# U.S. Climate Reference Network (USCRN) Program

## Official Algorithm for Precipitation

## Version 2.0

## Description

July 9, 2015

**Introduction**

The U.S. Climate Reference Network (USCRN) Program is designed to monitor the climate of the United States using research quality instrumentation located within pristine environments that are representative of regions and not likely to undergo human encroachment or land use change for the next 50 years. The primary observation variables are air temperature, precipitation, and soil moisture/soil temperature. As primary variables, each station collects three independent measurements for each type of observation, and uses this redundancy to ensure the quality and continuity of the climate record. In the case of precipitation, the redundancy is produced in the form of three independent measurements of precipitation depth in a weighing bucket gauge, supplemented by a wetness sensor (disdrometer) indicating the presence or absence of precipitation. Since 2004-2005, all these (precipitation) measurements have been made at 5-minute intervals. The existing algorithm for calculating precipitation from the USCRN station configuration was developed by Baker et al. (2005), and has performed adequately. However, cases of known difficulties have been documented, and this led to an effort to improve the algorithm for calculating precipitation by more fully utilizing the information available from the redundant measurements.

Requirements for a new Official Algorithm for Precipitation (OAP) were developed based on more than a year of fact finding and analysis from 2011 to 2013. The output from the existing algorithm was subject to intense scrutiny. Several weaknesses in the existing algorithm were identified, including a tendency to report slightly smaller total amounts of precipitation than other nearby gauges, and on occasion record unrealistic 5-minute intensity at the beginning of events. The exploration for a new OAP started with various attempts to modify the current algorithm; however, these approaches were not capable of resolving all issues. The specifications for OAP 2.0 were developed through testing against both station data and artificial event inputs. The final set of requirements was approved by the USCRN Program Configuration Control Board (CCB) in Configuration Change Request (CCR) 46 on 21 April 2014. The new OAP was documented in a refereed journal publication which was accepted in final form during May 2015 (Leeper et al. 2015a; Appendix 1).

During the review process, attention also was paid to events in the record when equipment malfunctioned but data still passed basic quality control, resulting in incorrect precipitation calculations. These events are being recorded as exceptions based on procedures adopted by the USCRN Program CCB CCR 42, and are being corrected or flagged as part of the OAP 2.0 project.

Figure 1. (a) Weighing gauge within metal wind shield; (b) 3 load cells measured bucket depth

**Station Configuration**

To address the challenges of accurate precipitation observation with an automated station, the USCRN adopted an innovative approach to monitor precipitation redundantly from a well-shielded enclosure (Figure 1a). The gauge is configured such that the reservoir (weighing bucket in Figure 1b) is suspended from three independent load sensors. The load cells observe reservoir weight by magnetically plucking an internal wire and monitoring its frequency of vibration, which will vary with wire tension caused by the weight of precipitation in the bucket. When calibrated, the Geonor gauge can reliably detect changes in gauge depth to a resolution of 0.1 mm, although field testing by the USCRN Program resulted in a finding that a minimum depth change increment of 0.2 mm is required to assure proper precipitation calculation for a 5-minute time interval. The redundant sensors improve the resilience of the precipitation observing system against single sensor degradation and failure as the additional sensors continue to monitor precipitation until the gauge can be repaired. The redundant monitoring also ensures the continuity of the data record, in addition to providing information that can be used to improve the detection of a precipitation signal from gauge noise.

The USCRN station engineering also endeavors to reduce environmental issues that can negatively impact precipitation measurements. To inhibit frozen precipitation from collecting on the interior walls and capping the gauge, a heating tape is applied to the weighing gauge throat for stations located in colder climates. The USCRN lessens the impacts of surface winds by observing precipitation from a well-shielded enclosure. A majority of USCRN gauges are surrounded by a small double fence inter-comparison reference shield (SDFIR) with an interior single alter shield (as in Figure 1a). In locations where siting and/or material transport become an issue (mostly in Alaska), the gauges are shielded solely by a double alter shield.

While clearly beneficial, redundant measurements of gauge depth increase the complexity of quality assurance (QA) systems by requiring both traditional quality control (QC) checks on raw gauge depths and a computational algorithm to compute an observation quantity from the redundant measures of gauge depth change. As an early adopter of redundant technology, the USCRN has pioneered the development of QA systems that process redundant measurements to enhance the quality of observations (air temperature, precipitation, and soil moisture/soil temperature) and continuity of the data record. The current QA system features a pairwise comparison of depth changes and a calculation that has evolved over time as a disdrometer (used to detect wetness due to precipitation) and higher time resolution (5-minute) observations were brought online during an earlier period in network history.

**Dataset Change**

There are more than 400 variables representing data types and time intervals in USCRN hourly station records. A subset of 61 variables is required for precipitation calculation each hour, consisting of values for two wetness sensor channels and three weighing bucket gauge depths for each 5-minute interval, a count of minutes that the data logger door was open during the hour, and the flags for these variables. The precipitation calculation output variables that will be changed by the new OPA are shown in Table 1. It is important to note that all raw inputs to the precipitation calculation have been and will be preserved permanently so that precipitation calculations can be revisited in the future.

Table 1. USCRN Database Elements affected by OAP 2.0.

| | | |
|---|---|---|
| 318 | P_OFFICIAL | calculated Geonor precip total for hour |
| 319 | P5_1 | calculated Geonor precip for 5 minutes ending at :05 |
| 320 | P5_2 | calculated Geonor precip for 5 minutes ending at :10 |
| 321 | P5_3 | calculated Geonor precip for 5 minutes ending at :15 |
| 322 | P5_4 | calculated Geonor precip for 5 minutes ending at :20 |
| 323 | P5_5 | calculated Geonor precip for 5 minutes ending at :25 |
| 324 | P5_6 | calculated Geonor precip for 5 minutes ending at :30 |
| 325 | P5_7 | calculated Geonor precip for 5 minutes ending at :35 |
| 326 | P5_8 | calculated Geonor precip for 5 minutes ending at :40 |
| 327 | P5_9 | calculated Geonor precip for 5 minutes ending at :45 |
| 328 | P5_10 | calculated Geonor precip for 5 minutes ending at :50 |
| 329 | P5_11 | calculated Geonor precip for 5 minutes ending at :55 |
| 330 | P5_12 | calculated Geonor precip for 5 minutes ending at :60 |

**New Official Algorithm for Precipitation (OAP 2.0)**

The procedures for the new precipitation calculation algorithm are shown in Figure 2. In the article comparing the new approach to the previous (see Appendix 1), the shorthand name for the method was wavgCalc, as the algorithm is based on a weighted average approach to combining the three wire depth change (delta) signals into a single value. Initially, several

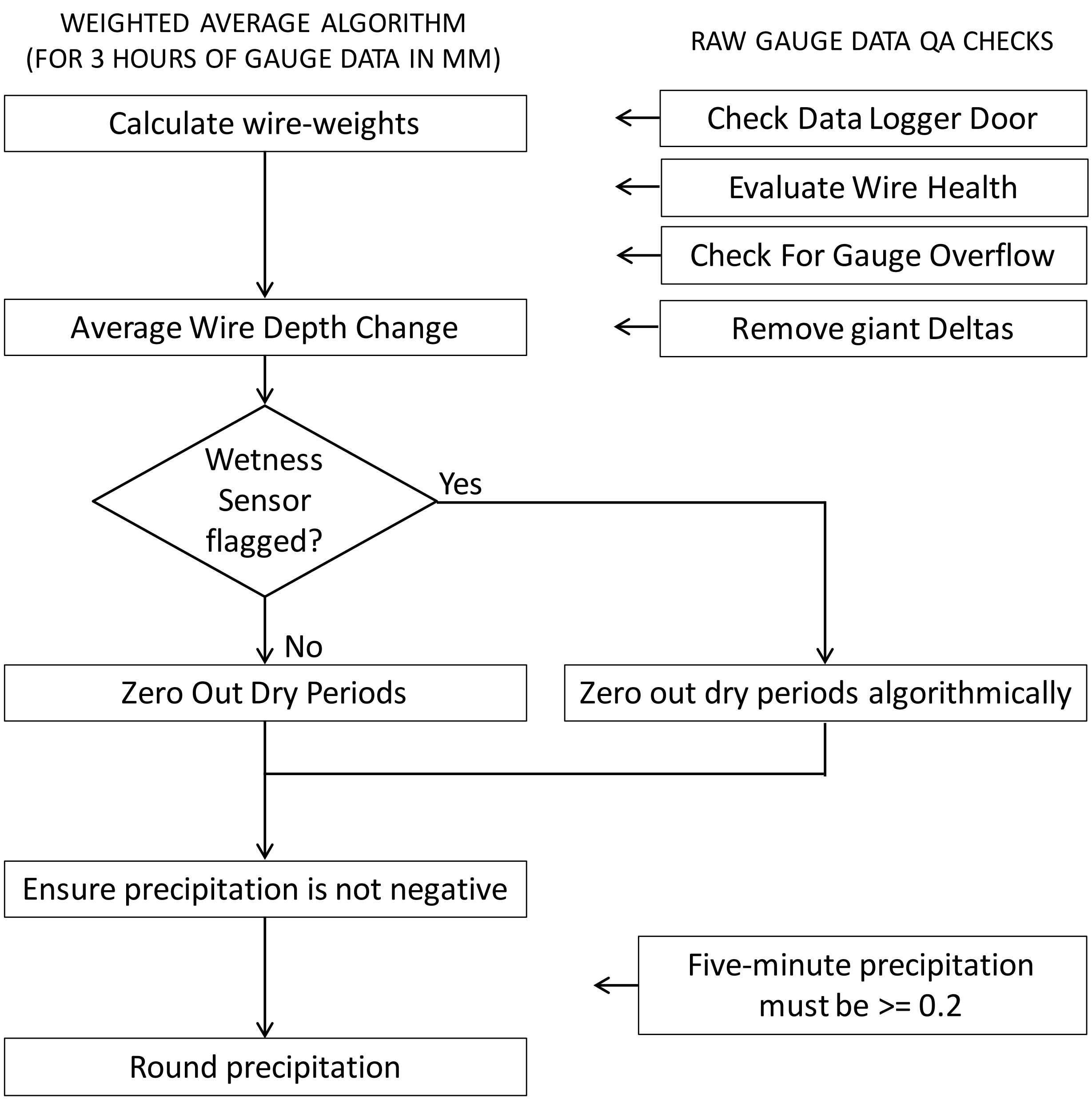



Figure 2. Flow chart detailing the order of OAP 2.0 (wavgCalc) procedures for calculating precipitation.

quality control steps are performed on the depth and wetness data directly. First, the data logger door variable is interrogated to see if anyone was performing maintenance on the station, which is indicated by opening the data logger door. Precipitation is not calculated during maintenance periods. Second, the individual wires that record depth changes are examined to ensure their operational status. When wires in the load cells break, they usually report a highly negative depth (less than -15 mm is flagged) as zero wire vibration input into the calibration equation for depth yields a number in the -60 to -80 mm range. Precipitation is calculated only using wires that pass this check with a minimum of at least two healthy wires remaining. Third, the depth values over the stated capacity of the gauge (either 600 mm or 1000 mm, depending on location) are flagged, and precipitation is not calculated. Finally, very large depth changes called "giant deltas" are flagged and not used in precipitation calculations. Positive depth change values over 25 mm in 5 minutes are flagged automatically, as changes this rapid are likely due to a cause unrelated to precipitation, such as placing antifreeze in the bucket without opening the data logger door, or drifted snow stuck to the collar slumping into the gauge. Values can be restored manually using the exception process if found to be real.

The non-flagged depth and wetness data from the previous two (where available; one previous hour at the minimum) and current hour are processed through the calculation system. Wire deltas are computed as a change in gauge depth between successive sub-hourly periods in a manner similar to the way manual evaluations might be performed (current sub-hourly depth minus previous sub-hourly depth). Wire deltas are then averaged using a weighted mean. Wire weights are determined based on the average delta variance of each wire over the three-hour period. The delta variance is calculated for each of the three wires $k$=1,2,3 (Eq. 1) where $X$ is delta, $i$ is the $i$th sub-hourly period, $\bar{X}$ is the three-wire delta mean, and $n$ is the total number of sub-hourly periods.

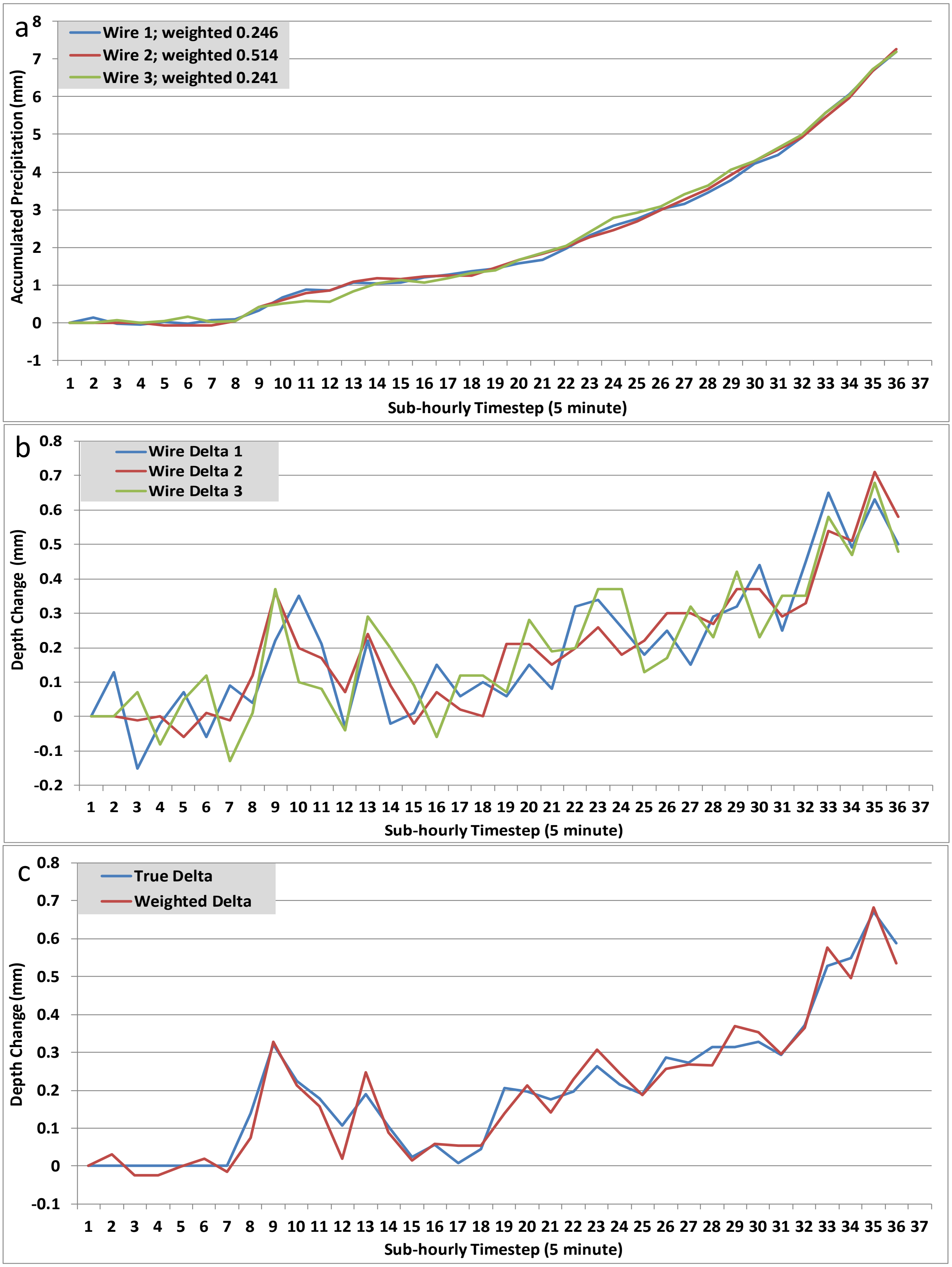
$$delta\ variance_k = \frac{\sum_i^n (X_{ik} - \bar{X}_i)^2}{n} \qquad \text{(Eq. 1)}$$

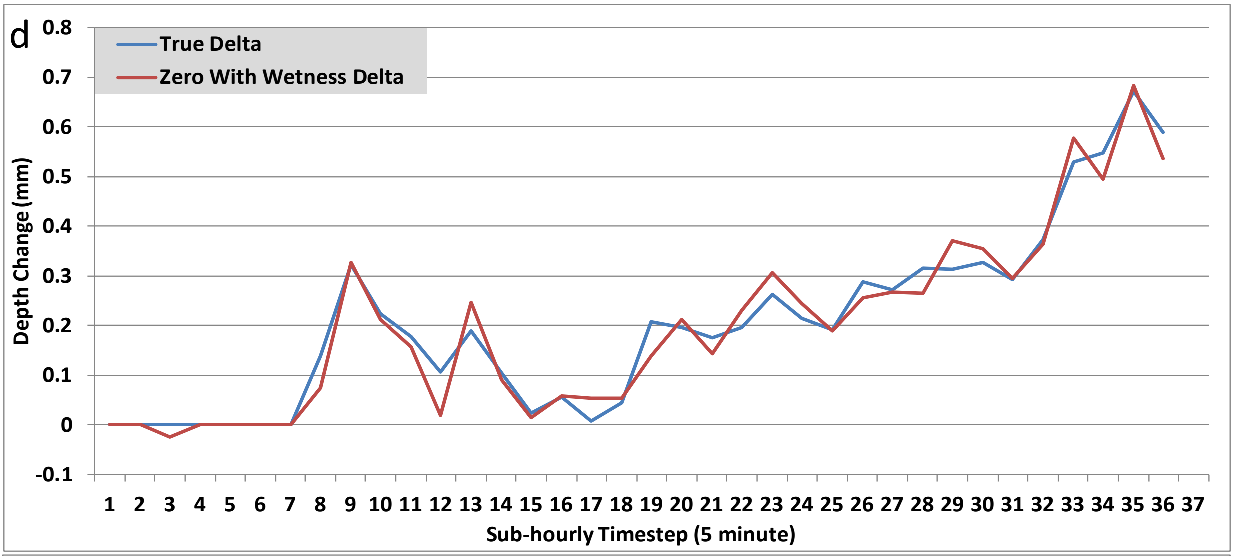
The delta variance is a relative measure of noisiness among the wires, with wire weights assigned inversely proportional to the delta variance such that noisier wires (more variance) are weighted less. Note that the delta variance is calculated using all three hours of data (or two if the first hour in the three-hour block is missing), but since precipitation is already known for the first two hours, calculations are only for the current hour. A minimum of two hours of depth data is required to calculate precipitation, so that any change in depth between the previous and current hour can be measured. The wire weights are normalized so that their sum is one. The normalized weight for each wire is then multiplied by the raw depth change value for each wire during a 5-minute period, and the components are added together to produce the weighted average depth change.

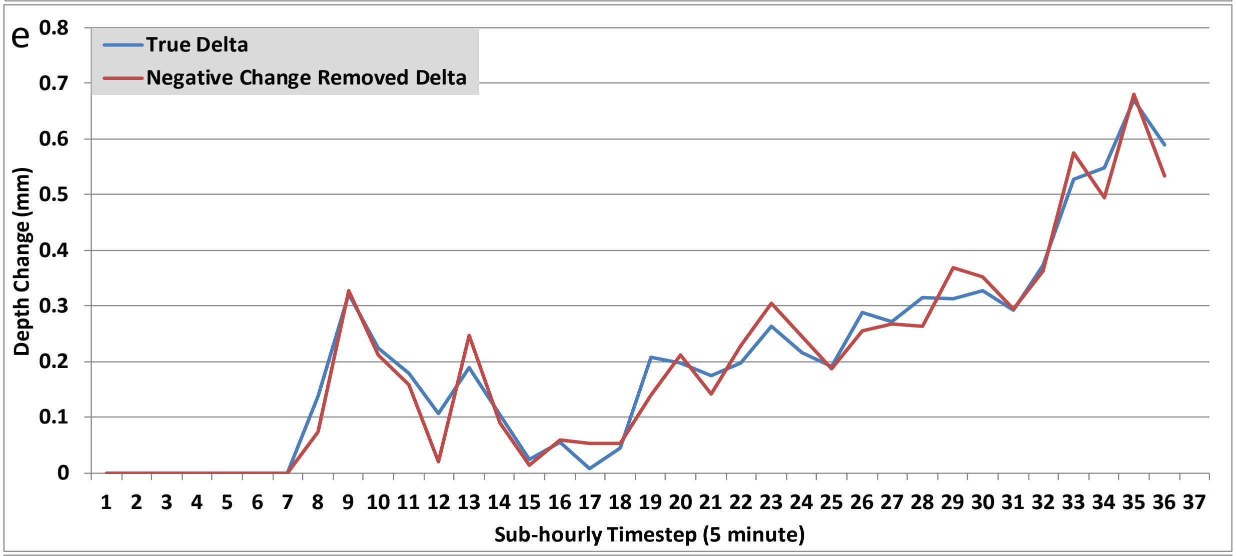
At this point, the 5-minute weighted average wire depth changes exist for the latest hour, but these are not yet calculated precipitation values. The wetness sensor information is then interrogated to see if precipitation falling from the sky was detected. If the wetness state is yes, the delta for that 5 minutes is retained; if the wetness sensor state is no, the delta is set to zero. If the wetness sensor data are flagged, a more complex approach requiring all three wire deltas to be simultaneously positive and within agreement to 0.5 mm is used to assure that noise is not considered precipitation in the absence of wetness sensor data. Deltas that fail this test are also set to zero. The wetness sensor performance is quite reliable, so this alternate branch is not often used.

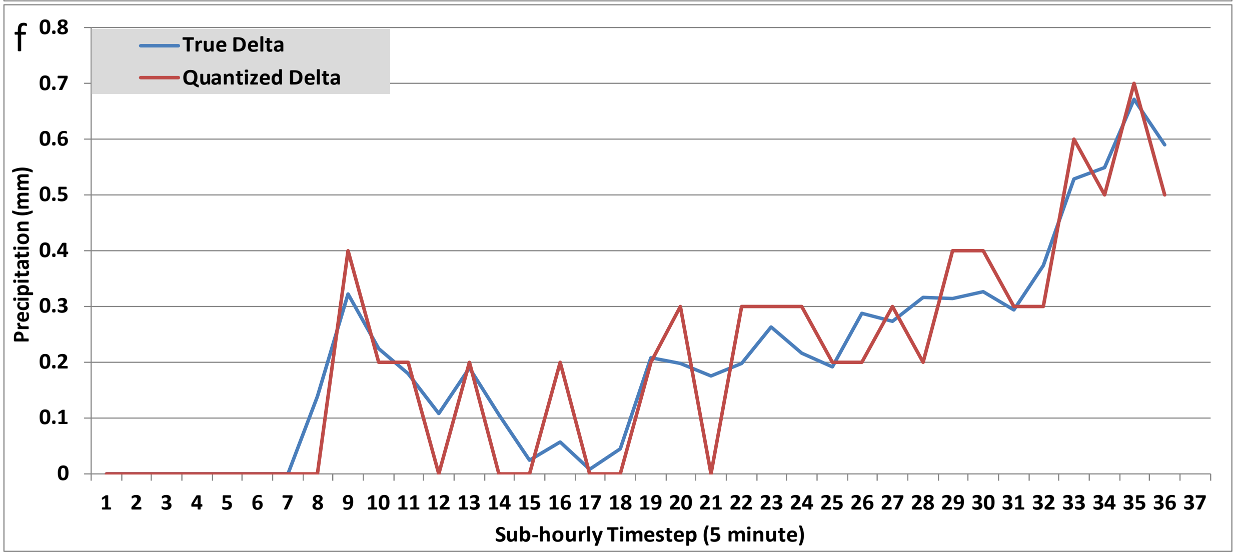
The remaining deltas are once more checked to see if any negative values (i.e., sensor noise) exist during periods when the wetness sensor indicates precipitation. If so, the particular negative delta is set to zero. However, if that is all that was done, removing negative wire depth noise without compensating for its removal would cause a positive bias in calculations due to positive noise being retained. Therefore, the total amount of negative precipitation is split among the other positive deltas and subtracted from them.

The final step is to quantize the existing delta values into calculated 5-minute precipitation amounts. As they are, the delta values have many places to the right of the decimal point, yet the precipitation must be given in tenths of millimeters, and must exceed 0.2 mm for any 5-minute period. If the first 5-minute period with precipitation has a positive delta of < 0.2 mm, that amount is added to the next 5-minute period with a positive delta, until the 0.2 mm threshold is exceeded and precipitation is reported as the number of tenths of millimeters accumulated up to that time. Following this, the residual is added to the next 5-minute period with a positive delta, and the procedure continues until the end of the most recent hour is reached. This approach also allows any residual from the previous hour to be transferred into the most recent hour so that precipitation is not lost in quantizing and rounding the precipitation values. Therefore, at the end of the procedure for a given hour, all the 5-minute calculated precipitation values are 0.2 mm or larger in tenths of millimeters, and any residual left at the end of the hour of less than 0.2 mm will eventually make its way into the next hour. These small residual amounts are important to account for in dry climates and/or very light precipitation events. This series of steps that describe the calculation of precipitation from OAP 2.0 are shown in Figure 3 for a typical precipitation event.


a
Wire 1; weighted 0.246
Wire 2; weighted 0.514
Wire 3; weighted 0.241
Accumulated Precipitation (mm)
Sub-hourly Timestep (5 minute)
b
Wire Delta 1
Wire Delta 2
Wire Delta 3
Depth Change (mm)
Sub-hourly Timestep (5 minute)
c
True Delta
Weighted Delta
Depth Change (mm)
Sub-hourly Timestep (5 minute)


d
True Delta
Zero With Wetness Delta
Depth Change (mm)
Sub-hourly Timestep (5 minute)


e
True Delta
Negative Change Removed Delta
Depth Change (mm)
Sub-hourly Timestep (5 minute)


f
True Delta
Quantized Delta
Precipitation (mm)
Sub-hourly Timestep (5 minute)

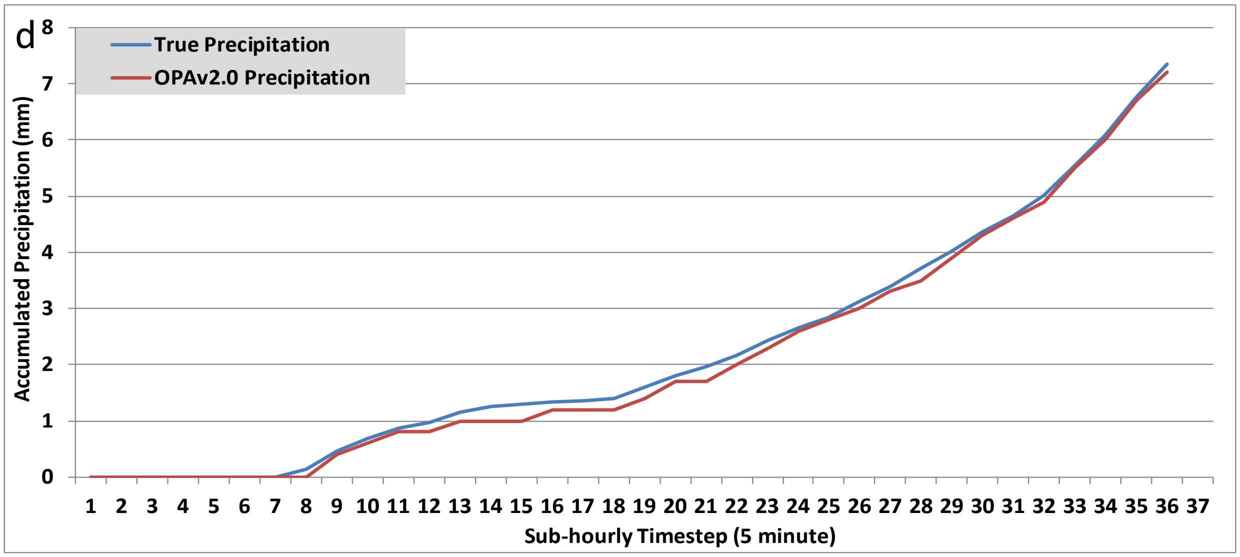



Figure 3. A precipitation event subjected to the OAP 2.0 process: a) the raw wire depths; b) the raw wire deltas; c) raw precipitation (labelled true precip) displayed with the weighted average delta values derived by wire delta variance weights of 0.246, 0.514, and 0.241, respectively; d) values zeroed according to the wetness sensor; e) values with negatives set to zero and the amount removed from the deltas; f) the quantized and rounded values, the final calculated precipitation; g) accumulated raw and OAP 2.0 calculated precipitation values, leaving a residual of -0.15511 mm for consideration in the next hour.

## Testing

OAP 2.0 (wavgCalc) and the original calculation algorithm (pairCalc) were evaluated through comparisons using both station data and synthetic precipitation events; this analysis is detailed in Appendix 1. The synthetic events assess algorithm performance against a known outcome under conditions of varying rates of gauge evaporation and randomized wire noise. These results were further evaluated in a field campaign study conducted over the summer of 2013 investigating the impacts of gauge evaporation in Leeper and Kochendorfer (2015). The station comparisons offer an opportunity to examine relative differences in measured precipitation with algorithm choice. The results from Appendix 1 will be summarized here.

The synthetic testing involved four precipitation events, based on the overall nature of the precipitation rate: heavy, very light, zero, and constant (Table 2). For each base event, there were 15 scenarios with each having 3 gauge evaporation rates and 5 wire noise settings. To fully evaluate algorithm performance, each scenario was simulated 100 times to generate mean absolute error statistics. In general, the OAP 2.0 had lower measures of error compared to the old algorithm (Table 3).

Table 2. Description of synthetic precipitation events duration, total accumulation, and peak and average event intensity

| Artificial Events | Duration (hr) | Accumulation (mm) | Peak Intensity ($mmhr^{-1}$) | Average Intensity ($mmhr^{-1}$) |
|---|---|---|---|---|
| Heavy | 4.4 | 98.90 | 58.80 | 22.5 |
| Very Light | 4.5 | 0.89 | 0.04 | 0.19 |
| Non-Precipitating | 4.0 | 0.00 | 0.00 | 0.00 |
| Constant | 10.0 | 36.00 | 0.30 | 3.60 |

Application of the new OAP 2.0 to large segments of real USCRN and regional network station data demonstrated an increase in total precipitation of 0.5% or more at 87% of stations, with only 5% showing a decrease of 0.5% and 8% staying near their current values. By totaling precipitation from all stations in the 5-minute precipitation era, about 1.6% more precipitation was calculated using the OAP 2.0 approach. This result is in general agreement with undercatch seen for liquid precipitation between USCRN and Cooperative Observer Program Network stations that are closely co-located (Leeper et al. 2015b).

Examining more closely individual precipitation events at a subset of 42 USCRN stations, it is clear that the increase in precipitation calculated by the new OAP 2.0 approach holds true for almost all air temperature, wind, and precipitation intensity states (Figure 4). In the case of precipitation events separated by an hour of dry conditions as determined by both algorithms, the new OAP 2.0 approach produced a high percentage of events with more precipitation, about 64%. In cold conditions below -5°C, the new algorithm does as well as it does in the general case, but with fewer cases having similar total precipitation between the two algorithms. This is probably due to the wetness sensor working less efficiently at very cold air temperatures, when some hydrometeors bounce off the disdrometer rather than being melted and detected. Only in the case of strong winds >=7$ms^{-1}$ was it equally likely for both calculation algorithms to produce the most precipitation, again probably related to inefficiencies in precipitation detection by the disdrometer in these circumstances.

Table 3. The old algorithm (pairCalc) and AOP 2.0 (wavgCalc) one-hundred member ensemble MAE average (mm) for synthetic heavy, very light, non-precipitating, and constant rate events by various levels of gauge evaporation (0.00 – 0.02) and wire noise (000,111,113,133,and 333).

| Generated Events | QA Variants | Gauge Evaporation | Noise Level Per Wire | | | | |
|---|---|---|---|---|---|---|---|
| | | | 0 | 111 | 113 | 133 | 333 |
| Heavy Event | pairCalc | 0.00 | 0 | 0.06 | 0.11 | 0.16 | 0.24 |
| | | 0.01 | 0.06 | 0.06 | 0.09 | 0.15 | 0.24 |
| | | 0.02 | 0.13 | 0.10 | 0.11 | 0.16 | 0.24 |
| | wavgCalc | 0.00 | 0 | 0.06 | 0.08 | 0.13 | 0.18 |
| | | 0.01 | 0 | 0.07 | 0.08 | 0.13 | 0.19 |
| | | 0.02 | 0 | 0.07 | 0.08 | 0.13 | 0.19 |
| Very Light Event | pairCalc | 0.00 | 0.49 | 0.34 | 0.30 | 0.28 | 0.52 |
| | | 0.01 | 0.78 | 0.47 | 0.40 | 0.23 | 0.30 |
| | | 0.02 | 0.88 | 0.57 | 0.50 | 0.33 | 0.24 |
| | wavgCalc | 0.00 | 0.08 | 0.11 | 0.13 | 0.16 | 0.22 |
| | | 0.01 | 0.10 | 0.16 | 0.16 | 0.17 | 0.21 |
| | | 0.02 | 0.25 | 0.23 | 0.23 | 0.22 | 0.23 |
| Non-Precipitation Event | pairCalc | 0.00 | 0 | 0 | 0 | 0.07 | 0.19 |
| | | 0.01 | 0 | 0 | 0 | 0.03 | 0.08 |
| | | 0.02 | 0 | 0 | 0 | 0.02 | 0.03 |
| | wavgCalc | 0.00 | 0 | 0 | 0 | 0.01 | 0.03 |
| | | 0.01 | 0 | 0 | 0 | 0 | 0.02 |
| | | 0.02 | 0 | 0 | 0 | 0 | 0.01 |
| Constant Rate Event | pairCalc | 0.00 | 0 | 0.02 | 0.70 | 1.65 | 1.75 |
| | | 0.01 | 0.09 | 0.04 | 0.67 | 1.54 | 1.67 |
| | | 0.02 | 0.11 | 0.12 | 0.58 | 1.49 | 1.60 |
| | wavgCalc | 0.00 | 0 | 0.06 | 0.08 | 0.13 | 0.18 |
| | | 0.01 | 0 | 0.07 | 0.09 | 0.13 | 0.18 |
| | | 0.02 | 0.01 | 0.07 | 0.10 | 0.14 | 0.19 |

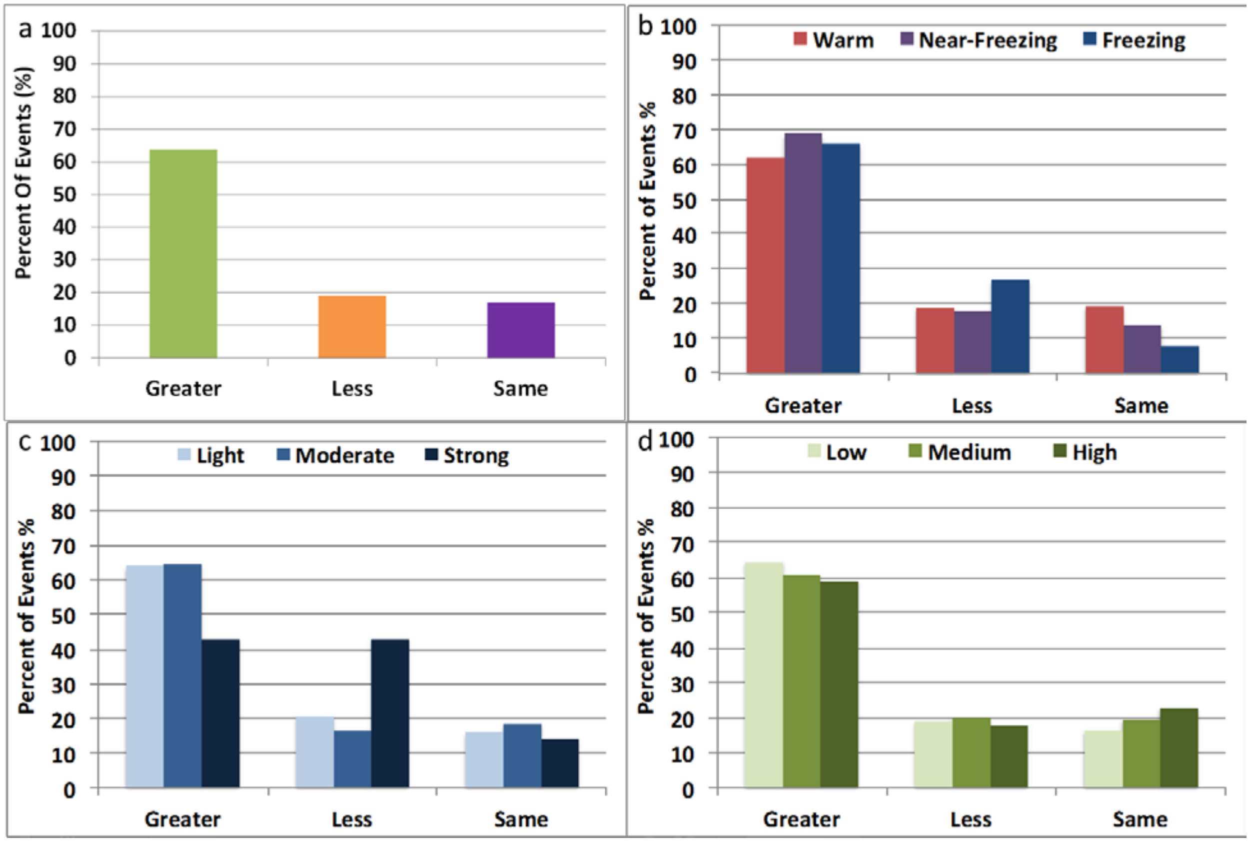



Fig 3. Percentage of precipitation events in which AOP 2.0 (wavgCalc) had greater (left), less (middle), and about the same (right) accumulations as the old algorithm (pairCalc) for: (a) all cases; (b) warm (avg. temperature > 5°C; red), near-freezing (avg. temperature < 5°C & > -5°C; purple), and freezing (avg. temperature <= -5°C; blue) temperature conditions; (c) light (avg. wind <= $2ms^{-1}$; light blue), moderate (avg. wind > 2 & < $7ms^{-1}$; medium blue), and strong (avg. wind >= $7ms^{-1}$; dark blue) wind conditions; and (d) low (avg. rate <= $0.5mmhr^{-1}$; light green), medium (avg. rate > 0.5 & < $2mmhr^{-1}$; medium green) and high (avg. rate >= $2mmhr^{-1}$; dark green) intensity conditions.

**Conclusion**

This document describes the process through which precipitation will be calculated by OAP 2.0, and summarizes the results of an examination of this approach compared to the existing method for precipitation calculation. The new method was shown to objectively have a lower measure of error in synthetic test scenarios with known outcomes, and resulted in 1.6% more precipitation from USCRN station observations overall, including an improved distribution of precipitation at the start of precipitation events. The development of the new algorithm has been successful; however, all input data will be retained for further improvements in the future.

## References

Baker, B. C., R. Buckner, W. Collins, and M. Phillips, 2005: Calculation of USCRN Precipitation from Geonor Weighing Precipitation Gauge. NOAA Technical Note, No. USCRN-05-1. (Available online at: http://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/program/technotes/TN05001GeonorAlgorithm.pdf).

Leeper, R. D. and J. Kochendorfer, 2015: Evaporation from weighing precipitation gauges: impacts on automated gauge measurements and quality assurance methods, Atmos. Meas. Tech., 8, 2291–2300, doi:10.5194/amt-8-2291-2015.

Leeper, R.D., M.A. Palecki, and E. Davis, 2015a: Methods to calculate precipitation from weighing bucket gauges with redundant depth measurements. *Journal of Atmospheric and Oceanic Technology*, **32**, 1179-1190, doi:10.1175/JTECH-D-14-00185.1.

Leeper, R.D., J. Rennie, and M.A. Palecki, 2015b: Observational perspectives from U.S. Climate Reference Network (USCRN) and Cooperative Observer Program (COOP) Network: Temperature and precipitation comparison. *Journal of Atmospheric and Oceanic Technology*, **32**, 703-721, doi:10.1175/JTECH-D-14-00172.1

# Appendix 1

## Methods to Calculate Precipitation From Weighing Bucket Gauges with Redundant Depth Measurements

Leeper, R.D., M.A. Palecki, and E. Davis, 2015a: Methods to calculate precipitation from weighing bucket gauges with redundant depth measurements. *Journal of Atmospheric and Oceanic Technology*, **35**, 1179-1190, doi:10.1175/JTECH-D-14-00185.1.

# Methods to Calculate Precipitation from Weighing-Bucket Gauges with Redundant Depth Measurements

RONALD D. LEEPER

*Cooperative Institute for Climate and Satellites–North Carolina, North Carolina State University, and NOAA/National Climatic Data Center, Asheville, North Carolina*

MICHAEL A. PALECKI

*NOAA/National Climatic Data Center, Asheville, North Carolina*

EGG DAVIS*

*NOAA/National Climatic Data Center, Asheville, North Carolina, and ERT Inc., Silver Spring, Maryland*

(Manuscript received 24 September 2014, in final form 31 January 2015)

ABSTRACT

The U.S. Climate Reference Network (USCRN) monitors precipitation using a well-shielded Geonor T-200B gauge. To ensure the quality and continuity of the data record, the USCRN adopted an innovative approach to monitor precipitation using redundant technology: three vibrating-wire load sensors measuring the liquid depth of a weighing-bucket gauge. In addition to detecting and flagging suboptimally operating sensors, quality assurance (QA) approaches also combine the redundant observations into a precipitation measurement. As an early adopter of this technology, USCRN has pioneered an effort to develop QA strategies for such precipitation systems.

The initial USCRN approach to calculating precipitation from redundant depth observations, pairwise calculation (pairCalc), was found to be sensitive to sensor noise and gauge evaporation. These findings led to the development of a new approach to calculating precipitation that minimized these nonprecipitation impacts using a weighted average calculation (wavgCalc). The two calculation approaches were evaluated using station data and simulated precipitation scenarios with a known signal. The new QA system had consistently lower measures of error for simulated precipitation events. Improved handling of sensor noise and gauge evaporation led to increases in network total precipitation of 1.6% on average. These results indicate the new calculation system will improve the quality of USCRN precipitation measurements, making them a more reliable reference dataset with the capacity to monitor the nation's precipitation trends (mean and extremes). In addition, this study provides valuable insight into the development and evaluation of QA systems, particularly for networks adopting redundant approaches to monitoring precipitation.

## 1. Introduction

Precipitation is a fundamental meteorological and climatological variable. Variations in precipitation patterns can disrupt agricultural productivity and in extreme cases foster drought or flood conditions. From 2011 to 2012, natural disasters related to extreme precipitation patterns (droughts and floods) cost the United States over 47 billion dollars (NOAA 2013). These extreme conditions contribute to secondary hazards that include mudslides, disease pandemics, heat waves, and forest fires, all of which further increase environmental impacts, property damage, and human casualties. Over climatological time scales, changing spatiotemporal patterns of precipitation can impact agricultural productivity (desertification/persistent floods), the availability and quality of water resources, and place further strain on already aging infrastructure (dams, levees,

* Current affiliation: Global Science and Technology, Inc., Greenbelt, Maryland.

*Corresponding author address:* Ronald D. Leeper, CICS-NC/NCDC, 151 Patton Avenue, Asheville, NC 28801.
E-mail: ronald.leeper@noaa.gov

DOI: 10.1175/JTECH-D-14-00185.1

bridges, etc.; Kunkel et al. 2013). High-quality in situ measurements of precipitation are fundamental to ensuring the quality (through validation) of radar and satellite precipitation estimates and quantitative precipitation forecasts. In an effort to improve U.S. precipitation monitoring, NOAA's National Climatic Data Center (NCDC) deployed the U.S. Climate Reference Network (USCRN) to support hydrological studies (flood and drought extremes) and to accurately monitor the nation's precipitation trends over climatological time scales.

Observing ground-based precipitation accurately is a challenging task (Rasmussen et al. 2012). Few observing networks are designed adequately to monitor the variety of hydrometeor types while mitigating the adverse effects of surface winds, sensor noise, gauge evaporation, and other sources of observation biases on precipitation measurements (Goodison et al. 1998; Rasmussen et al. 2012; You et al. 2007). In addition, quality assurance (QA) methods used to ensure the validity of detected precipitation signals and to mitigate these biases are often not well documented (Fiebrich et al. 2010; Shafer et al. 2000). To address these challenges, the USCRN adopted an innovative approach to monitor precipitation redundantly from a well-shielded enclosure (Fig. 1a).

The USCRN uses the all-weather Geonor T-200B weighing precipitation gauge to monitor precipitation redundantly. The gauge is configured such that the reservoir (weighing bucket in Fig. 1b) is suspended from three independent load sensors. The load cells observe reservoir weight by magnetically plucking an internal wire and monitoring its frequency of vibration, which will vary with wire tension as described by Duchon (2008). When calibrated, the Geonor gauge can reliably detect changes in gauge depth to a resolution of 0.1 mm. The redundant sensors improve the resilience of the precipitation-observing system against single-sensor degradation and failure, as the additional sensors continue to monitor precipitation until the gauge is repaired. The redundant monitoring also ensures the continuity of the data record, which is necessary to effectively monitor climate (Diamond et al. 2013; National Research Council 1999; Trenberth et al. 2002), in addition to providing information that can be used to improve the detection of a precipitation signal from gauge noise. For instance, intercomparisons of the redundant sensor depth change can be used to determine if an increase in gauge depth is the result of random noise detected by a single sensor or precipitation equally observed by all three sensors.

The USCRN also endeavors to reduce environmental issues that can negatively impact precipitation measurements. To inhibit frozen precipitation from collecting on the interior walls and capping the gauge, a heating tape is applied to the Geonor throat for stations located in colder climates. A better-known bias often discussed in the context of precipitation measurement is wind errors, which can bias measurements as much as 50% (Sevruk et al. 2009). The USCRN lessens the impacts of surface winds by observing precipitation from a well-shielded enclosure. The majority of USCRN gauges are surrounded by a small double fence intercomparison reference shield (SDFIR) with an interior single-alter shield (Fig. 1b). In locations where siting and/or material transport become an issue (mostly in Alaska), the gauges are shielded by a double-alter shield. These shielding arrangements were found to reduce wind-related errors in sensitivity tests with the all-weather Geonor T-200B precipitation gauge (Baker et al. 2005b). Finally, each station is also equipped with a Hydrological Services tipping-bucket rain gauge model TB-3 for added redundancy during liquid precipitation; however, real-time QA processes do not currently use tipping-bucket data.

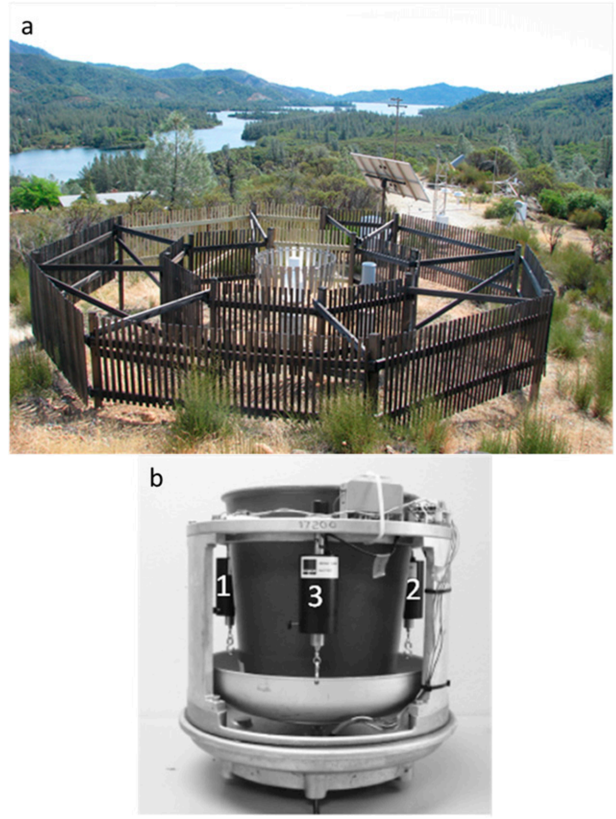



FIG. 1. Photographs of (a) USCRN Geonor-T-200B gauge within both a SDFIR shield and an alter shield near Merced, CA; and (b) a view of three vibrating-wire load sensors (redundant technology) used to monitor gauge depth.

While clearly beneficial, redundant measurements of gauge depth increase the complexity of QA systems by requiring both traditional quality control (QC) checks on raw gauge depths and a computational algorithm to compute an observation quantity from the redundant measures of gauge depth. In this study, the traditional definition of a QA system is extended to include this additional calculation. As an early adopter of redundant technology, the USCRN has pioneered the development of QA systems that process redundant measurements to enhance the quality of observations (temperature and precipitation) and continuity of the data record. The current QA system features a pairwise calculation (pairCalc) of depth changes that has evolved over time as a disdrometer (used to detect atmospheric wetness) and higher time-resolution (5 min) observations were brought online during the earlier period of the network's history, which is briefly described by Baker et al. (2005a). As operations stabilized, internal evaluation of USCRN precipitation measurements revealed pairCalc may have sensitivities to sensor noise and gauge evaporation. Comparing closely spaced members of the USCRN and the Cooperative Observer Program (COOP) network, Leeper et al. (2015) speculated that precipitation differences between the two networks were partially due to gauge attributes, such as wetting factor, sensor noise, and gauge evaporation, that at times adversely affected the performance of pairCalc. These issues in addition to COOP observer biases (Daly et al. 2007; Holder et al. 2006; Fiebrich and Crawford 2009) led to a small underreporting of USCRN precipitation compared to COOP.

In response, a new QA system has been developed based on a weighted average calculation (wavgCalc) that better utilizes the redundant information from the three load cells (wires) to mitigate the impact of both sensor noise and gauge evaporation on precipitation measurements. The purpose of this study is to outline the QA techniques USCRN has explored to process redundantly monitored gauge data and to document the performance of these methods using field data and synthetically generated precipitation events. The outcome of this comparison study not only validates the USCRN QA approach but also provides valuable insight into development and evaluation strategies for precipitation QA systems in general. This is particularly true for QA specialists of other networks considering the adoption of redundant observation systems as this approach to monitoring precipitation becomes more widespread.

## 2. Calculation methods for redundant systems

The two calculation methods, pairCalc and wavgCalc, were designed for the same USCRN precipitation system, which is configured to report raw gauge depths (1-min average of thirty 2-s samples) from the redundant load sensors every 5 min. The two approaches apply an identical set of QA checks on the raw gauge depth values to ensure data quality. These QA tests include a range check (each load sensor separately) that at the upper limit ensures gauge depths are within the operational capacity of the gauge (600 or 1000 mm) and that at the lower limit validates sensor health; failed sensors report negative depths (Fig. 2a). To limit false reports of precipitation that may arise from sensor noise (i.e., wind loading, electrical issues, temperature dependencies, among others; Figs. 2b and 2c), a detection threshold of 0.2 mm is applied. However, changes in gauge depth as small as 0.1 mm are discernable beyond this threshold (i.e., 0.3 mm is detectable). To handle instances of gauge maintenance, animal infestations, and/or electrical issues that result in large synchronous (among the three wires) increases in gauge depth (Fig. 2d), an upper threshold on depth change of 25 mm is enforced. For additional cases when these QA checks fail to prevent false precipitation, an independent disdrometer is also used to determine the presence of precipitation (wetness). If no wetness is observed, then any reported increases in gauge depth are not included in the precipitation calculations. As noted previously, these QA checks are applied to both computational algorithms.

While these methods have similar raw gauge data QA checks, the main distinction between them is how depth changes are computed and redundant observations are merged into a single precipitation measurement. Fundamentally, the current method relies on pairwise agreement of depth changes, using redundancy as a double or triple check on the measurement. This is the approach used by the USCRN program in calculating air temperature from redundant measurements (Palecki and Groisman 2011). However, air temperature measurements experience much less noise among the redundant sensors than that which exists with gauge measurements, which tend to have both diurnal and nonsystematic noise signals. The second approach pools available information from the redundant depth measurements to identify the most reasonable precipitation signal by giving greater weight to less noisy measurements. The following sections briefly describe the existing (pairCalc) and recently developed (wavgCalc) QA systems.

### a. PairCalc

PairCalc requires gauge depths from the preceding 2 h and the current hour (a total of 3 h) to evaluate subhourly depth changes. Subhourly depth change is computed for each wire separately over the current hour.

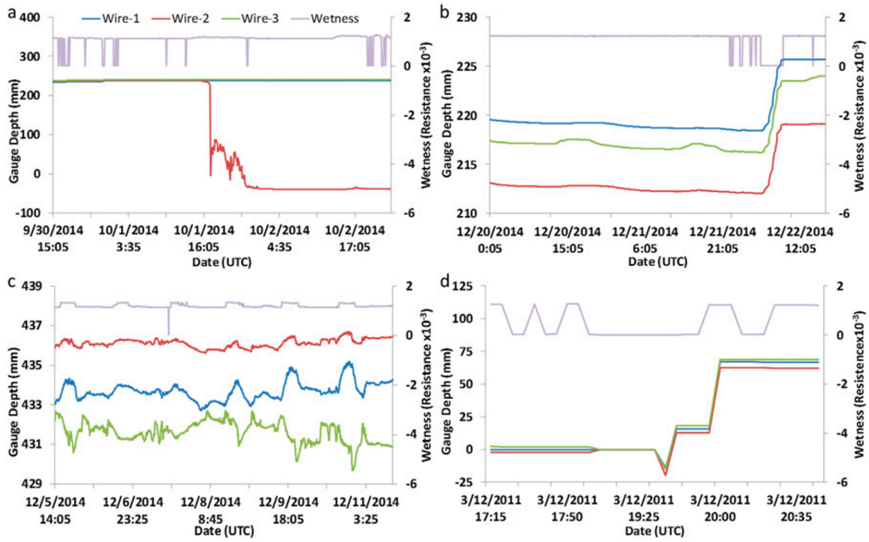


FIG. 2. Geonor gauge depths for wire1 (blue), wire2 (red), and wire3 (green), and wetness sensor (purple) data where observed resistance less (greater) than $1 \times 10^{-3}$ indicates atmospheric wetness (dryness) for a (a) failed wire scenario at Jamestown, ND; (b) diurnal noise pattern (most visible in wire3) embedded within an evaporation signal at Titusville, FL; (c) random noise embedded in diurnal variations at Sundance, WY; and (d) gauge maintenance event where an antifreeze mixture was added to winterize the gauge at Bowling Green, KY.

Depth changes are computed by differencing the current depth with a reference depth determined from the previous (2 h) depth measurements. Reference depths are derived in one of two ways. If precipitation was observed within the last 2 h, then the reference depth is set to the gauge depth when precipitation was last recorded; otherwise, the reference depth is an average of all wire depths over the previous 2 h. Each wire reference depth is then deducted from the current wire depth to quantify depth changes (wire deltas). Wire deltas are then compared in pairwise fashion (wire1–wire2, wire1–wire3, and wire3–wire2) as a consistency check to identify and remove poorly behaved (e.g., noisy and broken) wires. There are four possible outcomes from the pairwise comparison:

1) All three pairwise differences are less than 0.2 mm (all three wires pass).
2) A single pairwise difference is less than 0.2 mm (both wires in that pair pass).
3) Two pairwise differences are less than 0.2 mm (only the wire common to both pairs passes).
4) No pairwise difference is less than 0.2 mm (invoke the storm clause).

In case 4, intense precipitation can cause water within the gauge to occasionally slosh randomly, altering the load on each wire, resulting in failed pairwise agreement checks. In these cases, agreement is relaxed, allowing larger pairwise differences among the redundant measurements (up to 20% of the three-wire delta mean). Deltas that pass the consistency check are used to determine depth change from an arithmetic mean. For instance, if the three sensors (in order) reported a 5-min depth change of 3.2, 4.0, and 3.8 mm, all three wires would initially fail the pairwise check (none agrees to within 0.2 mm). The storm clause would relax the pairwise agreement to 0.7 mm for this subhourly period (mean delta of 3.6 mm $\times$ 20%). Using the new agreement threshold, two pairwise differences now pass this check (less than 0.7 mm) with the third wire common to both checks used to determine total precipitation (based on pairwise outcome case 3) of 3.8 mm. A flowchart describing this process has been provided in Fig. 3 with additional information for this algorithm provided in Baker et al. (2005a).

### b. WavgCalc

Similar to pairCalc, wavgCalc computes precipitation from the same 3-h window (previous 2 h plus current hour). However, wire deltas are computed as a change in gauge depth between successive subhourly periods in a manner similar to the way manual evaluations might be performed (current subhourly depth minus previous subhourly depth). Wire deltas are then averaged using a weighted mean. Wire weights are determined based on the average delta variance of each wire over the 3-h period. The delta variance is calculated for each of the three wires, $k = 1, 2, 3$ [Eq.(1)], where $X$ is delta, $i$ is the

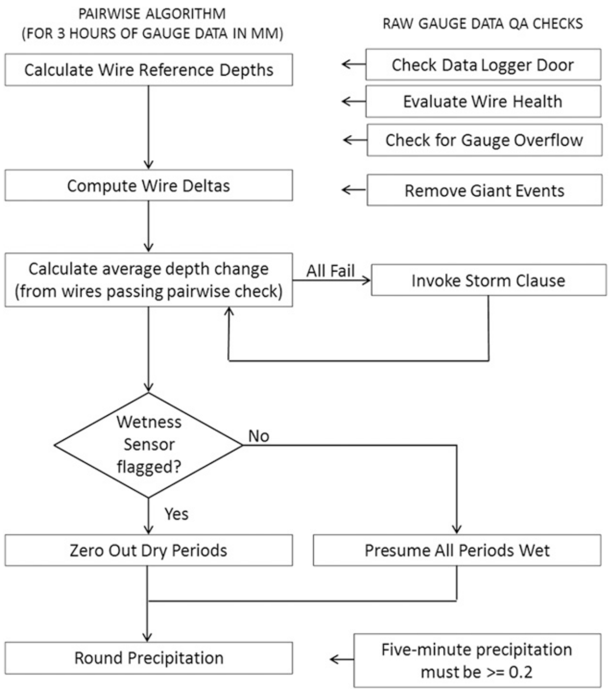



FIG. 3. Flowchart detailing the order of pairCalc procedures for calculating precipitation.

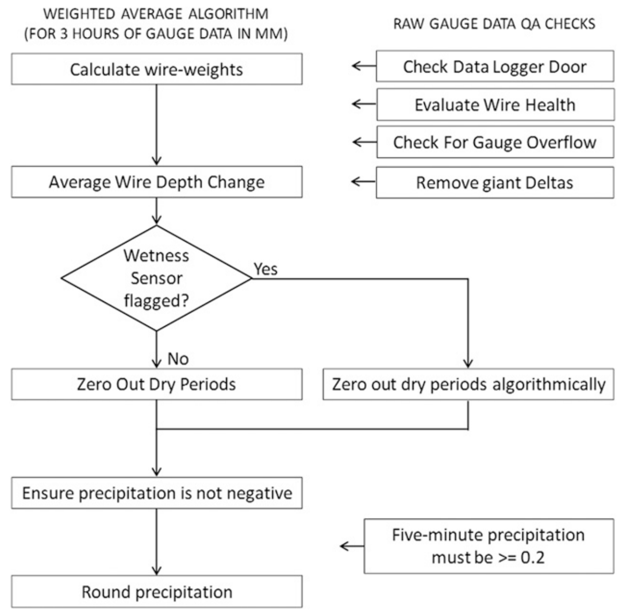



FIG. 4. Flowchart detailing the order of wavgCalc procedures for calculating precipitation.

$i$th subhourly period, $\overline{X}$ is the three-wire delta mean, and $n$ is the total number of subhourly periods:

$$\text{delta variance}_k = \frac{\sum_i^n (X_{ik} - \overline{X}_i)^2}{n}. \tag{1}$$

The delta variance is a relative measure of noisiness among the wires, with wire weights assigned inversely proportional to the delta variance such that noisier wires (more variance) are weighted less. A flowchart of this algorithm is provided in Fig. 4.

## 3. Methodology

Precipitation measurements from the two QA systems were evaluated against station data to compare relative differences and then by synthetically generated precipitation events to quantify QA performance against a known precipitation signal. The generated precipitation scenarios were designed to include sensor noise and gauge evaporation signals to evaluate the QA systems' sensitivity to these processes. Initially, QA calculations of precipitation were compared using all USCRN and U.S. Regional Climate Reference Network (USRCRN) stations. A more thorough investigation based solely on a USCRN subset of 42 stations (Fig. 5) was designed to explore how environmental conditions such as temperature, wind speed, and precipitation intensity impacted total precipitation. These analyses were conducted over the period of record, where observations were taken at a 5-min frequency (2006–07 for most stations to 2012).

Evaluations of method performance were carried out with synthetic precipitation events using a precipitation generator. Precipitation scenarios of known subhourly intensity and total accumulation were used to initialize the generator (Fig. 6a). The generator produces synthetic gauge data (depths from each wire and wetness) that match the precipitation scenario and can be processed through both QC algorithms. This approach allows the two methods to be evaluated against a known precipitation signal in much the same way a ''true'' dataset is used. Additionally, the generator has the capacity to embed defined levels of sensor noise (Fig. 6a) and gauge evaporation (Fig. 6b) for each redundant wire separately as is observed in the field. The magnitude of noise variations is randomly generated based on user-defined range specifications using a constrained random walk that limits the number of steps that can move away from the actual precipitation value. Gauge evaporation and sensor noise are two of the most important physical processes that QA systems mitigate to reduce measurement uncertainty. Generated data are then processed through each of the two QA methods for quantitative comparisons with respect to the known artificial signal.

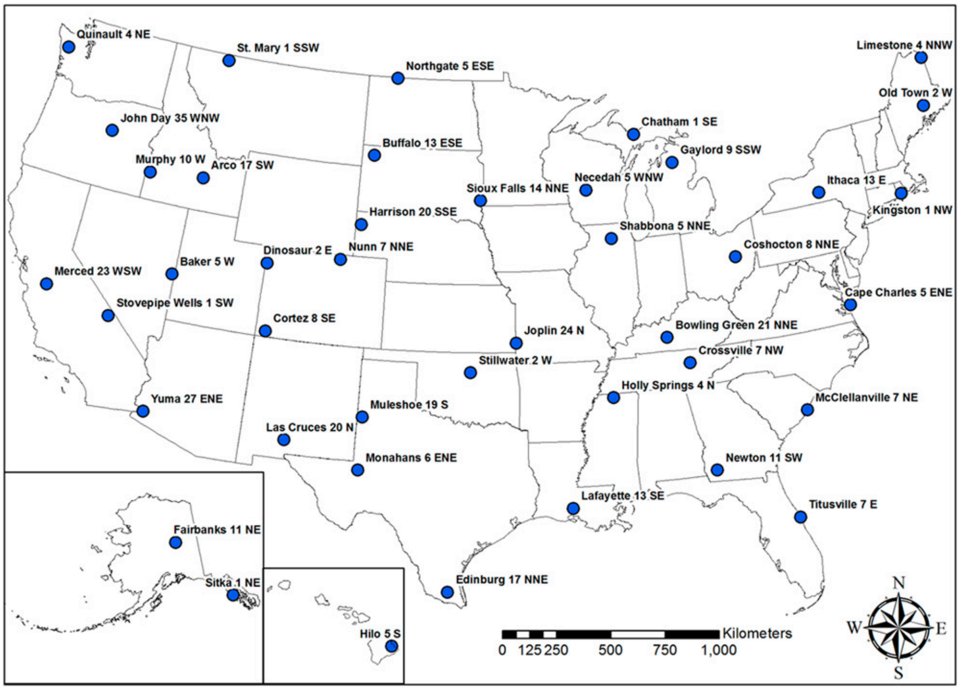



FIG. 5. Location of USCRN stations used in the subset analysis.

Given that generator-embedded signals are applied randomly for each 5-min time step, an ensemble of simulations for each artificial case was used to evaluate QA performance. In this study, synthetic cases were examined using 100-member ensembles for each gauge evaporation (0.00, 0.01, or 0.02 mm per 5 min) and wire noise (000, 111, 113, 133, and 333 in tenths of millimeters for each wire) combination. For instance, a scenario with an evaporation level of 0.02 mm and a 113 noise setting would have a randomly generated loss (from evaporation) between 0 and 0.02 mm for each wire in addition to variations in depth for wire1 and wire2 of ±0.05 mm (total range equaling 1.0) and a higher range (±0.15 mm) for wire3 due to noise. Each simulated case has a total of 1500 simulations (100 per scenario times 15 scenarios). The performance of QA calculations was evaluated using ensemble mean absolute error (MAE) as described by Legates and McCabe (1999), where $P_s$ is the known precipitation signal, $P_i$ is method-generated precipitation for the $i$th simulation, and $n$ is the number of simulations [Eq. (2)]:

$$\mathrm{MAE} = \frac{\sum_i^n \mathrm{abs}(P_s - P_i)}{n} \tag{2}$$

Four synthetic precipitation scenarios were included in this study that ranged from low to high precipitation intensities (Table 1). These event scenarios are described in section 4b.

## 4. Results

### a. Station observations

#### 1) All stations

The wavgCalc method calculated 1.6% more total precipitation than pairCalc on average over the combined networks (USCRN and USRCRN). The increase in reported precipitation was consistent across individual stations in the network (Table 2) with more than 87% of stations having an increase in accumulated precipitation of at least 0.5%. Those stations having a reduction in total precipitation by at least .5% represented less than 5% of the network. The increase in reported precipitation by wavgCalc relative to pairCalc was also consistent across annual and monthly time scales (Figs. 7a and 7b). Seasonally, QA system differences were slightly larger (>1.8%) from late winter to early spring (Fig. 7b). Seasonal trends in precipitation differences may be attributed to the performance of the auxiliary disdrometer used to detect falling precipitation by both methods. Tabler (1998) found the sensor type used by USCRN may fail to detect precipitation in colder conditions, as frozen hydrometeors can strike

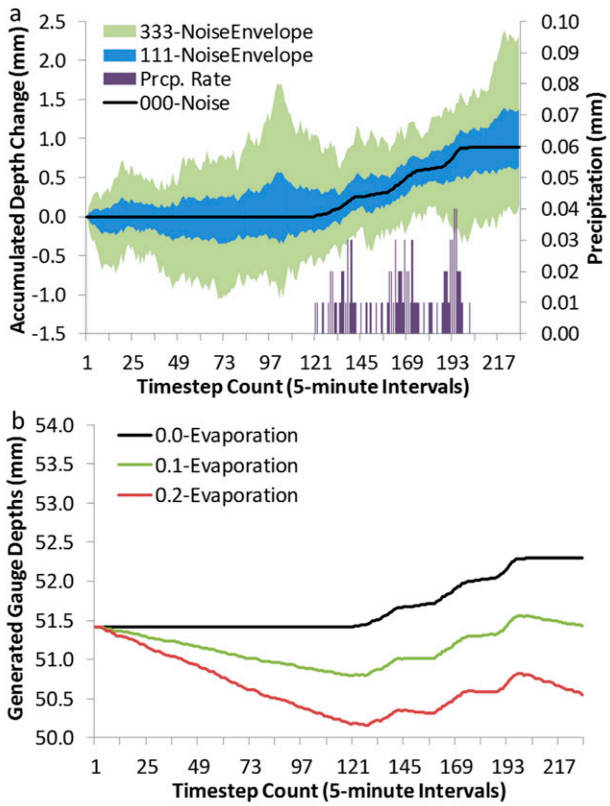



FIG. 6. Generator-produced (a) range of ensemble-accumulated depth change for noise levels 000 (black), 111 (blue), and 333 (green) for the very light precipitation scenario (purple). (b) Gauge depths from a single ensemble member with 0.0 mm (black), 0.1 mm (green), and 0.2 mm (red) evaporation settings.

the sensing plate and be bounced or be blown off before detection.

### 2) STATION SUBSET

A more detailed analysis using a subset of 42 USCRN stations shown in Fig. 5 was performed to evaluate calculation method differences with respect to surface conditions (air temperature, wind, and intensity). Periods of continuous precipitation (referred to as precipitation events) were defined as the time interval between the first and last hour both algorithms reported zero precipitation. Precipitation events were categorized by wavgCalc total precipitation and placed into greater than, less than, or equal to pairCalc bins.

From the 42 stations, 31 825 precipitation events were identified. The wavgCalc method observed more (less) precipitation than pairCalc for 63.9% (19.1%) of events with both QA systems reporting the same precipitation (within a tenth of a millimeter) for the other 17% (Fig. 8a). Distinguishing periods of precipitation between warm (average temperature greater than or equal to 5°C), near-freezing (between 5° and −5°C), and freezing (less than or equal to −5°C) conditions revealed the two calculation methods were more dissimilar in colder conditions as noted previously (Fig. 8b). The percent of precipitation events where both algorithms had the same accumulation diminished from a high of 19.3% for warm events to a low of 7.7% of events during freezing conditions. In addition, the percent of events in which wavgCalc was "greater than" (from 62.0% to 65.7%) and "less than" (from 18.7% to 26.5%) pairCalc both increased from warm to freezing conditions. The increase of precipitation dissimilarities between the two QA systems may be linked to the ineffectiveness of some collocated disdrometers during cold, snowy conditions. Failure to detect wetness when (within the 5-min window) increases in gauge depth occur seems to result in dissimilar QA responses that are sensitive to the way depth change is evaluated.

QA differences were also examined by surface wind speed and precipitation intensity. The percentage of events, during which wavgCalc observed more precipitation, dropped from 63.7% to 42.9% as winds speeds increased from light (<2 m s$^{-1}$) to strong (7 m s$^{-1}$) conditions (Fig. 8c). Similar to the temperature categories, the detection of wetness may be less reliable during windier surface conditions, where hydrometeors, if light enough, can be swept past the disdrometer and drive up QA differences. Precipitation intensity revealed the two QA methods were more similar during high- than low-intensity events (Fig. 8d). In addition, the wavgCalc method tended to report more precipitation during light-precipitation-rate events when sources of measurement error (sensor noise, gauge evaporation) make up a greater percentage of total precipitation.

Case studies of individual events revealed some additional insight into QA differences. For instance, pairCalc missed a precipitation event detected by both wavgCalc and a collocated tipping bucket at Yuma, Arizona (Fig. 9a), as a result of overly stringent wire agreement checks (failed pairwise check). Conversely,

TABLE 1. Description of synthetic precipitation events' duration, total accumulation, and peak and average event intensity.

| Artificial events | Duration (h) | Accumulation (mm) | Peak intensity (mm h$^{-1}$) | Avg intensity (mm h$^{-1}$) |
|---|---|---|---|---|
| Heavy | 4.4 | 98.90 | 58.80 | 22.5 |
| Very light | 4.5 | 0.89 | 0.04 | 0.19 |
| Nonprecipitating | 4.0 | 0.00 | 0.00 | 0.00 |
| Constant rate | 10.0 | 36.00 | 0.30 | 3.60 |

TABLE 2. Count and percent of USCRN and USRCRN stations that experienced a net reduction, no change, or increase in total precipitation.

| Difference | Difference criteria (%) | Station count | Stations (%) |
|---|---|---|---|
| Reduction | $<-0.5$ | 11 | 4.9 |
| Same | $\geq-0.5$ and $\leq0.5$ | 18 | 7.9 |
| Increase | $>0.5$ | 198 | 87.2 |

poor estimates of reference depth due to sensor noise led to a presumably false precipitation event reported by pairCalc at Joplin, Missouri (Fig. 9b) that neither the wavgCalc system nor the tipping bucket observed. In other cases, suspicious disdrometer behavior was found to result in sizable differences between the two methods (Figs. 9c and 9d), as noted previously during cold conditions. In these instances, an intermittent wetness signal (frequent change in sensor resistance) was observed, which generally resulted in lower wavgCalc precipitation totals compared to pairCalc despite both using the same sensor. For wavgCalc, increases in gauge depth during periods when wetness was *not* observed (Figs. 9c and 9d) are excluded from reported precipitation. While this is also true for pairCalc, this method was capable of capturing these depth increases later in time because depth change is computed with respect to a moving 2-h average (reference depth). By averaging previous depth changes, pairCalc has a limited memory of earlier depth increases when evaluating the current hour. This can negatively impact subhourly precipitation rates. Figure 9e illustrates this point. Stringent pairwise checks initially caused pairCalc to miss the initiation of a precipitation event, as observed by wavgCalc. However, pairCalc was able to recapture a portion of missed precipitation 2 h later when over 23 mm of precipitation were recorded in a single 5-min period at Durham, North Carolina. During that same subhourly period, wavgCalc reported 0.5 mm, which was in better agreement with observed depth change. Not only did pairCalc underreport total precipitation but also poorly distributed reported precipitation over time, affecting precipitation rates. While the intermittent wetness signal was usually observed during colder conditions, it should be noted that not all snowy periods had an intermittent wetness signal, as shown in a January precipitation event in Fairbanks, Alaska (Fig. 9f).

### *b. Synthetic precipitation scenarios*

Precipitation scenarios used to evaluate QA performance are listed in Table 1. In each of these events, levels of wire noise and gauge evaporation were allowed to vary between ranges observed in the field. The first case represents a heavy precipitation scenario that was 4.4 h in duration with an average and peak precipitation rate of 22.3 and 58.8 mm $h^{-1}$, respectively, for a total accumulation of 98.9 mm. Second, the very light event had an average intensity of 0.19 mm $h^{-1}$ over a 4.5-h period with a total accumulation of 0.89 mm. The non-precipitation event tested the methods against several hours of zero precipitation to evaluate their tendency to report false precipitation or type II errors. The final case was a constant-precipitation event that lasted 10 h with precipitation falling at a constant rate of 0.3 mm per 5-min period, for a total of 36 mm.

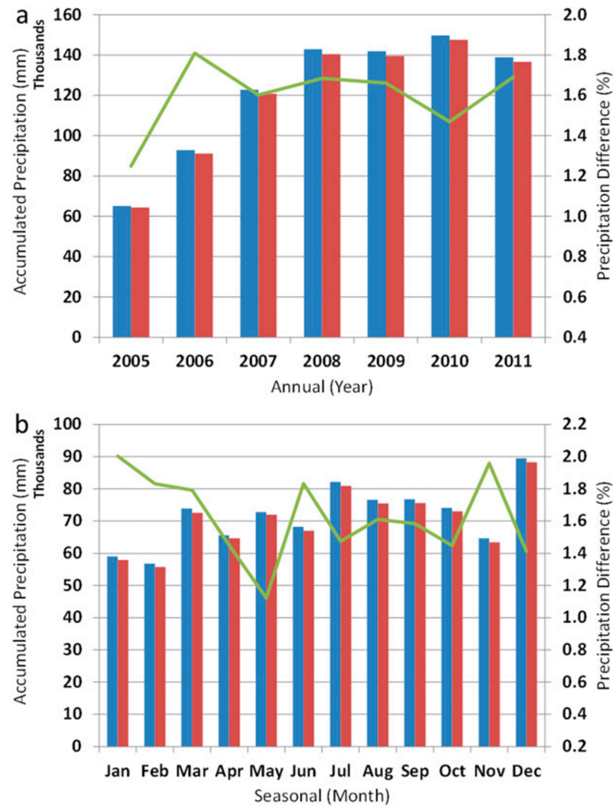



FIG. 7. USCRN total precipitation computed from wavgCalc (blue) and pairCalc (red) with percent differences (green) over (a) annual and (b) monthly time scales.

#### 1) HEAVY PRECIPITATION

For the no-evaporation and no-wire-noise case, pairCalc and wavgCalc were both error free with a MAE of zero for the heavy precipitation event (Table 3). As wire noise levels increased, both algorithms displayed higher levels of error (MAE). Of the two methods, pairCalc was more sensitive to elevated noise levels with an ensemble MAE range of 0–0.24 mm. Ensemble MAEs for wavgCalc were generally less, ranging from 0 to 0.19 mm for the same set of noise levels. Ensemble MAEs for pairCalc and to a lesser extent wavgCalc were found to reduce slightly with elevated evaporation signals. These

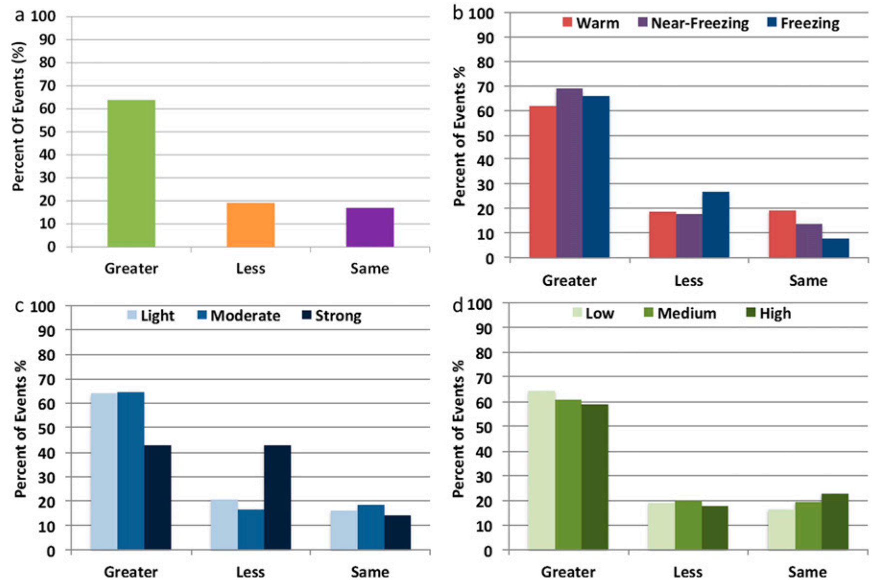



FIG. 8. Percentage of precipitation events in which wavgCalc had greater (green), less (orange), and the same (purple) accumulations as pairCalc for (a) all cases; (b) warm (avg temperature $\geq$ 5°C; red), near-freezing (avg temperature $<$ 5°C and $>$ $-$5°C; purple), and freezing (avg temperature $\leq$ $-$5°C; blue) temperature conditions; (c) light (avg wind $\leq 2\,\mathrm{m\,s^{-1}}$; light blue), moderate (avg wind $>$ 2 and $< 7\,\mathrm{m\,s^{-1}}$; medium blue), and strong (avg wind $\geq 7\,\mathrm{m\,s^{-1}}$; dark blue) wind conditions; and (d) low (avg rate $\leq 0.5\,\mathrm{mm\,h^{-1}}$; light green), medium (avg rate $>$ 0.5 and $< 2\,\mathrm{mm\,h^{-1}}$; medium green), and high (avg rate $\geq 2\,\mathrm{mm\,h^{-1}}$; dark green) intensity conditions.

results indicate that the positive errors (false precipitation) effect of wire noise may have been countered by the negative errors (missed precipitation) resulting from gauge evaporation within some of the simulations, resulting in overall lower ensemble MAE mean. Regardless, the wavgCalc system was less sensitive to evaporation and wire noise compared to pairCalc.

#### 2) VERY LIGHT PRECIPITATION

The lighter precipitation signal in this scenario resulted in higher MAEs relative to the heavy event, particularly as a percentage of total precipitation (Table 3). However, wavgCalc ensemble errors were still consistently lower than pairCalc. For the no-evaporation and no-wire-noise case, wavgCalc had a substantially lower MAE (0.08 mm) compared to pairCalc (0.49 mm, or more than 50% of total precipitation). PairCalc ensemble errors for the levels of noise and gauge evaporation ranged from 0.23 to 0.88 mm, or from 25% to 99% of event total precipitation. The range of MAEs for wavgCalc was much lower, between 0.08 and 0.23 mm, or 9%–25% of total precipitation. Higher MAEs were expected for lower-precipitation-rate events, as errors due to gauge noise and evaporation signals make up a greater percentage of total precipitation (lower signal-to-noise ratios), as shown in Figs. 6a and 6b.

#### 3) NONPRECIPITATING AND CONSTANT-PRECIPITATING CASES

In a similar manner, wavgCalc ensemble MAEs for both nonprecipitating and constant-precipitating scenarios were generally less than those for pairCalc (Table 3). When no gauge evaporation and no wire noise were included, both QA methods successfully reported no precipitation for the nonprecipitation case (MAEs of zero). However, this was not true, as noise levels were elevated. Ensemble MAEs for pairCalc and wavgCalc ranged between 0.00 and 0.19 mm and 0.00 and 0.03 mm, respectively. The smaller error suggests the wavgCalc system had fewer ensemble members reporting false precipitation based on noise than pairCalc. For the constant-precipitating case, the two calculation methods performed well (minimum error) when no evaporation and no wire noise were applied. However, when wire noise and gauge evaporation were included, the pairCalc MAEs ranged from 0 to 1.75 mm. Once again, wavgCalc MAEs were much lower, ranging between 0 and 0.19 mm.

## 5. Discussion and conclusions

Two distinct approaches to combining redundant gauge depth observations into a precipitation measurement were compared. Despite both methods using the

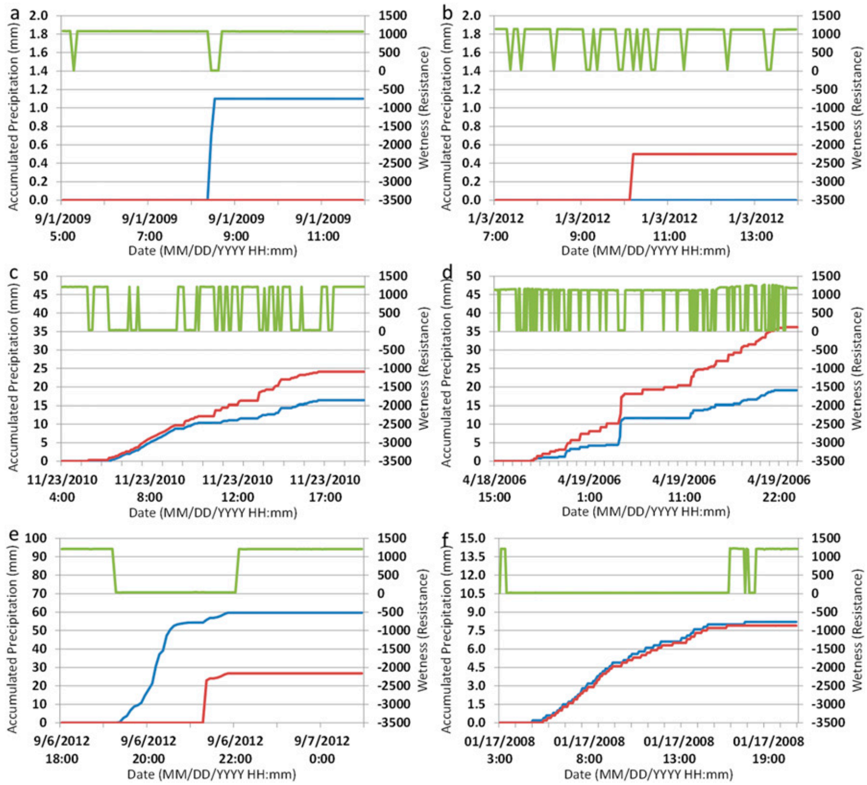


FIG. 9. Accumulated precipitation calculated using wavgCalc (blue) and pairCalc (red) systems with wetness (green; zero equal wet) for events in (a) Yuma, AZ; (b) Joplin, MO; (c) Arco, ID; (d) Buffalo, SD; (e) Durham, NC; and (f) Fairbanks, AK.

same set of QC checks on raw data, the manner in which redundant measures were combined had important impacts on reported precipitation with more than 90% of the network having some change in precipitation of at least 0.5% or more. Synthetically generated precipitation comparisons revealed that QA methods responded differently to simulated gauge evaporation and sensor noise signals. These differences were more pronounced for the very light precipitation scenario based on MAE as a percent of total precipitation. In every simulated case, the new weighted average calculation (wavgCalc) had a lower measure of error compared to the current pairwise calculation (pairCalc) regardless of sensor noise or gauge evaporation. This was also true for the nonprecipitating scenario, which indicates wavgCalc had a lower tendency to report false precipitation (type II errors). These results also suggest that wavgCalc was less sensitive to these nonprecipitating processes. Field comparisons revealed that the lessening of these sensitivities and easing of restrictive pairwise comparisons increased total station precipitation for more than 87% of the network. On average, USCRN stations reported 1.6% more precipitation using the wavgCalc method, which is similar in magnitude to the undercatch USCRN had with respect to COOP as reported by Leeper et al. (2015). A reevaluation of Leeper et al. (2015) using wavgCalc found USCRN reporting 0.21% more precipitation than collocated COOP stations.

Subhourly precipitation rates with wavgCalc were also found to be more realistic than pairCalc. The averaging approach of reference depths in pairCalc allows missed precipitation to be recaptured in subsequent subhourly periods. While this may improve total precipitation over longer time scales (e.g., monthly and annual), recaptured precipitation was found to negatively impact subhourly precipitation rates, creating unrealistic 5-min intensity values in some events. These scenarios were not common but more pronounced for colder precipitation events when disdrometer performance may be degraded (intermittent wetness signal), as frozen hydrometeors can fall undetected (Tabler 1998). However, disdrometer performance during snowy conditions was not always degraded, so additional research is currently being conducted at the precipitation test bed in Marshall, Colorado, to further

TABLE 3. PairCalc and wavgCalc 100-member-ensemble MAE average (mm) for synthetic heavy, very light, nonprecipitating, and constant-rate events by various levels of gauge evaporation (0.00–0.02) and wire noise (000, 111, 113, 133, and 333).

| | | | Noise level per wire | | | | |
|---|---|---|---|---|---|---|---|
| Generated events | QA variants | Gauge evaporation | 0 | 111 | 113 | 133 | 333 |
| Heavy | pairCalc | 0.00 | 0 | 0.06 | 0.11 | 0.16 | 0.24 |
| | | 0.01 | 0.06 | 0.06 | 0.09 | 0.15 | 0.24 |
| | | 0.02 | 0.13 | 0.10 | 0.11 | 0.16 | 0.24 |
| | wavgCalc | 0.00 | 0 | 0.06 | 0.08 | 0.13 | 0.18 |
| | | 0.01 | 0 | 0.07 | 0.08 | 0.13 | 0.19 |
| | | 0.02 | 0 | 0.07 | 0.08 | 0.13 | 0.19 |
| Very light | pairCalc | 0.00 | 0.49 | 0.34 | 0.30 | 0.28 | 0.52 |
| | | 0.01 | 0.78 | 0.47 | 0.40 | 0.23 | 0.30 |
| | | 0.02 | 0.88 | 0.57 | 0.50 | 0.33 | 0.24 |
| | wavgCalc | 0.00 | 0.08 | 0.11 | 0.13 | 0.16 | 0.22 |
| | | 0.01 | 0.10 | 0.16 | 0.16 | 0.17 | 0.21 |
| | | 0.02 | 0.25 | 0.23 | 0.23 | 0.22 | 0.23 |
| Nonprecipitating | pairCalc | 0.00 | 0 | 0 | 0 | 0.07 | 0.19 |
| | | 0.01 | 0 | 0 | 0 | 0.03 | 0.08 |
| | | 0.02 | 0 | 0 | 0 | 0.02 | 0.03 |
| | wavgCalc | 0.00 | 0 | 0 | 0 | 0.01 | 0.03 |
| | | 0.01 | 0 | 0 | 0 | 0 | 0.02 |
| | | 0.02 | 0 | 0 | 0 | 0 | 0.01 |
| Constant rate | pairCalc | 0.00 | 0 | 0.02 | 0.70 | 1.65 | 1.75 |
| | | 0.01 | 0.09 | 0.04 | 0.67 | 1.54 | 1.67 |
| | | 0.02 | 0.11 | 0.12 | 0.58 | 1.49 | 1.60 |
| | wavgCalc | 0.00 | 0 | 0.06 | 0.08 | 0.13 | 0.18 |
| | | 0.01 | 0 | 0.07 | 0.09 | 0.13 | 0.18 |
| | | 0.02 | 0.01 | 0.07 | 0.10 | 0.14 | 0.19 |

evaluate disdrometer performance and to identify sensor-related QC checks to better evaluate the quality of disdrometer measurements from the field.

One caveat of this study is the lack of a ''true'' precipitation dataset applied to the Geonor gauge. However, attempts were made to address this limitation by developing a precipitation generator to quantify QA performance with respect to a simulated known precipitation event. Generator simulations conducted without noise and gauge evaporation provide a true dataset equivalent from which to draw conclusions about the performance of both QA systems. With that said, further investigations evaluating both methods are ongoing, including a gauge evaporation field study conducted over the summer of 2013 and a disdrometer comparison study as noted previously.

In conclusion, two QA systems were extensively evaluated with the weighted average calculation (wavgCalc) system found to be less sensitive to wire noise and gauge evaporation, which from station comparisons generally resulted in increased precipitation and improved subhourly precipitation rates. Given the reliability of wavgCalc to detect artificial precipitation signals and the robustness of this QA system to withstand station irregularities (i.e., maintenance and broken wires), the wavgCalc system has proven to work well across the USCRN. Furthermore, by ensuring the quality of USCRN subhourly precipitation measurements, precipitation data from wavgCalc will be better suited for validation studies (model, radar, and satellite), hydrological forecasts (floods and droughts), and other high-temporal-resolution weather and climate impact studies in addition to accurately monitoring the nation's precipitation trends over climatological time scales from both mean and extreme perspectives. This study also provides an evaluation and testing outline that other networks can use to validate QA systems for precipitation in addition to highlighting techniques USCRN has explored while developing QA approaches for redundantly monitored precipitation systems.

*Acknowledgments.* This work was supported by NOAA through the Cooperative Institute for Climate and Satellites–North Carolina under Cooperative Agreement NA09NES4400006. The USCRN is supported by NOAA's Climate Program Office. We especially thank Scott Embler and Diana Kantor for their technical assistance and Michael Kruk, Scott Applequist, Russell Vose, Jay Lawrimore, Tom Peterson, and the external reviewers for their editorial suggestions. The views and opinions, and findings contained in this report are those of the authors and should not be construed

as an official NOAA or U.S. government position, policy, or decision.

## REFERENCES


Baker, B. C., R. Buckner, W. Collins, and M. Phillips, 2005a: Calculation of USCRN precipitation from Geonor weighing precipitation gauge. NOAA Tech. Note NCDC-USCRN-05-1, 27 pp. [Available online at http://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/program/technotes/TN05001GeonorAlgorithm.pdf.]

——, L. Larson, E. May, H. Bogin, and B. Collins, 2005b: Final report: Operational testing of various precipitation sensors in support of the United States Climate Reference Network (USCRN). NOAA Tech. Note NCDC-USCRN-05-2, 69 pp. [Available online at http://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/program/technotes/TN05002PrecipTesting.pdf.]

Daly, C., W. P. Gibson, G. H. Taylor, M. K. Doggett, and J. I. Smith, 2007: Observer bias in daily precipitation measurements at United States cooperative network stations. *Bull. Amer. Meteor. Soc.,* **88,** 899–912, doi:10.1175/BAMS-88-6-899.

Diamond, H. J., and Coauthors, 2013: U.S. Climate Reference Network after one decade of operations: Status and assessment. *Bull. Amer. Meteor. Soc.,* **94,** 485–498, doi:10.1175/BAMS-D-12-00170.1.

Duchon, C. E., 2008: Using vibrating-wire technology for precipitation measurements. *Precipitation: Advances in Measurement, Estimation and Prediction,* S. C. Michaelides, Ed., Springer, 33–58.

Fiebrich, C. A., and K. C. Crawford, 2009: Automation: A step toward improving the quality of daily temperature data produced by climate observing networks. *J. Atmos. Oceanic Technol.,* **26,** 1246–1260, doi:10.1175/2009JTECHA1241.1.

——, P. K. Hall, and R. A. McPherson, 2010: Quality assurance procedures for mesoscale meteorological data. *J. Atmos. Oceanic Technol.,* **27,** 1565–1582, doi:10.1175/2010JTECHA1433.1.

Goodison, B. E., P. Y. T. Louie, and D. Yang, 1998: WMO solid precipitation measurement intercomparison. WMO Instruments and Observing Methods Rep. 67, WMO/TD-872, 212 pp.

Holder, C., R. Boyles, A. Syed, D. Niyogi, and S. Raman, 2006: Comparison of collocated automated (NCECONet) and manual (COOP) climate observations in North Carolina. *J. Atmos. Oceanic Technol.,* **23,** 671–682, doi:10.1175/JTECH1873.1.

Kunkel, K. E., T. R. Karl, D. R. Easterling, K. Redmond, J. Young, X. Yin, and P. Hennon, 2013: Probable maximum precipitation and climate change. *Geophys. Res. Lett.,* **40,** 1402–1408, doi:10.1002/grl.50334.

Leeper, R. D., J. Rennie, and M. A. Palecki, 2015: Observational perspectives from U.S. Climate Reference Network (USCRN) and Cooperative Observer Program (COOP): Temperature and precipitation comparison. *J. Atmos. Oceanic Technol.,* **32,** 703–721, doi:10.1175/JTECH-D-14-00172.1

Legates, R. D., and J. G. McCabe Jr., 1999: Evaluating the use of ''goodness-of-fit'' measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.,* **35,** 233–241, doi:10.1029/1998WR900018.

National Research Council, 1999: *Adequacy of Climate Observing Systems.* National Academy Press, 51 pp.

NOAA, 2013: Billion-dollar weather and climate disasters: Table of events. National Oceanic and Atmospheric Administration, accessed 24 September 2014. [Available online at http://www.ncdc.noaa.gov/billions/events.]

Palecki, M. A., and P. Ya. Groisman, 2011: Observing climate at high elevations using United States Climate Reference Network approaches. *J. Hydrometeor.,* **12,** 1137–1143, doi:10.1175/2011JHM1335.1.

Rasmussen, R., and Coauthors, 2012: The NOAA/FAA/NCAR winter precipitation test bed: How well are we measuring snow? *Bull. Amer. Meteor. Soc.,* **93,** 811–829, doi:10.1175/BAMS-D-11-00052.1.

Sevruk, B., M. Ondrás, and B. Chvíla, 2009: The WMO precipitation measurement intercomparisons. *J. Atmos. Res.,* **92,** 376–380, doi:10.1016/j.atmosres.2009.01.016.

Shafer, M. A., C. A. Fiebrich, D. S. Arndt, S. E. Fredrickson, and T. W. Hughes, 2000: Quality assurance procedures in the Oklahoma Mesonetwork. *J. Atmos. Oceanic Technol.,* **17,** 474–494, doi:10.1175/1520-0426(2000)017<0474:QAPITO>2.0.CO;2.

Tabler, R. D., 1998: Improving the Wyoming road weather information system. Final Rep., Wyoming Dept. of Transportation Tech. Note FHWA-WY-98/02F, 81 pp. [Available online at http://ntl.bts.gov/lib/jpodocs/repts_te/7743.pdf.]

Trenberth, K. E., T. R. Karl, and T. W. Spence, 2002: The need for a systems approach to climate observations. *Bull. Amer. Meteor. Soc.,* **83,** 1593–1602, doi:10.1175/BAMS-83-11-1593.

You, J., K. G. Hubbard, S. Nadarajah, and K. E. Kunkel, 2007: Performance of quality assurance procedures on daily precipitation. *J. Atmos. Oceanic Technol.,* **24,** 821–834, doi:10.1175/JTECH2002.1.