NEON ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD) – WEIGHING GAUGE PRECIPITATION WITH BELFORT AEPG II 600M SENSOR

PREPARED BY	ORGANIZATION	DATE
Cove Sturtevant	Science	12/23/2024
Teresa Burlingame	Science	12/23/2024

APPROVALS	ORGANIZATION	APPROVAL DATE
Kate Thibault	SCI	07/10/2025

RELEASED BY	ORGANIZATION	RELEASE DATE
Frank Mocilac	CM	07/11/2025

See configuration management system for approval history.

The National Ecological Observatory Network is a project solely funded by the National Science Foundation and managed under cooperative agreement by Battelle. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Change Record

REVISION	DATE	ECO #	DESCRIPTION OF CHANGE
A	09/13/2017	ECO-04989	Initial Release
B	04/20/2022	ECO-06809	• Update to reflect change in terminology from relocatable to gradient sites • Updated logo
C	02/08/2024	ECO-07058	• Algorithm edits to improve product usability given extreme sensor noise. The algorithm no longer removes negative values from sums and adds wireNoiseQF to impact the finalQF
D	12/23/2024	ECO-07128	• Significant rework of precipitation algorithm using smoothing method to reduce impact of sensor noise • Changing data delivery to hourly and daily to account for reduced temporal accuracy with smoothing algorithm • Refactor of uncertainty calculations to account for new methodology

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

NEON uses two methods to determine bulk precipitation across sites. Bulk precipitation measurements at core sites are made with a weighing gauge sensor surrounded by a double fence inter-comparison reference (DFIR), while bulk precipitation measurements at gradient sites are collected using a tipping bucket sensor. Bulk precipitation with a DFIR, alter shield, and a weighing gauge is known to provide improved results over standalone tipping bucket measurements. This document provides the details for computing weighing gauge precipitation, which consists of the algorithms used to create a NEON Level 1 data product (DP) from Level 0 (raw) data. Additionally, ancillary data/inputs such as calibration data are defined in this document. Domains 1, 5, 9, 10, 12, 13, 17, 18, and 19 use the heated version (Product Number (P/N): CG07180010 and 0303440002), while all other domains use the non-heated version (DGD P/N: CG07180000 and 0303440001). The purpose of this document is to explain the measurement theory and implementation, in addition to the appropriate theoretical background, data product provenance, quality assurance and control methods, approximations and/or assumptions, and an exposition of uncertainty resulting in a cumulative reported uncertainty for this product is provided.

1.2 Scope

The scope of this document is the theoretical background and algorithmic process used to derive Level 1 data from Level 0 data for weighing gauge precipitation. The AEPG II 600M weighing gauge is the selected measurement system to measure precipitation at all core tower sites and select aquatic sites. In Award Year (AY) 2025, the NEON program is replacing the Belfort AEPG II 600 M with the Pluvio2 L 200 sensor. Due to significant disparities in the algorithms required to process the Belfort and Pluvio sensor data, a separate ATBD will be created for bulk precipitation collected from the Pluvio sensor. This document does not provide computational implementation details, except for cases where they stem directly from algorithmic choices explained herein.

2 RELATED DOCUMENTS AND ACRONYMS

2.1 Applicable Documents

AD[01]	NEON.DOC.000001	NEON Observatory Design
AD[02]	NEON.DOC. 002652	NEON Data Products Catalog
AD[05]	NEON.DOC.000782	ATBD QA/QC Data Consistency
AD[06]	NEON.DOC.011081	ATBD QA/QC plausibility tests
AD[07]	NEON.DOC.000783	ATBD QA/QC Time Series Signal Despiking for TIS Level 1 Data Products
AD[08]	NEON.DOC.000897	C^3 Primary Precipitation Gauge
AD[09]	NEON.DOC.000898	ATBD Primary Precipitation Gauge
AD[10]	NEON.DOC.000367	C^3 Secondary Precipitation Gauge
AD[11]	NEON.DOC. 003289	Primary Precipitation Sensor Field Calibration
AD[12]	NEON.DOC.000927	NEON Calibration and Sensor Uncertainty Values
AD[13]	NEON.DOC.000785	TIS Level 1 Data Products Uncertainty Budget Estimation Plan
AD[14]	NEON.DOC.000746	Evaluating Uncertainty (CVAL)
AD[15]	NEON.DOC.001113	Quality Flags and Quality Metrics for TIS Data Products
AD[16]	NEON.DOC.001213	Primary Precipitation Calibration Fixture Manual

2.2 Reference Documents

RD[01]	NEON.DOC.000008	NEON Acronym List
RD[02]	NEON.DOC.000243	NEON Glossary of Terms

2.3 Acronyms

Acronym	Explanation
ATBD	Algorithm Theoretical Basis Document
CVAL	NEON Calibration, Validation, and Audit Laboratory
DFIR	Double Fence Intercomparison Reference
DGD	Data generating device
DP	Data Product
L0	Level 0
L1	Level 1

2.4 Variable Nomenclature

The symbols used to display the various inputs in the Algorithm Theoretical Basis Document (ATBD), e.g., calibration coefficients and uncertainty estimates, were chosen so that the equations can be easily interpreted by the reader. However, the symbols provided will not always reflect NEON's internal notation or the notation that is used to present variables on NEON's data portal. Therefore, a lookup table is provided to distinguish what symbols specific variables can be tied to in the following document.

Table 1. Variable Nomenclature

Symbol	Internal/Portal Notation	Description
A	CVALA1	CVAL Strain gauge calibration coefficient
B	CVALA2	CVAL Strain gauge calibration coefficient
P_0	CVALP0	CVAL Strain gauge calibration coefficient for an empty collector
f_0	CVALF0	CVAL Strain gauge calibration coefficient for a collector with oil/antifreeze added, but otherwise empty
u_{A1}	U_CVALA1	Combined, relative calibration uncertainty of a strain gauge reading (%)
$V_{eff\,A1}$	U_CVALD1	Effective degrees of freedom relating to u_{A1} (unitless)

3 DATA PRODUCT DESCRIPTION

3.1 Variables Reported

The primary precipitation related L1 DPs provided by the algorithms documented in this ATBD are displayed in the publication workbook stored in the CI data store.

3.2 Input Dependencies

Table 2. List of weighing gauge precipitation related L0 DPs that are transformed into L1 DPs in this ATBD.

Data product	Sample Frequency	Units	Data Product ID
Strain Gauge 1 Frequency (f_1)	0.1 Hz	Hz	NEON.DOM.SITE.DP0.00006.001.01900.HOR.VER.000
Strain Gauge 2 Frequency (f_2)	0.1 Hz	Hz	NEON.DOM.SITE.DP0.00006.001.02072.HOR.VER.000
Strain Gauge 3 Frequency (f_3)	0.1 Hz	Hz	NEON.DOM.SITE.DP0.00006.001.02073.HOR.VER.000
Strain Gauge 1 Stability (S_1)	0.1 Hz	NA	NEON.DOM.SITE.DP0.00006.001.01897.HOR.VER.000
Strain Gauge 2 Stability (S_2)	0.1 Hz	NA	NEON.DOM.SITE.DP0.00006.001.02068.HOR.VER.000
Strain Gauge 3 Stability (S_3)	0.1 Hz	NA	NEON.DOM.SITE.DP0.00006.001.02069.HOR.VER.000
Inlet Temperature*	0.1 Hz	°C	NEON.DOM.SITE.DP0.00006.001.01905.HOR.VER.000
Internal Temperature*	0.1 Hz	°C	NEON.DOM.SITE.DP0.00006.001.01906.HOR.VER.000
Heater Flag (i.e., orificeHeaterFlag)*	0.1 Hz	NA	NEON.DOM.SITE.DP0.00006.001.02000.HOR.VER.000

Note: * Signifies that these data products pertain to heated models only.

3.3 Product Instances

One instance of the Precipitation - weighing gauge L1 product is produced at each core terrestrial site and at select aquatic sites.

3.4 Temporal Resolution and Extent

The L0 data for weighing gauge precipitation are recorded at 0.1 Hz by three strain gauges, which are used to determine hourly and daily bulk precipitation for the L1 DP. These data or similar measurements are intended to be collected for the lifetime of the NEON project.

3.5 Spatial Resolution and Extent

The distance of the precipitation gauge from the tower (terrestrial sites) or meteorological station (aquatic sites) depends on the local terrain and is therefore site specific. The opening of the precipitation gauge is 200 mm^2. Thus, at a minimum, the spatial resolution of the gauge reflects a surface area of 200 mm^2 at the point in space where the precipitation gauge is located. However, the precipitation measurements are likely to be representative of a larger area surrounding the sensor, depending on local topography and atmospheric flow patterns.

4 SCIENTIFIC CONTEXT

Precipitation records are fundamental to understanding water availability in an ecosystem. As such, precipitation data is often used as ancillary data for a broad array of investigations. For example, evapotranspiration and latent heat fluxes can be better understood with knowledge of the amount and timing of precipitation and climate conditions can be better modeled.

4.1 Theory of Measurement

The measurement of precipitation is relatively straight forward; however, it can easily become biased by wind. Wind generally leads to the undercatch of precipitation and is the main factor that induces uncertainty in the measurement. The presence of solid precipitation compounds this problem, and windy conditions can result in 20-50% undercatch (Rasmussen, R. et al., 2012). Therefore, to reduce uncertainty in the measurement, NEON follows the site selection guidelines of the U.S. Climate Reference Network (USCRN) for the installation of precipitation gauges (CRN, 2002). Additionally, NEON incorporates the small DFIR configuration to minimize the effects of wind on the measurement. Precipitation itself is determined via a weighing gauge. The weighing gauge, with a known surface area, monitors the change in weight of the collector over time, which is directly equated to an accumulation in precipitation.

The weighing gauge is housed within a polyethylene resin shell that serves to protect the sensor components as well as reduce wind effects. In climates where freezing temperatures are expected, heaters are installed in the inlet of the sensor. The heaters serve two main purposes. Heaters reduce the potential of the gauge becoming encased in ice and melt solid precipitation to provide precipitation estimates when solid precipitation is present.

The precipitation measurement consists of three strain gauges that monitor the weight of the collector. A strain gauge is a metal wire that has known resonation characteristics. When a current is applied to the strain gauge it causes the wire to resonate. The resonant frequency is proportional to the square of the tension in the wire (Bakkehøi, S. et al., 1985). A range of calibration weights are used to develop a relationship between strain gauge frequency and weight for the gauge. This in turn allows the frequency output from the strain gauges to be used to calculate a corresponding depth measurement.

4.2 Theory of Algorithm (Overview)

For each 0.1 Hz observation record, each of the three strain gauge frequencies is converted into a collector depth by applying their respective calibration functions. Quality control is applied to remove unstable or implausible depth measurements. For records in which all three collector depths pass quality control, they are averaged to form a single collector depth for the record. Calibrated and averaged collector depth is further averaged over time intervals ranging from 5 to 60 minutes depending on the noise level of the sensor, which is stored in the processing parameters for each site. A smoothing algorithm is then applied to eliminate spurious variation in strain gauge frequency (and thus the calibrated collector depth) induced by environmental variation in temperature, pressure, and wind. The smoothing algorithm was adapted from a tool provided by Dr. Ralph Wright (Barnes and Hopkinson, 2022). Positive increases in the smoothed depth measurements are then summed into hourly and daily precipitation totals. Further details of the algorithm are described in the following Sections.

4.3 Removing Unstable Strain Gauge Frequencies

The Belfort AEPG II 600M sensor reports stability information for each gauge (S_1, S_2, and S_3) indicating whether the depth measurement is valid. A value of “P” indicates the gauge has stabilized, “S” indicates it is searching for stability, and “F” indicates the gauge or associated temperature thermistor has failed (see AD[08] for more details). Stability information is converted on site to integer format where P = 1, S = 0, and F = -1. Frequencies that are searching for stability or unstable (i.e., S = 0 or -1) shall be set to NULL:

$$f_{k,i} = \boldsymbol{NULL}\ when\ \mathrm{S}_{k,i} = 0 \text{ or } -1 \tag{1}$$

f = is a 0.1-Hz frequency measurement (Hz)

S = Strain gauge stability

k = 1, 2, or 3 (i.e., the number of strain gauges in the precipitation sensor)

i = Running index

4.4 Converting Frequency to Depth

Stable frequency measurements are converted to precipitation depth via the calibration function developed by the NEON program's Calibration, Validation and Audit Laboratory (CVAL); see AD[11] and AD[16] for information on the calibration procedure)

$$D_{k,i} = \left(A_k\left(f_{k,i} - P_{0_k}\right) + B_k\left(f_{k,i} - P_{0_k}\right)^2\right) * 10$$

$$D_0 = \left(A_k\left(f_{0_{k,i}} - P_{0_k}\right) + B_k\left(f_{0_{k,i}} - P_{0_k}\right)^2\right) * 10 \tag{2}$$

$$D_{precip,k,i} = D_{k,i} - D_{0,i}$$

Where:

D_k = Collector depth measured by a given strain gauge, k (mm)
D_0 = Collector depth at zero calibration for a given strain gauge, k (with oil/antifreeze solution added) (mm)
D_{precip} = Total precipitation depth measured by a given strain gauge, k (after zero calibration weight removed) (mm)
A = Strain gauge specific calibration coefficient provided by CVAL (mm*sec)
B = Strain gauge specific calibration coefficient provided by CVAL (mm*sec^2)
f = Frequency for a given strain gauge (Hz)
P_0 = Frequency with an empty collector at calibration, strain gauge specific and provided by CVAL (Hz)
f_0 = Frequency with oil/antifreeze added to the collector at zero calibration, strain gauge specific and provided by CVAL (Hz)

The average depth measurement across all three strain gauges is calculated when calibrated depths are available from all three strain gauges:

$$\overline{D_\iota} = \frac{1}{3}\sum_{k=1}^{3} D_{precip,k,i} \tag{3}$$

Where:

$\overline{D}$ = The average precipitation depth across the three strain gauges.

Depth measurements are averaged over the site-specific 5-60 minute interval:

$$\overline{D_n} = \frac{1}{n}\sum_{i=1}^{n} \overline{D_k} \tag{4}$$

Where:

$\overline{D_n}$ = Is the 5 to 60-minute average precipitation depth over all strain gauges (mm)

$\overline{D_k}$ = The average precipitation depth across three strain gauges at 0.1-Hz.

n = number of 0.1-Hz measurements in the averaging window

4.5 Smoothing Algorithm

4.5.1 Core Logic

The purpose of the smoothing algorithm is to differentiate depth changes due to precipitation from those induced by the susceptibility of the strain gauges to environmental variation. Raw frequency data from strain gauge-based precipitation collectors, and especially the wire strain gauges of the Belfort AEPG II 600M, can be very noisy and are particularly sensitive to diel temperature variability. The smoothing algorithm described here was adapted from a tool provided by Dr. Ralph Wright used to process weighing gauge precipitation measurements made by the Government of Alberta Environment and Parks service (Barnes and Hopkinson, 2022).

Prior to applying the smoothing algorithm, site-specific parameters are determined based on observed data that characterize the spurious variability in the strain gauge measurements (**Figure 1**). A *Window* is determined for each site that represents the periodicity of spurious variability (typically 24 hours but it can be up to 72 hours), and an *Envelope* is determined that represents the magnitude of spurious variability observed over the *Window* when no precipitation has occurred. Other site-specific parameters used in the algorithm and described below are denoted with *italics*.

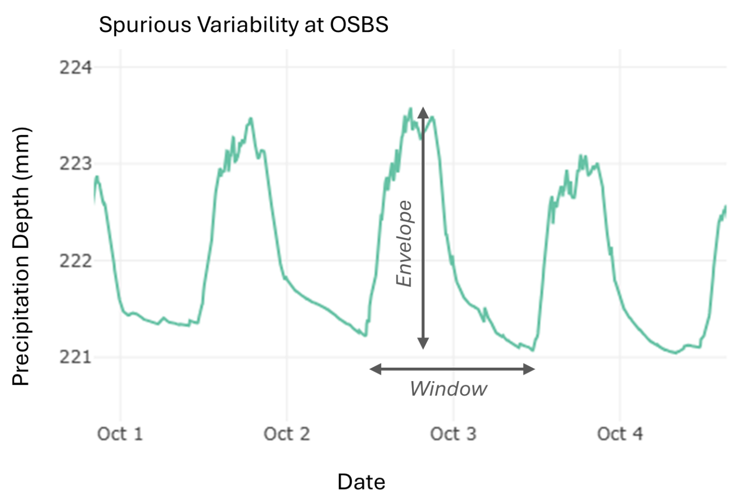

Figure 1. Example characterization of spurious variability in 5-min average calibrated precipitation depth from the Belfort AEPG II 600M sensor at the NEON OSBS site. No precipitation occurred over the period shown.

The smoothing algorithm is applied to rolling 7-day spans of the 5 to 60-minute average precipitation depths as calculated in Eq. (4) (denoted as Raw Depth in the following description). A 7-day span provides sufficient data to avoid edge effects in the center three days of the span, for which precipitation estimates are output. Each subsequent computation span rolls forward by one day. Thus, for any given day there are three overlapping computation spans which are averaged to produce final precipitation estimates.

At the start of each span, the Benchmark Depth is established as a particular *Quantile* (typically the median) of the Raw Depth distribution within the first *Window* of measurements (**Figure 2**). Beginning one measurement after the first *Window* and progressing forward one measurement at a time, the Benchmark is compared to the Raw Depth. If the Raw Depth measurement is greater than the Benchmark Depth, it is marked toward the count for valid precipitation. If the Raw Depth stays consistently above the Benchmark Depth for a specified number of hours (*Threshold Count*, typically 15 hours), the Benchmark Depth is adjusted to match increasing values of the Raw Depth since the start of the count, reflecting precipitation. If at any point the difference between the Raw Depth and the Benchmark Depth is greater than the *Envelope*, the Benchmark Depth is immediately adjusted to the Raw Depth to reflect precipitation. If the Raw Depth is less than the Benchmark Depth, the count toward determining valid precipitation is reset. The evaluation proceeds each subsequent measurement until the end of the computation span, at which point the increases in Benchmark determine bulk precipitation (**Figure 2**).

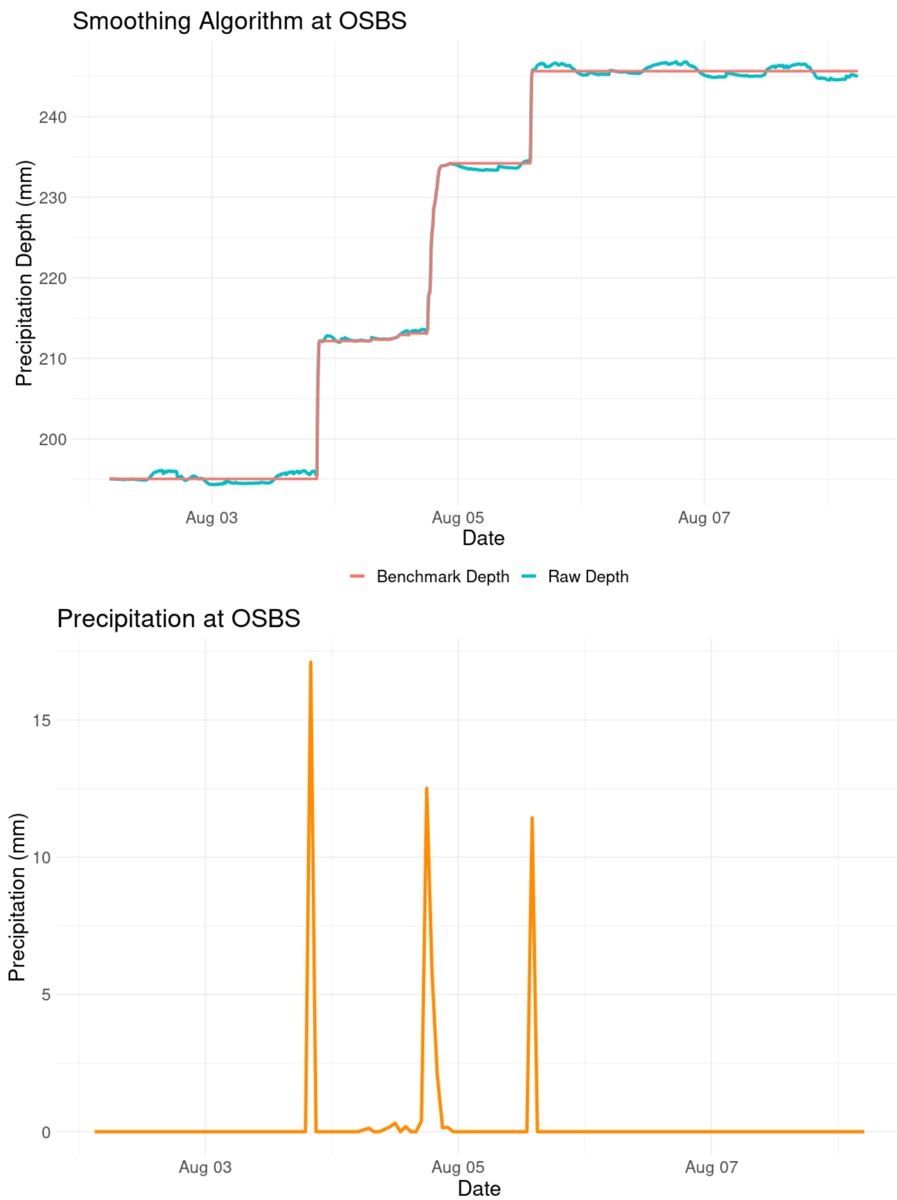

Figure 2: Above: Sample computation span at NEON site OSBS showing the Raw Depth readings and the Benchmark that was determined via the smoothing algorithm. Below: The same data converted to bulk precipitation.

Throughout the computation described above, several checks are performed to handle missing data, modify the Benchmark Depth to account for evaporation of the precipitation within the collector as well as ensure that increases in the Benchmark Depth are not temporary. Some of these additional algorithm details are described below, and the full algorithm code is freely available in the NEON-IS-data-processing GitHub repository hosted by the NEONScience organization.

4.5.2 Evaporation

The Belfort AEPG II 600M sensors are deployed with both an anti-freeze solution and an oil cap to prevent evaporation. While this helps to mitigate losses, it does not prevent them, particularly in warm arid regions. The smoothing algorithm handles evaporative conditions by checking whether the Raw Depth is significantly below the Benchmark Depth. If so, and it has been more than 24 hours since a precipitation event, the Benchmark Depth is adjusted downward to compensate. When this occurs, an informational flag indicating evaporation was detected is raised. An example of how the smoothing algorithm handles evaporation is shown in **Figure 3**.

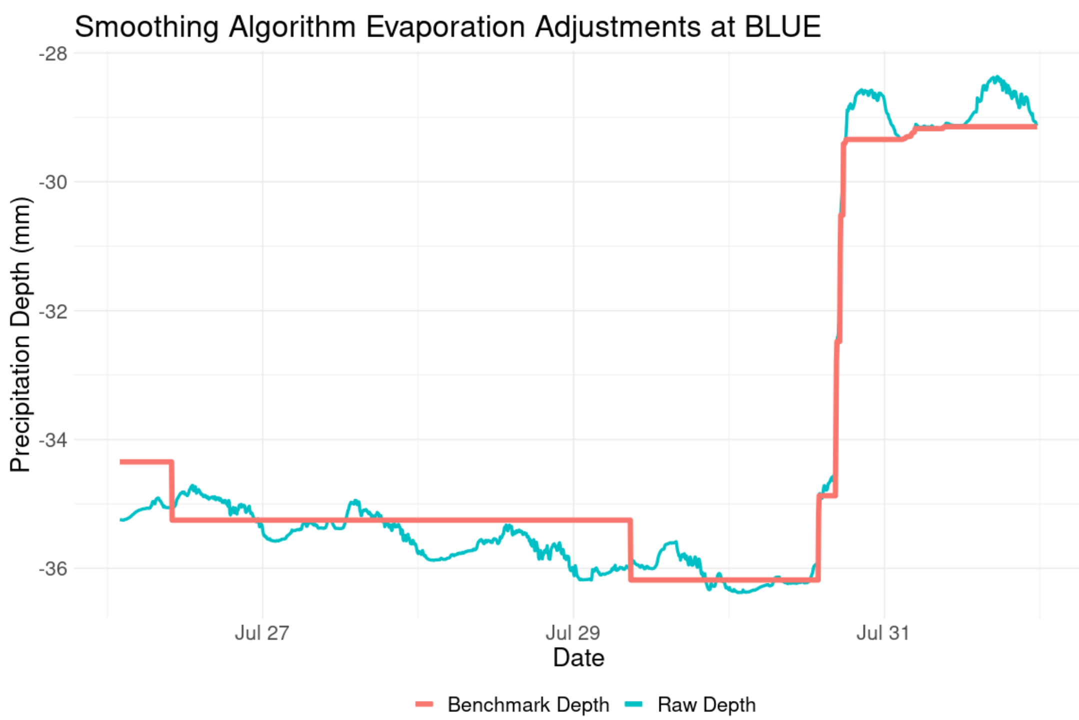

Figure 3: Example of evaporative loss from the precipitation collector at NEON site BLUE. The Benchmark Depth is adjusted down before measuring precipitation event.

4.5.3 Missing Data

Depth values are only calculated when all three strain gauges are available and have passed Quality Control (QC) tests in AD[06] and in the Quality Assurance(QA)/QC procedures described in section 5.2 below. Because the collector depth measurements reflect cumulative precipitation, the algorithm is robust to short data gaps. However, if the evaluation period contains long gaps, the Benchmark Depth and therefore bulk precipitation values are not calculated. If greater than 50% of the data points are missing for the *Window*, the smoothing algorithm progresses forward until a window with sufficient data can be established. A quality flag is raised for the periods of insufficient data. An example of this scenario is shown in **Figure 4**.

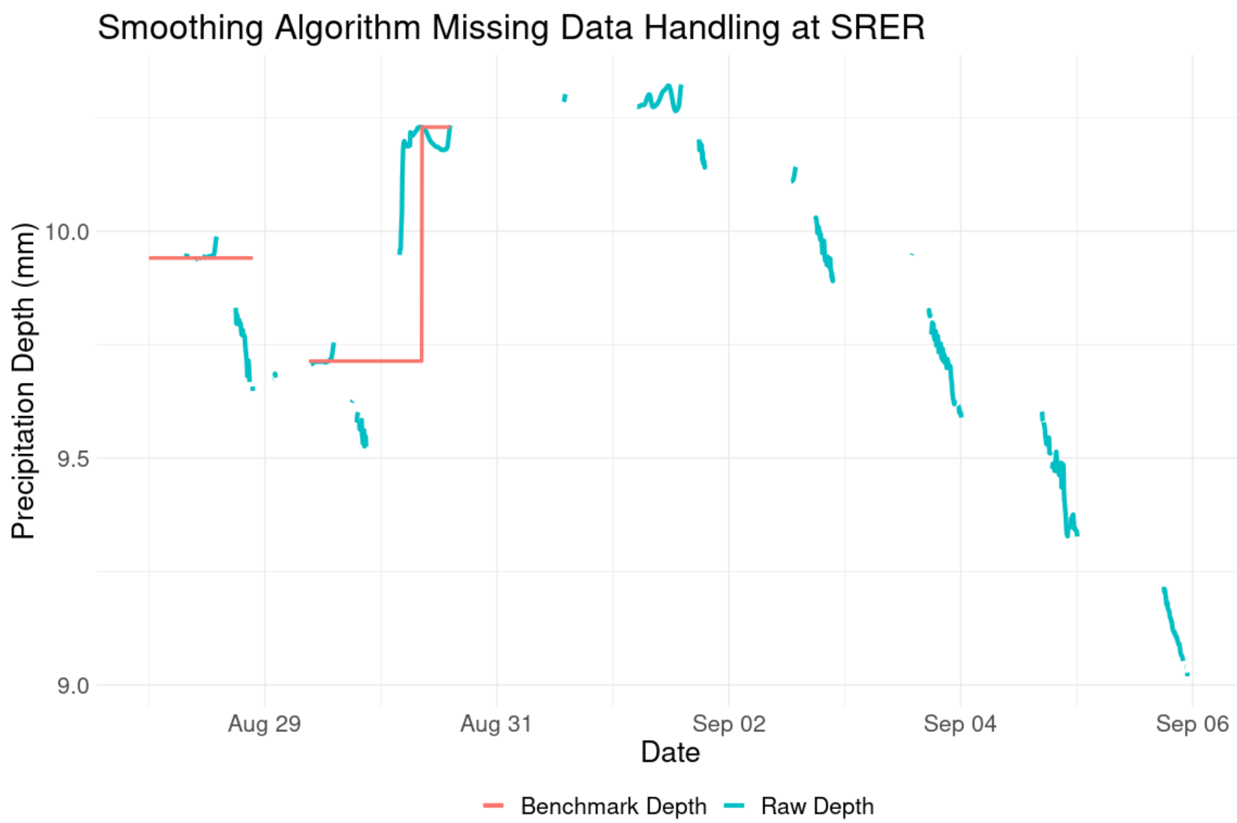

Figure 4: Example of sparse data availability at NEON site SRER. Precipitation values could not be adequately determined, and the insufficient data quality flag will be raised.

4.5.4 Output Reporting Interval

The smoothing algorithm is excellent at estimating precipitation at time scales at or greater than the *Window* (typically 24 hours), even with data exhibiting larger *Envelopes* (> 1 mm) which are often found in the Belfort AEPG II 600m sensor data (recall **Figure 1** for an example of these parameter characterizations). However, precipitation estimates at time intervals less than the *Window* and especially for depth changes within the *Envelope* have greater temporal uncertainty. This is because the algorithm evaluates persistence depth changes relative to a quantile of the data over the *Window* and disregards some changes in collector depth to remove inaccurate sensor noise. Thus, the exact timing of precipitation within the *Window* is not easy to determine. To more confidently represent both the timing and amount of precipitation, precipitation estimates are summed to hourly and daily outputs. A detailed description of the methodology to robustly estimate uncertainty in the smoothing algorithm outputs is provided in Section 6.

5 ALGORITHM IMPLEMENTATION

5.1 Processing Summary

Data flow for signal processing of L1 DPs is treated in the following order.

1. QA/QC tests are applied at the native sensor reporting interval according to the QA/QC Procedure below (Section 5.2) and in accordance with AD[06].
2. Precipitation depth is calculated across all three gauges (Eqs. (2)-(3)).
3. Five-minute to 60-minute average depths is calculated (Eq. (4)).
4. Bulk precipitation at hourly and daily intervals is calculated (Section 4.5).
5. QA/QC tests are applied to the calculated precipitation according to the QA/QC Procedure section detailed below (Section 5.2).
6. Uncertainty for the computed precipitation is calculated (Section 6).
7. Quality flags and metrics are produced for hourly and daily precipitation values in accordance with AD [15].

5.2 Quality Assurance/Quality Control (QA/QC) Procedure

5.2.1 Plausibility Tests

Gauge stability (Eq. (1)), Range, Persistence, Gap, Step, Spike, Null, and Calibration Validity quality tests (AD[06]) are assessed on individual strain gauge measurements. If a data point fails the stability, range, step or spike tests, it is removed from processing. If any of the three gauge values are missing or removed for a given record, the remaining gauge values are also removed. These quality tests are reported for informational purposes only, in that they determine what data is used in further processing but are not used to determine the final quality flag.

5.2.2 Sensor Specific Tests

Multiple sensor specific tests were created to inform the quality assessment of the weighing gauge precipitation data. Computation of quality flags are outlined below as well as whether they contribute to the final quality flag. If any of the tests fail to run, the output for the test is -1.

i. strainGaugeStabilityQF – Indicates whether any of the strain gauges reported unstable status over the reporting interval. The affected records are not used in further processing. This flag is informational.

$$\text{strainGaugeStabilityQF} = \left| \begin{array}{l} 1 \;\; if \; \sum_{k=1}^{3} unstable_k \geq 1 \\ \\ 0 \;\; if \; \sum_{k=1}^{3} unstable_k = 0 \end{array} \right. \quad (5)$$

ii. dielNoiseQF - Overall, the smoothing algorithm adequately separates spurious variability in precipitation depth from true precipitation. However, in instances where the noise *Envelope* is very large, an informational flag is raised to communicate the difficultly in discerning the timing and quantity of small precipitation events. This flag is informational and does not contribute to the final quality flag.

$$\text{dielNoiseQF} = \begin{vmatrix} 1\ when\ Envelope > 10 \\ \\ 0\ \ when\ Envelope \leq 10 \end{vmatrix} \tag{6}$$

iii. evapDetectedQF – An indication that evaporation was detected and accounted for over the duration of the reporting interval (Section 4.5.2). This flag is informational.

$$\text{evapDetectedQF} = \begin{vmatrix} 1\ if\ \ Evaporation\ Detected \\ \\ 0\ \ if\ no\ evaporation\ detected \\ \end{vmatrix} \tag{7}$$

iv. extremePrecipQF – Indicates whether very large values of precipitation are output over the reporting interval and unlikely to have occurred for the given site. The threshold for this test is site-specific basis for hourly precipitation. If the flag is raised for hourly precipitation, daily precipitation is flagged as well. This test contributes to the final quality flag.

$$\text{extremePrecipQF} = \begin{vmatrix} 1\ when\ any\ hourly\ precipitation\ \ > threshold \\ \\ 0\ \ when\ any\ hourly\ precipitation\ \ \leq threshold \end{vmatrix} \tag{8}$$

v. insuffDataQF – An indicator of whether enough data was present to be able to confidently report precipitation over the reporting interval. If the flag is raised for hourly precipitation, daily precipitation is flagged as well. This test contributes to the final quality flag.

$$insuffDataQF = \begin{vmatrix} 1\ when \geq 50\%\ data\ points\ missing\ in\ Window \\ \\ 0\ \ when < 50\%\ data\ points\ missing\ in\ Window \end{vmatrix} \tag{9}$$

vi. heaterErrorQF – Indicates whether a heating error occurred during the reporting interval. This is computed by comparing the inlet temperature (where the heaters are located) to the internal sensor temperature (assumed to represent ambient temperatures) when conditions exist that should result in heater operation. If the inlet temperature is less than the internal temperature, then the heaterErrorQF is raised. If the heater status indicates heaters are enabled but the temperature is above the heater set point, then the heaterErrorQF is raised. See **Figure 5** for details. If more than 50% of the reporting interval (hourly or daily) has a heater error flag, the reporting interval is flagged. This test contributes to the final quality flag.

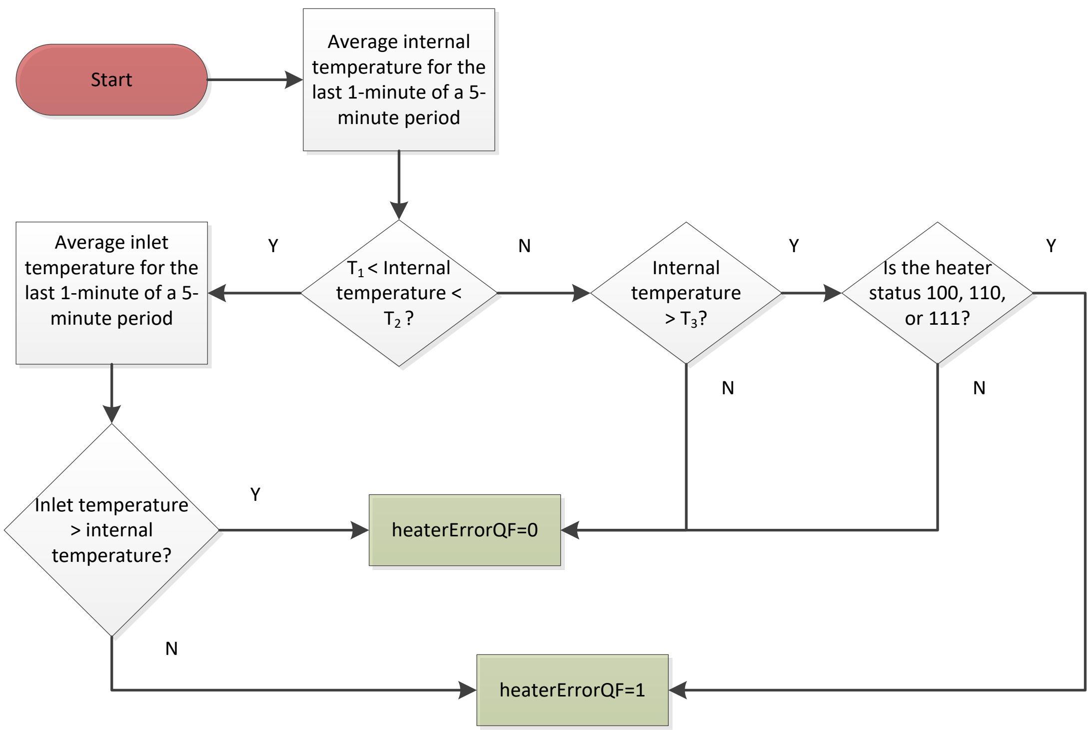

Figure 5. Heater error flag logic, temperature thresholds of T_1 = -6 °C, T_2 = 2 °C, and T_3 = 6 °C.

5.2.3 Quality Flags (QFs) and Quality Metrics (QMs)

In accordance with AD [15], the quality flags listed in **Table 3** will accompany each weighing gauge precipitation L1 DP. Inlet heater quality metrics summarize the inputs to the inletHeaterQF quality flag. There are three inlet heater quality metrics that correspond to the number of heaters that were operational during that period and a fourth that corresponds to the percent of measurements that had no heater information (i.e., missing). Accordingly, inletHeaters1QM corresponds to the percent of inletHeaterQF=100, inletHeaters2QM to the percent of inletHeaterQF=110, inletHeaters3QM to the percent of inletHeaterQF=111 over the reporting interval, and inletHeaterNAQM is the percent of missing heater data over the measurement period (see **Figure 5**). The final quality flag (finalQF) is determined according to Eq.10 for hourly and daily bulk precipitation output.

$$finalQF = \begin{cases} 1\ if\ any(\boldsymbol{QF}^{*}\ in\ Table\ 2 = 1\ or\ \boldsymbol{QM}^{*} > 0\%) \\ 0\ otherwise \end{cases} \quad (10)$$

Table 3. Quality metrics and flags associated with hourly and daily bulk precipitation output. Bolded and starred entries contribute to the final quality flag.

Quality Metric (QM) / Quality Flag (QF)
spikeQM
stepQM
rangeQM
gapQM
persistenceQM
nullQM
validCalibrationQM
suspectCalibrationQM*
inletHeaters1QM
inletHeaters2QM
inletHeaters3QM
inletHeaters4QM
evapDetectedQF
dielNoiseQF
strainGaugeStabilityQF
extremePrecipQF*
heaterErrorQF*
insuffDataQF*
finalQF

6 UNCERTAINTY

Uncertainty of measurement is inevitable; therefore, measurements should be accompanied by a statement of their uncertainty for completeness (JCGM 2008; Taylor 1997). To do so, it is imperative to identify all sources of measurement uncertainty related to the quantity being measured. Quantifying the uncertainty of measurements provides a measure of the reliability and applicability of individual measurements and data products. This portion of the document serves to identify, evaluate, and quantify sources of uncertainty relating to individual, calibrated depth measurements as well as the L1 bulk precipitation data product. It is a reflection of the information described in AD [13] and is explicitly described for the weighing gauge precipitation assembly in the following sections.

Sources of identifiable uncertainties include those arising from the sensor, calibration procedure, evaporation, wind, wetting, representativeness (Nemec 1969; Humphrey *et al.* 1997; Brock and Richardson 2001; WMO 2008), and the application of the smoothing algorithm. The DFIR setup (i.e., NEON's weighing gauge precipitation assembly) provides more accurate cumulative precipitation than other measurement techniques such as a tipping bucket (Rasmussen 2012). All types of identified uncertainties are detailed in the following sections.

6.1 Measurement Uncertainty

The following subsections present the uncertainties associated with individual collector depth values. It is important to note that the uncertainties presented in the following subsections are *measurement uncertainties*, that is, they reflect the uncertainty of an *individual* collector depth measurement. These uncertainties should not be confused with those presented in Section 6.2 which describes the procedure for estimating uncertainty in reported bulk precipitation values. We urge the reader to refer to AD[11] for further details concerning the difference between quantification of measurement uncertainties and L1 uncertainties.

Measurement uncertainties are qualified according to recommendations of the Joint Committee for Guides in Metrology (JCGM) 2008. In essence, if a measurand *y* is a function of *n* input quantities x_i $(i = 1, \dots, n)$, $i.e., y = f(x_1, x_2, \dots, x_n)$, the combined measurement uncertainty of *y*, assuming the inputs are independent, can be calculated as follows:

$$u_c(y) = \left(\sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) \right)^{\frac{1}{2}} \quad (11)$$

Where:

$\frac{\partial f}{\partial x_i}$ = partial derivative of *y* with respect to x_i

$u(x_i)$ = combined standard uncertainty of x_i.

Thus, the uncertainty of the measurand can be found be summing the *quantifiable* input uncertainties in quadrature. The calculation of these *quantifiable* input uncertainties is discussed below.

6.1.1 Calibration

An individual collector depth measurement is a combination of three calibrated strain gauge measurements. NEON's CVAL applies unique calibration coefficients to each strain gauge (Eq.(2)) but provides a measurement uncertainty estimate that collectively accounts for all three gauges. In other words, the estimate represents an individual depth reading (Eq. (2)), with the assumption that all three gauges are stable and equally weighted (AD [16]) such an individual gauge may estimate the entire collector depth. The measurement uncertainty is a relative value that provided by CVAL (AD [12]), stored in the CI data store, and applied to each Raw Depth measurement (Eq. (3)) .

$$u(D_{precip,k,i}) = u_{A1} * \; D_{precip,k,i} \quad (12)$$

Where:

u_{A1} = relative uncertainty of calibrated depth measurement (%)
$u(D_{precip,k,i})$ = standard uncertainty of calibrated depth measurement (mm)

6.1.2 Data Acquisition System (DAS)

The weighing gauge quantifies precipitation in units of Hz captured through the serial port of the Data Acquisition System (DAS). Thus, the signal, although analog, is treated as a digital signal and uncertainties introduced by the DAS are considered negligible.

6.1.3 Expanded Measurement Uncertainty

The expanded measurement uncertainty is calculated as:

$$U_{95}\left(D_{precip,k,i}\right) = k_{95} * u(D_{precip,k,i}) \quad (13)$$

Where:

$U_{95}\left(D_{precip,k,i}\right)$ = expanded measurement uncertainty at 95% confidence (mm)
k_{95} = 2; coverage factor for 95% confidence (unitless)

6.2 Evaporative Loss

The Belfort AEPG II 600M sensors are deployed with both an anti-freeze solution and an oil cap to prevent evaporation. While this helps to mitigate losses, it does not prevent them, particularly in warm arid regions. In addition, the use of heaters and solar radiation can cause a weighing gauge's inlet to be warmer than the ambient environment thus causing a chimney effect (Rasmussen 2012). If this occurs for prolonged periods before or during precipitation events, evaporative losses and associated measurement uncertainty can amplify (Brock and Richardson 2001; WMO 2008). Currently, NEON cannot confidently quantify the specific contribution of evaporation to the measurement uncertainty. However, the uncertainty estimates for the L1 DP as described in Section 6.2 below are inclusive of the uncertainty due to evaporation.

6.3 Wind

The measurement of precipitation is particularly sensitive to wind (WMO 2008). Laminar and turbulent flows can result in a reduction of catch at the inlets of precipitation gauges, thus resulting in underestimations of precipitation measurements. Brock and Richardson (2001) note that catch reductions can be up to 20% with winds ranging from 5 to 10 m s^{-1} and nearly 80% for winds >10 m s^{-1} during light rainfall and most snowfall events. Wind speeds near the weighing gauge can be reduced and catch reduction can be mitigated by shielding the precipitation gauge with shields such as fencing (WMO 2008). Such is the case for the NEON's weighing gauge precipitation assembly, which comprises an alter shield and two fences around the weighing gauge.

6.4 Wetting

Wetting is a term used to describe a buildup of precipitation at the inlet of a precipitation sensor (Groisman and Legates 1994). In most cases such precipitation would evaporate before falling *into* the weighing gauge and would not be quantified, thus causing an underestimation of precipitation due to *wetting loss.* Such losses are small (Sevruk 1982), and given the magnitude of other uncertainties (i.e., wind induced), NEON considers wetting losses to be negligible.

6.5 Representativeness

It is argued that any type of precipitation gauge is unrepresentative of precipitation over large areas – caution should be executed when spatially interpolating and extrapolating precipitation measurements. It is considered poor sampling when one precipitation gauge is used to represent precipitation characteristics of a larger surrounding area (e.g., 200 km^2); this is especially true during thunderstorms (Rinehart 2004; WMO 2008). Passing of a localized rainstorm can grossly overestimate (if directly over the gauge) or underestimate (if storm misses gauge completely) precipitation characteristics for a mesoscale sized region (Brock and Richardson 2001). With the aid of radar imagery, representativeness can be better understood.

6.6 Uncertainty of Bulk Precipitation

The characterization of individual measurement uncertainty described in Section 6.1 above does not account for the susceptibility of the strain gauge measurements to environmental variation and resultant propagation of that uncertainty through the smoothing algorithm to compute hourly and daily bulk precipitation. Due to the complexity of the approach applied in the smoothing algorithm, uncertainty estimates for the L1 bulk precipitation are generated using a numerical Monte Carlo approach with surrogate depth timeseries. This approach is common in science and engineering applications where analytical computation of uncertainty is impractical or impossible (Kroese et al. 2014). In this approach, synthetic timeseries of collector depth data are generated that exhibit the characteristics of the measured data but are otherwise random. These surrogates are then processed with the smoothing algorithm with the same algorithm parameters as applied to the measured data. The variability in the output of the surrogate data is used to compute the uncertainty in the reported bulk precipitation. The following sections describe these procedures in greater detail, and the full algorithm code is freely available in the NEON-IS-data-processing Github repository hosted by the NEONScience organization.

6.6.1 Surrogate timeseries

Thirty surrogate depth timeseries are generated for each 7-day computation span. First, the Benchmark Depth is computed with the smoothing algorithm for the original data (see Section 4.5). The Benchmark Depth is subtracted from the Raw Depth to remove precipitation events and isolate spurious and otherwise random variability. Thirty random representations of the spurious variability are generated via the Iterated Amplitude Adjusted Fourier Transform (IAAFT) method (Schreiber and Schmitz 1996). IAAFT surrogates match the periodicity, autocorrelation structure, and marginal distribution of the original timeseries but are otherwise random. The Benchmark is then added back to the IAAFT surrogates. An example of the resultant surrogate depths is show in **Figure 6A**. The surrogate depth timeseries are processed along with the measured precipitation depth timeseries with the same algorithm settings, resulting in thirty surrogate Benchmark Depth timeseries (examples in **Figure 6B)**.

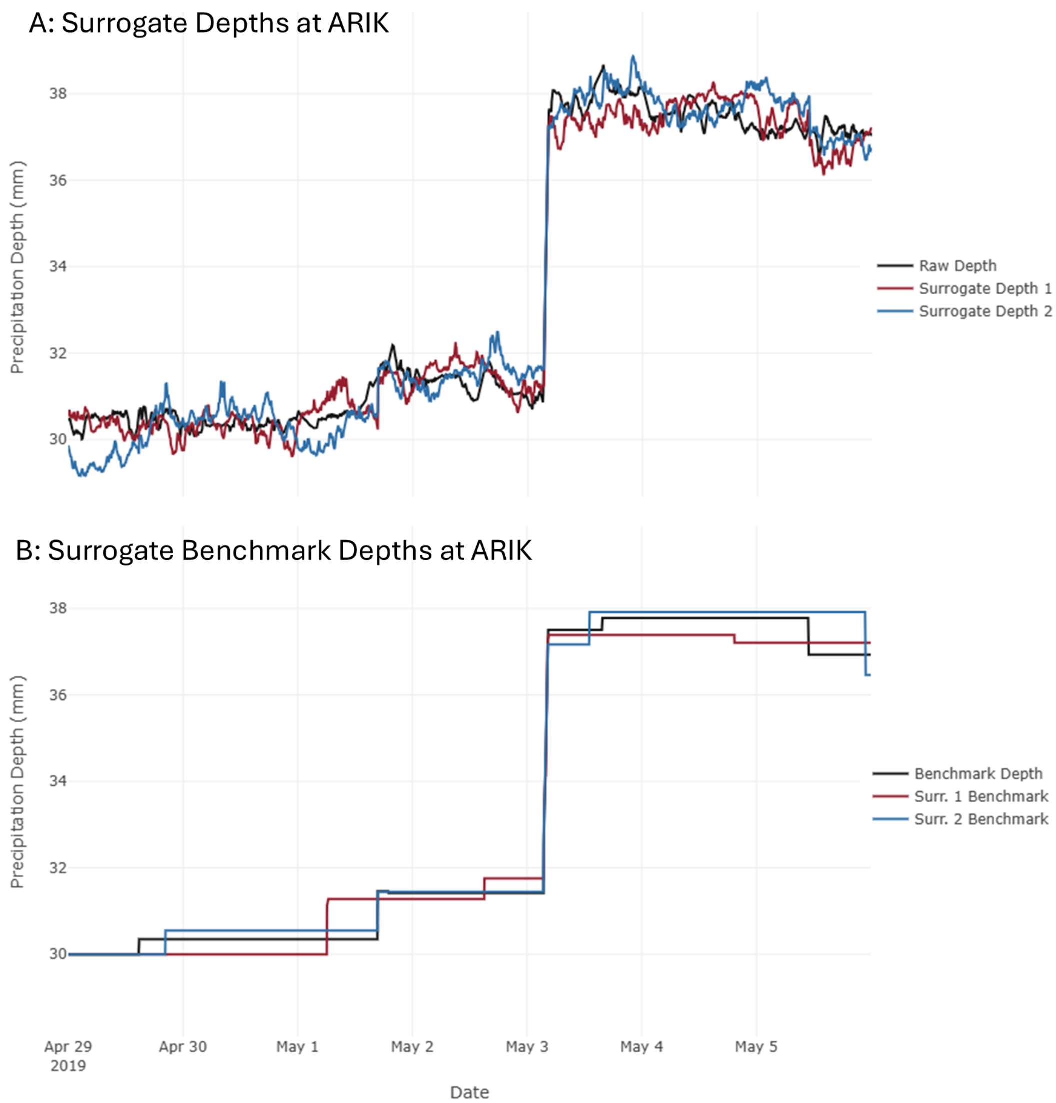

Figure 6. Example of surrogate testing for uncertainty computations at ARIK. **(A)** The measured Raw Depth along with two surrogates (a total of 30 are generated and tested). **(B)** The computed Benchmark Depths after processing the measured and surrogage precipitation depths in subplot A with the smoothing algorithm.

6.6.2 L1 Uncertainty estimation

The uncertainty in the Benchmark Depth at any given time point is computed as the standard deviation of the surrogate Benchmark Depths:

$$u(D_{bench,t}) = \sqrt{\frac{\sum_{k=1}^{n} D_{bench,k,t}}{n-1}} \tag{14}$$

Where:

$D_{bench,k,t}$ = Benchmark Depth of surrogate k at time point t (mm)
$u(D_{bench,t})$ = uncertainty in Benchmark Depth at time point t, expressed as a standard deviation (mm)
n = number of surrogates

L1 bulk precipitation at hourly and daily intervals is computed by subtracting the Benchmark Depth at the start of the interval from that at the end of the interval. Assuming that the errors in the Benchmark Depths at any two time points can be treated as random variables, the uncertainty in the difference of the Benchmark Depths is equal to the sum of the uncertainties at the two time points, added in quadrature:

$$u(P_{bulk,t-x}) = \sqrt{u\left(D_{bench,t}\right)^2 + u\left(D_{bench,t-x}\right)^2} \tag{15}$$

Where:

$u(P_{bulk,t-x})$ = uncertainty in bulk precipitation occurring between timepoint $t-x$ and timepoint t, expressed as a standard deviation (mm)

The computation in Eq. (15) is valid between any two continuous increases in Benchmark Depth, as the start and end points can be directly differenced. When there is a decrease in the Benchmark Depth timeseries, due to bucket emptying or evaporation handling, the uncertainty in bulk precipitation is computed for each continuously increasing span, and the uncertainties for each continuous span again added in quadrature for time ranges that include multiple continuous spans. Note that the uncertainty in bulk precipitation generated by this procedure includes random measurement uncertainty as well as model uncertainty, as the surrogates intentionally incorporate variability observed in the depth measurement, which is then propagated through the model.

6.6.3 Expanded Uncertainty

The expanded uncertainty is calculated as:

$$U_{95}\left(P_{bulk,t-x}\right) = k_{95} * u(P_{bulk,t-x}) \tag{16}$$

Where:

$U_{95}\left(P_{bulk,t-x}\right)$ = expanded uncertainty in bulk precipitation at 95% confidence (mm)

k_{95} = 2; coverage factor for 95% confidence (unitless)

6.7 Communicated Precision

The communicated precision of L1 bulk precipitation and expanded uncertainty is 0.01 mm.

6.8 Uncertainty Budget

The uncertainty budget is a visual aid detailing i) quantifiable sources of uncertainty, ii) means by which they are derived, and iii) the order of their propagation.

Table 4. Uncertainty budget for individual precipitation measurements.

Source of measurement uncertainty	**measurement uncertainty component $u(x_i)$**	**measurement uncertainty value**	$\frac{\partial f}{\partial x_i}$	$u_{x_i}(Y) \equiv \left\lvert\frac{\partial f}{\partial x_i}\right\rvert u(x_i)$ **(mm)**
Precipitation depth measurement	$u(D_{precip,k,i}) = u_{A1} * D_{precip,k,i}$	Eq. (12)	n/a	n/a

Table 5. Uncertainty budget for bulk precipitation measurements.

Source of uncertainty	**uncertainty component $u(x_i)$**	**uncertainty value**	$\frac{\partial f}{\partial x_i}$	$u_{x_i}(Y) \equiv \left\lvert\frac{\partial f}{\partial x_i}\right\rvert u(x_i)$ **(mm)**
Precipitation depth benchmark	$u(D_{bench,t}) = \sqrt{\frac{\sum_{k=1}^{n} D_{bench,k,t}}{n-1}}$	Eq. (14)	n/a	n/a
Bulk precipitation	$u(P_{bulk,t-x}) = \sqrt{u(D_{bench,t})^2 + u(D_{bench,t-x})^2}$	Eq. (15)	n/a	n/a

7 ACKNOWLEDGEMENTS

NEON would like to acknowledge Dr. Ralph Wright and his generosity in sharing his weighing gauge and time series analysis processing tool processing code (Wright 2021). His work helped improve the data quality of the observatory. NEON would also like to thank Josh Roberti and Derek Smith for previous contributions to algorithm code and documentation.

8 BIBLIOGRAPHY

Baker, C. B., Buckner, R., Collins, W., and M. Phillips. (2005) Calculation of USCRN precipitation from Geonor weighing precipitation gauge. NOAA Tech. Note NCDC No. USCRN-05-1. [Online]. Available: http://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/program/technotes/TN05001GeonorAlgorithm.pdf [December, 2013].

Bakkehøi, S., Øien, K., and E.J. Førland. (1985) An automatic precipitation gauge based on vibrating-wire strain gauges. Nordic Hydrology. **16**, 193-202, DOI: 10.2166/nh.1985.015.

Barnes, C., & Hopkinson, C. (2022). Quality Control Impacts on Total Precipitation Gauge Records for Montane Valley and Ridge Sites in SW Alberta, Canada. *Data*, *7*(6), 73. https://doi.org/10.3390/data7060073.

Belfort. (2014) Universal all environment precipitation gauge model APEG II 1200 instruction manual MK III. Belfort instrument Company. pp. 55, Manual Number 90008-3

Brock F. V., and Richardson S. J. (2001) Meteorological measurement systems. Oxford University Press, pp. 290.

Campbell Scientific. (2011) Geonor T-200B series precipitation gauge. Rev. 10.7 [Online]. Available: http://s.campbellsci.com/documents/ca/manuals/geonort-200b_man.pdf [December, 2013].

Collins, W. G., Baker, C. B., and T. Wilson. (2005) The new precipitation algorithm for the three-wire Geonor gauge of the U.S. climate reference network – objectives, description and performance. AMS 13th Symposium on Meteorological Observations and Instrumentation. [Online]. Available: https://ams.confex.com/ams/pdfpapers/ 93695.pdf [December, 2013].

CRN. (2002) Climate reference network site information handbook. NOAA/NESDIS CRN Series X030. [Online]. Available: http://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/program/X030FullDocumentD0.pdf [December, 2013].

Groisman, P. Y., and Legates, D. R. (1994). The accuracy of United States precipitation data. Bulletin of the American Meteorological Society., **75**, 215-227.

Humphrey M. D., Istok J. D., Lee J. Y., Hevesi J. A., and Flint A. L. (1997) A new method for automated dynamic calibration of tipping-bucket rain gauges. *J. Atmos. Ocean Tech.*, **14**, 1513 - 1519

Joint Committee for Guides in Metrology. (2008) Evaluation of measurement data – Guide to the expression of uncertainty in measurement. pp. 120.

Joint Committee for Guides in Metrology. (2012) International vocabulary of metrology – Basic and general concepts and associated terms (VIM). 3rd Edition. pp. 92

Kroese, D.P., Brereton, T., Taimre, T., Botev, Z.I. 2014. Why the Monte Carlo method is so important today. WIREs Computational Statistics 6, 386–392. https://doi.org/10.1002/wics.1314

Lamb, H. H. and J. Swenson. (2005) Measurement errors using a Geonor weighing gauge with a Campbell Scientific datalogger. AMS 16th Conference on Climate Variability and Change. [Online]. Available: https://ams.confex.com/ams/Annual2005/techprogram/paper_84196.htm [December, 2013].

Leeper, R.D., Palecki, M.A., and E. Davis. (2015) Methods to calculate precipitation from weighing-bucket gauges with redundant depth measurements. American Meteorological Society. **32**, 1179-1190, DOI: 10.1175/JTECH-D-14-00185.1.

Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A.P., Black, J., Theriault, J.M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., and E. Gutmann. 2012. How well are we measuring snow? The NOAA/FAA/NCAR winter precipitation test bed. Bull. Amer. Meteor. Soc., **93**, 811-829, DOI: 10.1175/BAMS-D-11-00052.1.

Schreiber, T., Schmitz, A. 1996. Improved surrogate data for nonlinearity tests. Phys. Rev. Lett. 77, 635–638. https://doi.org/10.1103/PhysRevLett.77.635

Sevruk, B. (1982) Methods of correction for systematic error in point precipitation measurement for operational use. Operational Hydrology Report 21. World Meteorological Organization. pp. 91.

World Meteorological Organization (WMO) (2008) Guide to meteorological instruments and methods of observations: Chapter 6 - Measurement of precipitation. pp. 23

Wright, R. *Weighing Gauge Time Series Analysis and Noise Filtering Tool*; Alberta Agriculture and Forestry: Edmonton, AB, Canada, 2021.