty instrument(s). Because of this described ambiguity of fault, as with SERI QC, this method will assign similar uncertainties to each of the three components even though it is very likely that not all three are at fault.
This ambiguity of fault and the necessary painting of uncertainty with a broad brush can cause concern among users. If a measurement campaign focuses on one component—for example, the direct beam—and the inclusion of the other two holds the potential to cast doubt on the primary measurement, there can arise skepticism or even opposition; however, there are many circumstances that could exonerate the primary measurement if additional maintenance or operational information points to a malfunction that affects one of the other instruments. This information is often readily available at a well-run measurement station, and its existence will add to the defense of data quality rather than distract from it.

8.4 Application of the Operational Uncertainty

The uncertainty estimated by this process will add to the fundamental instrument and measurement uncertainty determined by the NREL Solar Resource Uncertainty Application. The application demonstrates that instruments themselves have a base uncertainty determined by knowledge of the measurement technology and the calibration method (which itself carries forward an uncertainty from the reference instruments used for the calibrations).

Instruments are issued a calibration uncertainty that reflects instrument performance under carefully controlled conditions, and measurements performed under calibration conditions may reasonably take on the calibration uncertainty; however, field instruments are subjected to environmental and operational factors that introduce many instrument response characteristics that must add to the uncertainty of a measurement. The operational uncertainty described here attempts to capture some of that additional uncertainty.

There is no doubt that the operational uncertainty apparent in field measurements includes the fundamental calibration uncertainty from the NREL uncertainty application, and, therefore, the operational uncertainty is expected to be nonzero even under calibration conditions. For this reason, some account should be made to not double count the fundamental uncertainty when considering the operational uncertainty. This issue will be investigated in Task 4.2.

8.5 A Foundation of Best Practices

In the context of determining the measurement uncertainty, the importance of the best practices information summarized in Section 1.1.2 and expressed in much greater detail by Sengupta (2021) cannot be overstated. Best practices generally speak for themselves as a positive endeavor, but, as revealed in this discussion of determining uncertainty, the best practices themselves are critical to providing a basis for directly evaluating the data.
9 Modifications to SERI QC for Deriving Uncertainty

SERI QC as it currently operates is not sufficient to meet the goals of this project. Some modifications will be required, and a decision must be made to determine how to best accomplish the goal of translating SERI QC flags to uncertainty.

9.1 Required Modifications to SERI QC

The flags issued by SERI QC represent either a condition where data have conformed to an expected value (passed a quality check) or the magnitude of a discrepancy between the measurement and the expected value. In the three-component test, however, SERI QC has a threshold of 0.03 K-space before agreement is rejected. In the context of adding operational uncertainty to the fundamental instrument uncertainty, this step at the threshold is too coarse.

To correct this design characteristic, two approaches should be considered: modify the SERI QC flag scheme or calculate the operational measurement uncertainties within SERI QC.

9.1.1 Modify the SERI QC Flag Scheme

The values in the operational uncertainty reference equations (12), (13), and (14) are embedded in SERI QC flags 10–93 but only if the discrepancy is greater than 0.03 K-space. The values less than or equal to 0.03 are assigned a “good” flag of 03, yet they might be very relevant to the fine-tuning of uncertainty because this value represents a realm close to the base uncertainty itself.

The solution to this would be to modify the SERI QC flag scheme by assigning flag 03 to measurements less than 0.01 K-space and starting flag 10 at a threshold of 0.01 K-space. This would allow an external program to decode the flags and extract the $K_t$, $K_n$, and $K_d$ values. Calculating the actual operational uncertainty would require calculating the extraterrestrial solar values for the station location, date, and time of the measurement. Because SERI QC uses these values in its evaluation, they could be included as additional passed parameters by reference back to the calling program, which would perform the calculations.

Coding changes would be minimal to slide the three-component test flag threshold from 0.03 to 0.0 and include additional variables in the parameter list. Because the original editable version of the SERI QC user’s manual no longer exists, there would be considerable revision work in multiple sections to describe the new scheme.

9.1.2 Calculate the Operational Measurement Uncertainties within SERI QC

SERI QC has all the information to satisfy the needs of equations (12), (13), and (14) and could perform the operational uncertainty calculations internally. This would require that three additional parameters holding the uncertainty values be passed by reference back to the calling program. The coding change would be minimal to perform the calculations and add three more variables to the function parameter list.

This approach would hold intact the existing SERI QC flag scheme for backward compatibility to accommodate the needs of other analyses that examine the flags. Additional documentation would be required, but it could be in the form of an addendum to the existing user’s manual.
9.2 Additional Changes to SERI QC to Consider

Since SERI QC was developed, the preponderance of measured solar data has shifted from hourly integration to 1-minute. The Gompertz curves developed at the time fit the characteristics of the hourly data quite well but fail to consider some of the short-lived, high-resolution measurements. An additional one or two curves on both the left and right could be considered, assuming that the Gompertz function is capable of producing suitable curves.

This is a low-priority item in the context of this project because, as described in Section 3.5, the three-component test has little reliance on the Gompertz curves.
10 Conclusions and Recommendations

10.1 Suitability of SERI QC for Supporting Operational Uncertainty

SERI QC has a long history in measurement network operations. Using either the SERI QC flags themselves or the same information that creates the flags, the software can easily and effectively be used to refine the stated uncertainty of a measurement by supplementing the fundamental instrument uncertainty.

The approach described in Section 8.2 can be effectively implemented without the use of SERI QC if the necessary supporting K-space information is available; however, SERI QC not only provides all supporting information, but its data quality capabilities are a necessary step in filtering poor data from the uncertainty process. SERI QC also provides robust error checking and input validation, of which much would need to be duplicated in an external process.

10.2 Recommendations for Implementation

Two options were discussed in Section 9 for implementing an operational uncertainty calculation. The second option has the advantage of simple modifications and minimum documentation changes. Any program calling SERI QC would require a minor change to the function parameter list to accommodate the returned uncertainty values for each measurement; however, for existing programs, these parameters could be ignored, and the original program would operate the same.

Including the calculations within SERI QC would provide a more transparent method of obtaining operational uncertainty, providing a standard method for the process.

This method heavily relies on the implementation of best practices. Any data collection effort used for major financial decisions should incorporate best practices from its inception in the design phase and carry the approach through to routine station operations. Any measurement campaign that does not include an emphasis on best practices cannot fully benefit from the uncertainty determination methods described in this report.
References


### Appendix A. SERI QC Flagging Convention

<table>
<thead>
<tr>
<th>Flag</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Untested (raw data)</td>
</tr>
<tr>
<td>01</td>
<td>Passed one-component test; data fall within max-min limits of Kt, Kn, or Kd</td>
</tr>
<tr>
<td>02</td>
<td>Passed two-component test; data fall within 0.03 of the Gompertz boundaries</td>
</tr>
<tr>
<td>03</td>
<td>Passed three-component test; data come within ±0.03 of satisfying ( Kt = Kn + Kd )</td>
</tr>
<tr>
<td>04</td>
<td>Passed visual inspection; <em>not used</em> by SERI_QCI</td>
</tr>
<tr>
<td>05</td>
<td>Failed visual inspection; <em>not used</em> by SERI_QCI</td>
</tr>
<tr>
<td>06</td>
<td>Value estimated; passes all pertinent SERI_QCI tests</td>
</tr>
<tr>
<td>07</td>
<td>Failed one-component test; lower than allowed minimum</td>
</tr>
<tr>
<td>08</td>
<td>Failed one-component test; higher than allowed maximum</td>
</tr>
<tr>
<td>09</td>
<td>Passed three-component test but failed two-component test by &gt;0.05</td>
</tr>
<tr>
<td>10-93</td>
<td>Failed two- or three-component tests in one of four ways. To determine the test failed and the manner of failure (high or low), examine the remainder of the calculation (flag + 2)/4.</td>
</tr>
</tbody>
</table>

#### Failure Table

<table>
<thead>
<tr>
<th>Rem</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Parameter too low by three-component test (( Kt = Kn + Kd ))</td>
</tr>
<tr>
<td>1</td>
<td>Parameter too high by three-component test (( Kt = Kn + Kd ))</td>
</tr>
<tr>
<td>2</td>
<td>Parameter too low by two-component test (Gompertz boundary)</td>
</tr>
<tr>
<td>3</td>
<td>Parameter too high by two-component test (Gompertz boundary)</td>
</tr>
</tbody>
</table>

The magnitude of the test failure (distance in K-units) is determined from:

\[
d = \frac{(\text{INT} (\text{flag} + 2)/4) \times 100}{100}
\]

Examples and further discussion of the meaning of flags 10-93 are given in the text in this chapter and in Section 9.3, page 153.

| 94-97 | Data fall into a physically impossible region where \( Kn > Kt \) by K-space distances of 0.05 to 0.10 (94), 0.10 to 0.15 (95), 0.15 to 0.20 (96), and \( \geq 0.20 \) (97). |
| 98   | Not used |
| 99   | Missing data |

---

Figure AA-1. SERI QC flagging convention

From Maxwell et al. (1993)
Appendix B. SERI QC Flowchart

Figure AB-1. SERI QC flowchart
From Maxwell et al. (1993)
Appendix C. QC0 File

Site Identifier: SR2, Solar Radiation Research Laboratory
--- Latitude: 39.7400
-- Longitude: -105.1800
-- Time Zone: -7.0000

---Max Code--- ---Low Airmass--- --Medium Airmass--- ---High Airmass---
KN ----KT----- Left ----Right---- Left ----Right---- Left ----Right----
1 5 15 60 S P S 1 5 15 60 S P S 1 5 15 60 S P S 1 5 15 60

JAN: 87-99/00/00/00; 0-00 0-00/00/00/00; 4-08 2-17/00/00/00; 4-08 1-15/00/00/00
FEB: 87-99/00/00/00; 0-00 0-00/00/00/00; 5-08 1-17/00/00/00; 5-08 1-15/00/00/00
MAR: 85-99/00/00/00; 4-10 1-15/00/00/00; 5-08 3-17/00/00/00; 5-08 1-15/00/00/00
APR: 83-99/00/00/00; 4-10 1-18/00/00/00; 5-09 1-18/00/00/00; 5-08 1-15/00/00/00
MAY: 84-99/00/00/00; 4-10 1-17/00/00/00; 4-09 1-17/00/00/00; 4-08 1-16/00/00/00
JUN: 82-99/00/00/00; 4-10 1-16/00/00/00; 4-09 1-16/00/00/00; 4-08 1-14/00/00/00
JUL: 82-99/00/00/00; 4-10 1-16/00/00/00; 4-09 1-16/00/00/00; 4-08 1-13/00/00/00
AUG: 81-99/00/00/00; 4-10 1-16/00/00/00; 4-09 1-16/00/00/00; 4-08 1-14/00/00/00
SEP: 85-99/00/00/00; 4-10 1-15/00/00/00; 6-07 1-16/00/00/00; 5-07 1-12/00/00/00
OCT: 86-99/00/00/00; 0-00 0-00/00/00/00; 5-08 1-16/00/00/00; 5-06 1-14/00/00/00
NOV: 89-99/00/00/00; 0-00 0-00/00/00/00; 5-08 1-16/00/00/00; 5-07 1-13/00/00/00
DEC: 87-99/00/00/00; 0-00 0-00/00/00/00; 5-08 1-16/00/00/00; 5-07 1-14/00/00/00

DEFAULT CONFIGURATION
Integration (minutes): 1
Data Folder: d:\NSRDB2\Validation Data\SRRL
Plane: 1 Kt-Kd
3-Component Filter: -1.000000
Appendix D. Gompertz Curve Definitions

float curveLeft[5][100] = {
    {72.0, 74.0, 76.0, 78.0, 80.0, 82.0, 84.0, 86.0, 88.0, 90.0},
    {57.0, 59.0, 61.0, 63.0, 65.0, 67.0, 69.0, 71.0, 73.0, 75.0},
    {42.0, 44.0, 46.0, 48.0, 50.0, 52.0, 54.0, 56.0, 58.0, 60.0},
    {37.0, 39.0, 41.0, 43.0, 45.0, 47.0, 49.0, 51.0, 53.0, 55.0},
    {22.0, 24.0, 26.0, 28.0, 30.0, 32.0, 34.0, 36.0, 38.0, 40.0},
};

float curveRight[5][100] = {
    {70.0, 72.0, 74.0, 76.0, 78.0, 80.0, 82.0, 84.0, 86.0, 88.0},
    {62.0, 64.0, 66.0, 68.0, 70.0, 72.0, 74.0, 76.0, 78.0, 80.0},
    {47.0, 49.0, 51.0, 53.0, 55.0, 57.0, 59.0, 61.0, 63.0, 65.0},
    {32.0, 34.0, 36.0, 38.0, 40.0, 42.0, 44.0, 46.0, 48.0, 50.0},
    {17.0, 19.0, 21.0, 23.0, 25.0, 27.0, 29.0, 31.0, 33.0, 35.0},
};