

Interactive comment on “Evaporation from weighing precipitation gauges: impacts on automated gauge measurements and quality assurance methods” by R. D. Leeper and J. Kochendorfer

R. D. Leeper and J. Kochendorfer

ronald.leeper@noaa.gov

Received and published: 27 March 2015

Author Responses

The authors would like to thank each of the reviewers for their efforts in reviewing our manuscript entitled “Evaporation from Weighing Precipitation Gauges: Impacts on Automated Gauge Measurements and Quality Assurance Methods”. Reviewer comments and suggestions have culminated in a manuscript that is now more inclusive and thorough with the addition of algorithm descriptions and a more complete explanation of

C5263

diurnal depth variations over dry periods. They also led to the correction of figure 5 (previously 4) that originally displayed the incorrect column of data for the Geonor-NonEvap gauge, improving the manuscript further. These modifications have strengthened a manuscript that investigates the importance of quality control techniques on precipitation measurements that should appeal to both QC development and data user communities, resulting in what we hope will be a useful and well-cited manuscript. Provided below are author responses (italicized) and when necessary manuscript revisions (in purple) to each reviewer comment.

AMT-D Reviewer 1

1)The QA method evaluation is really the focus of this paper yet the description of these methods, especially the algorithms, are quite vague, somewhat confusing, and largely left to a reference of another submitted manuscript. a) Starting on page 12856 line 12, you note that the 1-minute data were “aggregated” into 5-minute values but this sentence is confusing and doesn’t explain how this was done. Was it an average of the 1-minute data or a re-sample? b) You need to provide at least one paragraph explanation for each of the two algorithms, instead of the 4 sentences on page 12856.

1a). The one-minute gauge depths were not averaged over the 5-minute clock intervals. Instead, the gauge depth at the top of each 5-minute clock period (0, 5, 10, . . . ,55). This may be better described as a subsampling of 1-minute gauge data. The sentence was modified to reflect this change as noted below.

To conform to the USCRN five-minute temporal resolution that is required by the QA methods, the one-minute observations from the experimental gauges were subsampled to five-minute periods by simply taking the 1-minute observation corresponding to each fifth minute period of the hour (0, 5, 10, 15, . . . , 55)

1b). This is a valid point that will make the manuscript more inclusive with additional QA details provided in submitted Leeper et al. manuscripts. A paragraph briefly describing each of the QA processes has been included in a new section entitled “Algorithm

C5264

Descriptions” as noted below.

2. Algorithm Descriptions Since detailed descriptions of both methods and comparisons are provided in Leeper et al. (2015b), brief descriptions of the two QA methods are provided here as a reference. The two QA methods evaluated in this study were both designed for the same USCRN precipitation system. For this reason, both methods share the same QA checks as identified in Table 1. These checks ensure the gauge is operating normally; within gauge capacity (600 mm for the gauges used in this study) and with functioning load sensors (failing sensors report negative gauge depths). To identify instances of gauge maintenance (which often includes calibration processes), a maximum depth change of 25 mm per five-minute period is also applied. A minimum depth change (detection) limit of 0.2 mm is also used to reduce the occurrence of false precipitation estimates due to sensor noise (Baker 2015; personal correspondence). However, the accuracy of the gauge above this minimum detection limit is 0.1 mm (i.e. QA methods can report increases in gauge depth of 0.3 mm). In both algorithms, if any of the load sensors fail these checks it is not used to evaluate precipitation. Moreover, the primary difference between the two QA algorithms is how a single depth change is computed from the triplicate measurements from each gauge as described in the proceeding paragraphs. The initial QA methodology computes depth change for each load sensor as the current depth minus a two-hour average of previous gauge depths (to smooth sensor noise), provided no precipitation has been observed in the preceding two hours. If precipitation has been observed, depth change is calculated as the difference between the current depth and the depth at the time of last reported precipitation. Once the individual depth changes (deltas) are computed for each of the three load sensors, they are compared to one another using a pairwise approach to determine which load sensors (also referred to as wires) will be used in the final calculation. This is similar to the approach USCRN uses to integrate redundant measures of atmospheric temperature. Only wire deltas that agree to within 0.2 mm pass this check and are used to calculate precipitation. The wires that pass the pairwise check are averaged together using a simple arithmetic mean. As a final check, if a collocated rain

C5265

detector has detected precipitation for this period, the mean is rounded to the nearest 10th decimal place and recorded as precipitation. This approach is referred to as the pairwise method throughout the remainder of this study with additional information provided in Leeper et al. (2015b). A more recent approach was developed that computes depth change over the five-minute sub-hourly period (current depth minus most recent depth) for each load sensor. This is similar to the way a human observer might process the gauge data and is one of the fundamental differences between the two methods. Previous gauge data (two-hours) is used in this approach to weight each wire by the inverse of its variance (a measure of sensor noise) from the three-wire delta mean. Precipitation is then calculated from the weighted average of the three load sensors and rounded to the nearest 0.1 mm, provided precipitation has been detected by the collocated rain detector. This approach will be referred to as the wAvg method for the remainder of this study. Additional details regarding the function and performance of this approach is also available in Leeper et al. (2015b).

2) On page 12858, you speculate that the increase in bucket weight in the evap gauge during low vpd and surface wind to possible condensation inside the gauge. This doesn't show up in the nonEvap gauge so I assume that you are inferring that this moisture is derived from evaporation of the liquid in the bucket. Can you comment on the impact of this on (false) precip estimates as a product of either of the algorithms? You could look at low vpd/wind speed periods to confirm your statements, or at least suggest that this could be done. Also on page 12858 lines 6-8, you comment that wind has less of an impact on the depth change in the evap gauge but I clearly see a decrease in depth with increasing wind in Fig 4d. You should revise this statement.

Further inquiry revealed the wrong data for the Geonor-NonEvap gauge was presented in Figure 4 (now Fig. 5). Figure panels have been updated accordingly. These changes do not affect the outcome of the study or presented results. However, the text was modified slightly.

For the Geonor-NonEvap gauge, changes in gauge depth were generally smaller than

C5266

Geonor-NonEvap and symmetric about zero regardless of atmospheric conditions.

Condensation buildup on the gauge-reservoir could be derived from two sources; the atmosphere or the previously evaporated internal-reservoir moisture. Since increases in gauge depth from the Geonor-NonEvap gauge are not as discernable, the internal-reservoir is likely a more dominant moisture source, as you suggested. However, it is also important to realize that during the early morning hours the evaporation signal, which dominates throughout the day resulting in mostly depth losses, is likely negligible at this time. This suggests that variations in gauge depth from the Geonor-Evap gauge may also be influenced by sensor noise, which we know from the Geonor-NonEvap gauge can vary about zero with positive and negative depth changes. Moreover, increases in gauge depth from the Geonor-Evap gauge during low winds and VPD may be a combination of both sensor noise and on some days condensation buildup. The timing of coincident peaks between both gauges during the early morning hours are likely caused by those days with condensation buildup.

Hourly increases in gauge depth at lower wind speeds (between 0 and 2 m s⁻¹), which reduced with increasing wind speed, may be related to the lack of an evaporation signal where gauge depth changes would be driven by sensor noise (slight positive and negative changes as shown from Geonor-NonEvap) and or condensation buildup on the gauge reservoir during the early morning hours when surface winds were generally calm.

As for condensation increases in gauge depth triggering false precipitation, this is unlikely since both algorithms require an independent sensor to also detect precipitation at the time of depth change. Recall, the "Dry" section only consists of periods where the precipitation detector did not sense precipitation. Reported increases in gauge depth under such conditions will be ignored by both QA methods.

As for wind speed, the authors drew this conclusion from the fact that losses in gauge depth were observed regardless of wind speed. For all wind speed conditions observed

C5267

(0 to 8 ms⁻¹), reductions in gauge depth of up to 0.5 mmhr⁻¹ were observed. However, hourly positive depth change did reduce with increasing wind speed, which may be due to both changes in dominant bias (sensor noise versus evaporation) and early morning condensation as noted previously. Additional analysis was performed that showed the inclusion of wind speed as a parameter in a linear model with VPD did not significantly improve the model's performance. The linear model root mean squared error (RMSE) with and without the wind speed parameter were 0.08035 and 0.08114 respectively. Based on these results, we hypothesize that the lack of an apparent relationship with wind speed is due to the fact that the air volume within the gauge is relatively well-shielded from the wind, and that convection caused by the solar heating of the gauge dominates the exchange rather than wind-driven mixing. This concept was introduced into the manuscript.

The lack of an apparent relationship with wind speed may be due to the fact that air volume within the gauge is well shielded from the wind, and that convection caused by the solar heating of the gauge dominates the exchange rather than wind-driven mixing.

3) On page 12858 lines 13-22, you attribute the amplified diurnal trend in the evap gauge to the occurrence of evaporation. I don't think this is the case. I'm wondering if this has more to do with the inherent noise in the sensors rather than an evaporation/condensation signal. The difference between hour 1 and hour 24 is more important I would think. Similarly, you make a comment on page 12859 lines 5-7 attributing the variability in sign to spatial variability in precipitation but could this also be a result of signal noise and the processing technique?

This is an important question about the source of depth variations (sensor noise or gauge evaporation) in the Geonor gauges over the dry period. The authors contend that variations in gauge depth are driven by both gauge evaporation and sensor noise. However, the dominant source is different between the Geonor-Evap and Geonor-NonEvap gauges. For instance, if sensor noise was the primary signal the diurnal variations in gauge depth between Geonor-Evap and Geonor-NonEvap gauges would

C5268

have similar magnitude, diurnal trends, and distribution. However, they are neither similar in magnitude, distribution, nor diurnal timing; see figures 4 and 6. Instead, the largest negative depth change over the diurnal period from the Geonor-Evap gauge coincided in time with the diurnal trend of the highest VPD (higher evaporative demand), suggesting that depth variations from this gauge were primarily influenced by gauge evaporation rather than sensor noise. In addition, sensor noise would consist of random positive and negative depth changes that should result in a net loss of \sim zero. While this was true of the Geonor-NonEvap gauge (see figure 4), this was not the case of depth variations from the Geonor-Evap gauge, which was biased towards negative values. The obvious explanation for this is gauge evaporation. After all, this gauge did report an accumulated loss of 228.5 mm, which cannot be explained by sensor noise alone.

Given the significant evaporation signal from the Geonor-Evap gauge, a QA method that is less sensitive to gauge evaporation should report similar precipitation totals between the two gauges. This was more true of the wAvg approach than the pairwise method. While sensor noise may have had some impact, the processing technique, which is the same between both gauges with the same algorithm, should have little impact. Based on the panel of figures presented in figure 7 (now 8), the pairwise algorithm consistently reported negative depth change prior to precipitation from the Geonor-Evap gauge, which is evidence of an evaporation bias (correct direction). This was not true of pairwise calculated depth change for the Geonor-NonEvap gauge as shown below in figure 1 as a reference. These results suggest that gauge evaporation had a greater role in gauge differences by QA algorithm than sensor noise, which would have been random with both positive and negative depth variations. In fact, a closer look at the provided figure below does reveal some of these random variations, which further indicates that sensor noise is secondary to the evaporative signal.

Fig. 1 Pairwise calculated depth change for Geonor-Evap and Geonor-NonEvap gauge data for precipitation events 25 (top) and 18 (bottom), which illustrate the impacts of

C5269

combined evaporation and sensor noise.

We thank the reviewer for making a good point regarding the statement on page 12859 lines 5-7. Differences between gauge measurements may not be due to only to spatial variability. The text has been modified accordingly.

Conversely, wAvg differences were variable in sign, which was likely caused by natural variations in the spatial distribution of precipitation and other sources of gauge (sensor noise) error.

4)On page 12860, lines 23-24, you note that evaporative biases shift the timing and impact the intensity of precipitation events. Can you comment on why this happens, assuming that it is a product of the algorithm?

The two methods quantify depth change differently. The pairwise algorithm uses a smoothing approach (two-hour average) to determine depth change. When gauge evaporation occurs within the two-hours prior to precipitation, calculated depth change becomes negative, resulting in an evaporative deficit that obviously impacts total precipitation (under-report). Examples of this are provided in figure 7. For precipitation events starting at lower intensities, depth change in the first couple five minute periods may not be sufficient enough to overcome the evaporation deficit and exceed the minimum detection limit of the gauge; 0.2 mm. The wAvg method computes depth change over a smaller period (5-minutes), which makes this method less sensitive to gauge evaporation (smaller evaporative deficit). In this scenario, precipitation start times can become shifted in time between the two methods due to the varying degrees of sensitivity to gauge evaporation. This has been better explained in the text as noted below.

... The wAvg method computes depth change over a smaller 5-minute period, and therefore the magnitude of evaporative losses over the calculation window is smaller than pairwise method. As a result, the wAvg method has a much smaller evaporation deficit compared to pairwise evident from figures 8a-f. In event 18, the overestimation of reference depths using the pairwise method resulted in missed precipitation that

C5270

was captured by the wAvg method. For this same event, precipitation was calculated with the pairwise method having little of no negative depth change (not shown) prior to precipitation from the Geonor-NonEvap (control) gauge, suggesting that QA processes were not the cause of the precipitation bias. For the wAvg method, the same total precipitation (0.3 mm) for event 18 was reported from both Geonor-Evap and Geonor-NonEvap gauges, suggesting this method was less sensitive to gauge evaporation. In addition to under-reporting total precipitation, poor evaluations of reference depths due to gauge evaporation bias caused the pairwise algorithm to report precipitation later in time for event 25 (Fig 8c and d) and with different sub-hourly intensities for event 16 (Fig 8e and f) relative to wAvg. Shifts in the timing and intensity of precipitation can occur when precipitation events start off light (low precipitation rates) and changes in gauge depth over the initial couple of 5-minute periods using the pairwise approach are not sufficient to overcome the evaporation deficit and exceed the 0.2 mm threshold.

5) In your conclusion, page 12861 lines 8-16, you basically suggest that evaporative suppressants are not required if you use the correct algorithm. I can understand this may be the case for measuring liquid precipitation events BUT many of the past studies that you cite in earlier sections have focused on solid precipitation, and with solid precipitation measurement, you require an antifreeze solution that usually contains methanol to inhibit stratification and freezing. Without an evaporative suppressant, this will not work. Be careful with your broad recommendations without caveats. Also, your statement in lines 17-19 is a bit misleading since you are suggesting that scientists using the data will not have to correct the data sets for evaporation: :but that's because the algorithm has already provided this adjustment. Maybe state this in a different way.

We realize that evaporative suppressants are necessary to measure solid precipitation, as described in the methodology section, and we did not intend to suggest otherwise. The text has been amended accordingly. However it is worth noting that the algorithm does not actually adjust the data for evaporation. The choice of algorithm affects the length of the calculation period, making it more or less susceptible to evaporation bi-

C5271

ases. The text has been modified to clarify this.

In addition, this does not suggest that evaporative suppressants are unnecessary year round. In a network like the USCRN that is designed to measure both solid and liquid precipitation, evaporation suppressants must still be used during the winter to suppress the evaporation of antifreeze chemicals necessary to measure the liquid equivalent of frozen hydrometeors.

6) In Table 1, the pairwise algorithm not only results in less precipitation in the evap gauge, but also in the control (nonEvap) gauge. This isn't really addressed in the paper but perhaps it should be.

The wAvg algorithm was developed in response to studies that identified additional limitations of the pairwise approach, beyond gauge evaporation, which are highlighted in Leeper et al. (2015a). These two approaches have been extensively evaluated against one another across the USCRN network and also using synthetic precipitation events in another manuscript (Leeper et al. 2015b). In that study, the wAvg approach was found to be less sensitive to simulated gauge evaporation and sensor noise, which resulted in greater total precipitation. This is in line with the differences discerned by the reviewer in Table 1 (now Table 2). The present manuscript is a follow up study investigating only the effects of evaporation on algorithm performance. While other differences between the QA methods are indeed interesting and significant, they have already been extensively investigated, and such an analysis is beyond the scope of this study.

Abstract: 1) Abstract has some wording and formatting issues that confuse the reader. The sentence from lines 12-14 is confusing. Try "Two Geonor gauges were collocated with one gauge using an evaporative suppressant (referred to as nonEvap) and the other with no suppressant (referred to as evap) so that...". Also, you use "suppressant" often but you don't describe what the suppressant is. 2) Suggest changing the names of the gauges to "Geonor-Evap" and "Geonor-NonEvap" might add to the readability of the

C5272

paper. I found that the interjection of “evap” into sentences when referring to the gauge was confusing when talking about evaporation at the same time (are you referring to the gauge or to what’s happening to the gauge?). Maybe it’s just me. 3)The way that you interjected the numbers in lines 17-20 impacted the readability of the sentence. 4)Line 21 should be “: : design affect gauge evaporation rates, computational methods can: : :” and Line 22 should be “It is hoped that this study: : :”

Abstract 1: Sentence on lines 12 – 14 was modified as suggested.

Two Geonor gauges were collocated, with one gauge using an evaporative suppressant (referred to as Geonor-NonEvap) and the other with no suppressant (referred to as Geonor-Evap) to evaluate evaporative losses and evaporation biases on precipitation measurements.

Abstract 2: Good point; the manuscript has been updated to reflect this new reference System as noted in our response to abstract comment 1 above.

Abstract 3: Moved first set of interjected number to improve readability of the sentence.

In general, the pairwise method that utilized a longer time series to smooth out sensor noise was more sensitive to gauge evaporation (-4.6% bias with respect to control) than the weighted-average method that calculated depth change over a smaller window (< +1% bias).

Abstract 4: A comma was added on line 21 and line 22 revised as suggested.

These results indicate that while climate and gauge design affect gauge evaporation rates, computational. . .

Introduction: 1)Page 12854, line 6: “: : can bias observations even when taken at subhourly : : :” 2)Page 12854, lines 13-16: The sentence “Addition analysis: : :” has several issues to correct and should be re-worded for clarity.

Introduction 1: Sentenced revised as suggested.

C5273

However, recent research comparing USCRN with COOP stations indicate gauge evaporation can bias observations even when taken frequently at a sub-hourly rate (Leeper et al. 2015a).

Introduction 2: The sentence was reworked to improve clarity.

Further analysis of the QA system, using synthetic precipitation events of a known precipitation signal revealed the method used to calculate depth change was sensitive to gauge evaporation bias and sensor noise (Leeper et al. 2015b).

Methodology: 1)Page 12855, lines 10-12: The Geonor and Tretyakov are not really similar. Instead of saying this, why don’t you simply describe the Geonor gauge. 2)Page 12855, line 13 and onward: you refer to “gauge depth”. What you are actually referring to is bucket weight which is translated by the logger into precipitation depth. Perhaps either call it bucket weight or explain why you call it “gauge depth”. 3)Page 12855, line 12: why don’t you reference the gauge sensitivity from the Geonor manual? 4)Page 12855, line 24: change “evaporation biases” to “biases as a result of evaporation”. 5)Page 12856, lines 1-10: lots of detail in the instrument descriptions that can be omitted.

Methodology 1: Sentence was reworded to remove the comment describing the Geonor as being similar to the Tretyakov gauge. Instead this section is now dedicated to describing the Geonor gauge.

In addition to supporting redundant monitoring, the Geonor gauge is also designed with an open vertical shaft (Fig. 1b) to limit splashing out (hydrometers splashing out of the gauge) and snow/ice capped (orifice) errors described by (Sevruk et al. 2009).

Methodology 2: The process by which load sensor data is converted to gauge depth is now described in this paragraph.

Each of the load sensors has an internal wire that when plucked vibrates at a frequency that changes with the tension on the wire as described by Duchon (2008). When

C5274

calibrated, the resonant frequency of the wire is converted to gauge depth, which is used to monitor precipitation (Leeper et al. 2015b).

Methodology 3: Based on field testing, the USCRN determined that the gauge can detect changes in depth of up to 0.1 mm; however, a minimum detection threshold of 0.2 mm is necessary to limit false precipitation due to noise (C.B. Baker 2015; personal correspondence). For instance, precipitation at a precision of 0.1 mm is reported as long as it exceeds 0.2 mm. This differs from the Geonor manual. This sentence was expanded to explain this further.

A minimum depth change (detection) limit of 0.2 mm is used to reduce the occurrence of false precipitation estimates due to sensor noise (C.B. Baker 2015; personal correspondence). However, the precision of the gauge beyond this detection limit is 0.1 mm (i.e. QA methods can report increases in gauge depth of 0.3 mm)

Methodology 4: Sentence modified as suggested.

...changes in gauge depth from the paired gauges were used to quantify evaporative loss (during dry periods) and evaluate biases as a result of evaporation on precipitation measurements during rainy conditions.

Methodology 5: Some of the detail has been removed, but some description of the sensors is still necessary.

To evaluate the sensitivity of gauge evaporation to atmospheric conditions, air temperature (Thermometrics 1000 Ω Platinum Resistance Thermometer housed in a MetOne 0766B Fan Aspirated Radiation Shield), humidity (Vaisala HMT337), and wind speed (MetOne 014A Wind Speed Sensor) were also monitored throughout the study period from a height of 1.5 meters. The sensors that monitor these variables were located on separate towers near the study gauges with temperature, humidity, and wind speed measurements taken approximately 10, 37, and 67 meters away respectively. In addition, a USB temperature logger (EL-USB-1) was submerged within the reservoir of

C5275

the Geonor-Evap gauge to monitor the internal water temperature of the Geonor-Evap gauge.

Results: 1)Page 12857, line 17: "gauge" should be plural. 2)Page 12859, line 17: where does 0.2mm come from and why?

Results 1: Former Line 17 on page 12857 has been revised as suggested

The results include two subsections: Dry Conditions to quantify evaporative loss from Geonor gauges, and Wet Conditions to evaluate the impact of gauge evaporation on reported precipitation.

Results 2: The 0.2 mm was determined from a field study conducted by Baker (2015; personal correspondence). This is a minimum threshold requirement to limit errors due to false precipitation as noted previously and described in the algorithm description and methodology sections. See author response to your third Methodology comment.

Conclusions: 1)Page 12860, line 2: ": : :was extensive when no suppressant is used."

Conclusions 1: Sentence modified to compare the Geonor-Evap and Geonor-NonEvap gauges.

The gauge evaporation field campaign revealed that evaporation from the Geonor T200B all-weather precipitation gauge, used by the USCRN network, was extensive and larger (greater RMSE values) than variations in gauge depth due to sensor noise observed from the Geonor-NonEvap gauge.

3)Page 12860, line 9-10: ": : :which likely challenges QA processes that distinguish between: : :"

Conclusions 3: Sentence revised as suggested.

The larger variations may further challenge QA processes that try to distinguish between noise and precipitation without a collocated instrument (precipitation detector) that can monitor the presence of precipitation.

C5276

4)Page 12860, lines 18-21: The sentence starting with “However: : :” is confusing and could be stated in a better way. Re-word.

Conclusions 4: This particular sentence was removed along with revisions to previous sentences as shown below.

Gauge evaporation impacted precipitation measurements from the two QA methods differently.

5)Page 12861, line 8: You can probably omit “National trends” as the data is likely used for much more than this.

Conclusions 5: “National trends” removed.

6)Page 12861, line15: Should “wildness” be wilderness?

Conclusions 6: Thank you, this has been corrected.

7)Page 12861, lines 19-23: The sentence beginning with “It is hoped: : :” is a bit confusing and should be reworded or made more concise.

Conclusions 7: The sentence has been reworded for clarity as shown below.

This study will help guide the development and evaluation of precipitation algorithms for weighing bucket gauges that are prone to evaporation.

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 12851, 2014.

C5277

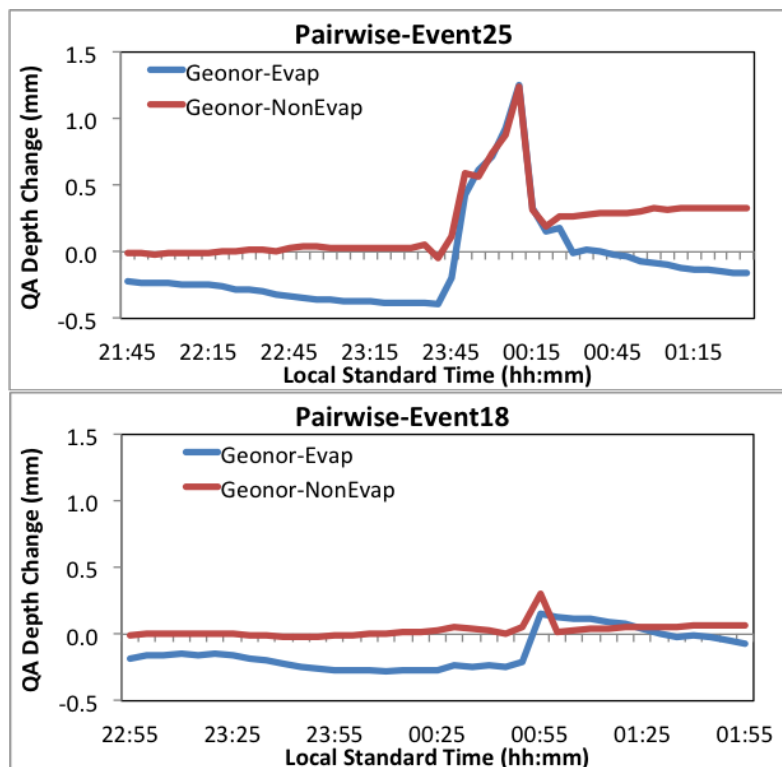


Fig. 1. Pairwise calculated depth change for Geonor-Evap and Geonor-NonEvap gauge data for precipitation events 25 (top) and 18 (bottom), which illustrate the impacts of combined evaporation and sensor

C5278