Synthesis of ARM User Facility Surface Precipitation Datasets to Construct a Best Estimate Value Added Product (PrecipBE)

Israel Silber[1], Jennifer M. Comstock[2], Adam K. Theisen[3], Michael R. Kieburtz[2], Zeen Zhu[4], Jenni Kyrouac[3]

[1] Atmospheric, Climate, and Earth Sciences Division, Pacific Northwest National Laboratory, Richland, WA, 99352, USA
[2] Advanced Computing, Mathematics, and Data Division, Pacific Northwest National Laboratory, Richland, WA, 99352, USA
[3] Environmental Science Division, Argonne National Laboratory, Lemont, IL 60439, USA
[4]Environmental Science and Technologies Department, Brookhaven National Laboratory, Upton, NY 11973, USA

Correspondence to: Israel Silber (israel.silber@pnnl.gov)

Abstract. Surface precipitation measurements are essential for Earth system model (ESM) evaluation and understanding cloud processes. An ever-growing need for robust, temporally evolving, and easy-to-use statistical datasets provides motivation for a baseline ground-based precipitation properties data product. The U.S. Department of Energy Atmospheric Radiation Measurement (ARM) user facility operates an extensive suite of precipitation instruments with various sensitivities and operating mechanisms, which render the decision of which instrument to use based on one or more fixed thresholds challenging and prone to errors and bias. Using a long-term instrument inter-comparison from a unique per-precipitation event perspective, rather than instantaneous sample comparison, we demonstrate that ARM precipitation instruments are generally consistent with each other at the statistical level. Inter-instrument deviations at the single event level can be large, especially at specific precipitation event properties such as maximum precipitation rates. A machine-learning (ML) analysis indicates that in some cases (e.g., certain instruments or deployments), atmospheric state variables influence the measured quantities and therefore the observed deviations between instruments. These results motivate the design of the ARM precipitation best-estimate (PrecipBE) value-added product, which incorporates all valid precipitation data while considering data quality and other instrument limitations.

PrecipBE consists of time series and tabular statistics datasets in an easy-to-use and insightful per-precipitation event format. It provides a large set of precipitation event properties supplemented with ancillary data from various ARM datasets that correspond to the detected precipitation events. We describe the PrecipBE algorithm and demonstrate its use via the examination of a single-day output as well as a long-term trend analysis of precipitation events at the ARM Southern Great Plains (SGP) site, covering more than 30 years of data. The trend analysis tentatively suggests a long-term tendency for mainly shorter and less intense precipitation events at the SGP site, but a long-term increase in annual rainfall by more than 36 mm (5%) per decade. This rainfall trend is catalyzed primarily by more extreme event properties of relatively rare, intense precipitation events, with event total and 1-minute maximum precipitation rate at a 1-year timeframe increasing up to 5 mm and 9 mm/hr (several percent) per decade, respectively. While the currently available PrecipBE datasets (at https://adc.arm.gov/discovery/) cover multiple ARM deployments up until March 2025, PrecipBE will soon become an

operational product with a several-day lag from real-time, and we invite the ARM user community to leverage this new product and welcome user feedback to enhance it further.

1 Introduction

Surface precipitation measurements serve as a crucial benchmark in Earth system model (ESM) evaluation (e.g., Emmenegger et al., 2022; Mikkelsen et al., 2024; Zhang et al., 2017) and aerosol-cloud interactions (ACI) studies (e.g., Christensen et al., 2024; Martin et al., 2017), among other process understanding efforts. Detailed case studies using surface precipitation data often require temporally evolving precipitation rate and accumulation data to account for the dynamic nature and short time scales of cloud evolution relative to the typically slower-evolving atmospheric state (e.g., Bretherton et al., 2010). These time series data serve as target quantities (benchmarks) for model simulations or analytical models. Certain precipitation-characterizing disdrometers, such as laser and video disdrometers, provide additional observational constraints on the precipitation properties, such as hydrometeor particle size distributions (PSDs). ESM evaluation studies, on the other hand, often rely on bulk statistics or data subsets and, therefore, utilize isolated precipitation event statistics after conditioning on quantities such as surface temperature, for example.

The U.S. Department of Energy Atmospheric Radiation Measurement (ARM) user facility (Mather, 2024; Mather et al., 2016) operates multiple types of precipitation-measuring instruments, including impact (Bartholomew, 2016a), video (Bartholomew, 2020b), and laser disdrometers (Bartholomew, 2020a), as well as tipping and weighing bucket rain gauges (Bartholomew, 2019; Kyrouac and Tuftedal, 2024). Each instrument tends to have higher sensitivity and/or better accuracy at certain precipitation conditions (e.g., Ciach, 2003; Fehlmann et al., 2020; Ro et al., 2024; Wang et al., 2021). For example, the Pluvio2 weighing bucket operated by ARM tends to be robust at high precipitation rates (Ro et al., 2024; Saha et al., 2021). The OTT Parsivel2 (LDIS; Bartholomew, 2020a), distributed in many ARM sites, is generally considered robust, but has been shown to suffer from biases at a specific drop size range (e.g., Raupach and Berne, 2015) and to underestimate the vertical velocity of drops larger than 1 mm, which translates to precipitation rate underestimation (Tokay et al., 2014). Similarly, the two-dimensional video disdrometer (VDIS; Bartholomew, 2020b) is often treated as a reference precipitation instrument, specifically when the drop PSDs are of interest (e.g., Tokay et al., 2020). However, this instrument is more likely to underestimate precipitation amounts in cases with drops smaller than roughly 0.3 mm (corresponding to its first size bin) or when large drops (> ~2.4 mm; often commensurate with heavy precipitation) are observed, due to terminal velocity underestimation (e.g., Tokay et al., 2013).

The availability of independent studies evaluating the performance of precipitation instruments under strict laboratory conditions (e.g., Colli et al., 2013; Saha et al., 2021) is still scarce. Moreover, comprehensive analyses of precipitation errors as a function of various background conditions (high wind, etc.) and deployment configurations (e.g., Montero-Martínez et al., 2016; Montero-Martínez and García-García, 2016; Wang et al., 2021), let alone snowy conditions (e.g., Battaglia et al., 2010; Milewska et al., 2019; Yuter et al., 2006), is still limited and requires additional research. In the interim, however, determining the "*true*" precipitation properties or weighting different ARM instrument samples based on the current literature is prone to unpredictable errors and biases. Therefore, as comprehensively discussed below, straightforward statistics combining data

from measurements collected (per deployment) would ostensibly provide the best estimates of precipitation event properties (onset and ending, accumulation, precipitation rates, etc.).

Here, we first present a long-term multi-instrument inter-comparison of precipitation event data collected at the ARM Southern
70 Great Plains (SGP; Sisterson et al., 2016) observatory (Section 2). Supported by the application of a machine learning (ML) algorithm, this analysis underscores the challenge in such cases of multi-instrument data without a clear and consistent *"true"* benchmark. The results from this comparison serve as a strong motivation for a best-estimate data product implementing straightforward statistics. These comparison results are also used to guide the design of the ARM precipitation best-estimate (PrecipBE) value-added product (VAP), the processing algorithm of which is elaborated on in Section 3. Section 4 describes
75 PrecipBE's data structure, and Section 5 presents a brief trend analysis using more than 30 years of ARM precipitation data from the ARM SGP site, available on the ARM Data Discovery (https://adc.arm.gov/discovery/). Conclusions and a short outlook are given in Section 5.

2 Instrument Inter-Comparison as Motivation for a Best-Estimate Data Product

2.1 Data Processing

80 Which precipitation instrument has the most reliable precipitation readings and should be used by default in given conditions? An answer to this question is not trivial. First, precipitation instruments have different sensitivities, which are influenced by ambient conditions and are often impacted by the same variables they aim to measure, namely, precipitation amount, rate, or particle properties, as noted above. In addition, those instruments have minimum quantization sizes, which could result in inconsistencies concerning precipitation event onset and ending times, leading to differences in event totals. As such, data
85 mining efforts aimed at determining those instrument strengths and weaknesses require a baseline definition of precipitation events instead of typical instantaneous sample comparisons. In this section, we perform an inter-comparison on a per-event basis by examining inter-instrument differences in precipitation event properties.

The analysis focuses on precipitation data collected at the ARM SGP site's co-located central (C1) and extended facility 13 (E13) over a 14-year period, from January 10, 2011, to January 10, 2025. A list of the instruments and data products analyzed
90 is provided in Table 1. For a given instrument, we define a precipitation event as a set of positive accumulated precipitation samples with gaps between neighboring precipitation readings (samples) shorter than 30 min. (larger gaps in event definition such as 60 min were tested and exhibited minor changes; not shown). Instrument events continuing to the next day are concatenated as long as they follow the same 30-min maximum gap logic. If the total accumulation in a given instrument event is smaller than 0.1 mm, it is omitted from this analysis. Instrument events that failed quality control (QC) checks (for calibration
95 issues, bad samples, etc.) in some or all event samples are also omitted from this analysis. Finally, a given event is also omitted if it indicates highly unlikely statistics; specifically, event total > 300 mm, event period > 5 days, mean precipitation rate > 120 mm/hr, and/or 1-min average maximum precipitation rate > 300 mm/hr. Some of these thresholds have been met and confirmed in recorded history (e.g., Koralegedara et al., 2019; Lagouvardos et al., 2013), but to our knowledge, have not

previously occurred during ARM deployments. However, these thresholds are rarely exceeded in instrument samples, for
100 various reasons, and account for up to a few percent of precipitation events detected using all ARM SGP instruments (counting from 2011), except for the optical rain gauge (ORG; Bartholomew, 2016b), with nearly 9% of detected events having one or more variables exceeding these thresholds. We note that ARM is in the process of retiring the ORG, which will not serve as a data source going forward.

datasets of the PWD (Kyrouac et al., 2021), AOSMET (Kyrouac and Tuftedal, 2010), DISDROMETER (Wang, 2006),
105 VDISQUANTS (Hardin et al., 2021), LDQUANTS (Hardin et al., 2019), TBRG (Kyrouac et al., 2006, 2021), RAINWB (Shi et al., 2010), and ORG (Kyrouac et al., 2021) from the ARM SGP site are available on the ARM Data Discovery (https://adc.arm.gov/discovery/; last access: 10 March 2025).

Table 1: Precipitation instruments and data products included in the analysis presented in Section 2 and incorporated in the PrecipBE value-added product.

Abbreviated name	Description	Reference
PWD[1,2]	Vaisala RAINCAP acoustic sensor as part of the Present Weather Detector, a component of the surface meteorological system (MET) at the main observatory	(Kyrouac et al., 2021; Kyrouac and Tuftedal, 2024)
AOSMET[1,2]	Vaisala RAINCAP acoustic sensor as part of the meteorological station associated with the Aerosol Observing System (~10-meters above ground)	(Kyrouac, 2019a; Kyrouac and Tuftedal, 2010)
DISDROMETER[1,2,3]	Joss-Waldvogel impact disdrometer	(Bartholomew, 2016a; Wang, 2006)
VDISQUANTS[1,2]	Joanneum Research two-dimensional video disdrometer quantities value-added product	(Bartholomew, 2020b; Hardin et al., 2020, 2021)
LDQUANTS[1,2]	OTT Parsivel2 laser disdrometer quantities value-added product	(Bartholomew, 2020a; Hardin et al., 2020, 2021)
WBPLUVIO2[1,2]	OTT Pluvio2 weighing bucket rain gauge	(Bartholomew, 2019; Zhu et al., 2016)
TBRG[1,2]	Novalynx Tipping bucket rain gauge; commonly part of the MET system	(Kyrouac et al., 2021; Kyrouac and Tuftedal, 2024)
METWXT, PRECIPMET, MARINEMET, and ABMMET[2]	Vaisala RAINCAP acoustic sensor as part of the Vaisala WXT520 or WXT530 meteorological instrument systems installed at various ARM and ARM-related facilities	(Holdridge and Kyrouac, 2012; Howie et al., 2016; Kyrouac, 2019b; Kyrouac et al., 2017; Kyrouac and Shi, 2018; Reynolds et al., 2017)

PWS[2,4]	Vaisala FD12P Present Weather Sensor meteorological system	(Kyrouac and Tuftedal, 2001; Ritsche, 2008)
RAINWB[1,4]	Belfort weighing bucket rain gauge	(Bartholomew, 2016c; Shi et al., 2010)
ORG[1,4]	Optical Scientific, inc optical rain gauge; commonly part of the MET system	(Bartholomew, 2016b; Kyrouac et al., 2021)

110 [1]Included in the analysis presented in Section 2.
[2]Incorporated in PrecipBE (where available)
[3]ARM changed the DISDROMETER code name to IDIS starting 2025-04-08, outside the date span examined in this study
[4]Retired instrument

115 To streamline the interpretation of analysis results, we select a "reference" instrument to examine deviations of events from one instrument to another. Thus, we inter-compare pairs of instruments, with one of them being the "reference" instrument. This "reference" instrument is not a *"true"* benchmark, and its related biases can still be characterized. For example, in cases where most or all other precipitation instruments show a consistent deviation from the reference, we can tentatively conclude that the observed bias originates in the reference instrument.

120 Ideally, the best reference instrument would be the tipping bucket rain gauge (TBRG; see Kyrouac and Tuftedal, 2024), because it was the first deployed precipitation instrument at the ARM SGP site (since 1993), and is still operational, covering the whole operation period of all other precipitation instruments. However, the TBRG has a very coarse precipitation amount least count (minimum detection of 0.254 mm; 0.1 inch), rendering its sensitivity and general accuracy (in weak events) inadequate for serving as a reference instrument (as demonstrated below), especially compared to other instruments such as disdrometers.

125 Therefore, we chose to use the Present Weather Detector (PWD), which is integrated in the ARM Surface Meteorological System (MET; Kyrouac and Tuftedal, 2024), as the reference instrument. The PWD has a very long record at the ARM SGP site, starting on January 10, 2011, enabling inter-comparison with a wide range of instruments.

One of the main challenges in a per precipitation event multi-instrument inter-comparison is associating individual instrument precipitation events with the reference instrument event, primarily due to the different onset and event duration times. This

130 challenge is exemplified in the simplified diagram shown in Figure 1. In this case, three precipitation events are identified in the PWD data (reference instrument). One or more events detected with other instrument data can be aggregated and become associated with a given reference instrument event (as a single event). For example, events 1 and 2 detected using the LDIS are associated with the PWD's event number 1, while events 3, 4, and 5 detected using the TBRG data are associated with the PWD's event number 2. However, to prevent event conflicts in the inter-comparison, multiple reference instrument events

135 cannot be associated with a single event detected using a different instrument. In such cases, the instrument events are omitted from the inter-comparison. For example, event 1 detected using the VDIS or the LDIS event 4. In the latter case, we have interlacing conditions, resulting in the exclusion of LDIS event 3 as well since including it would likely result in a negative bias when comparing it to the PWD's event number 2.

140
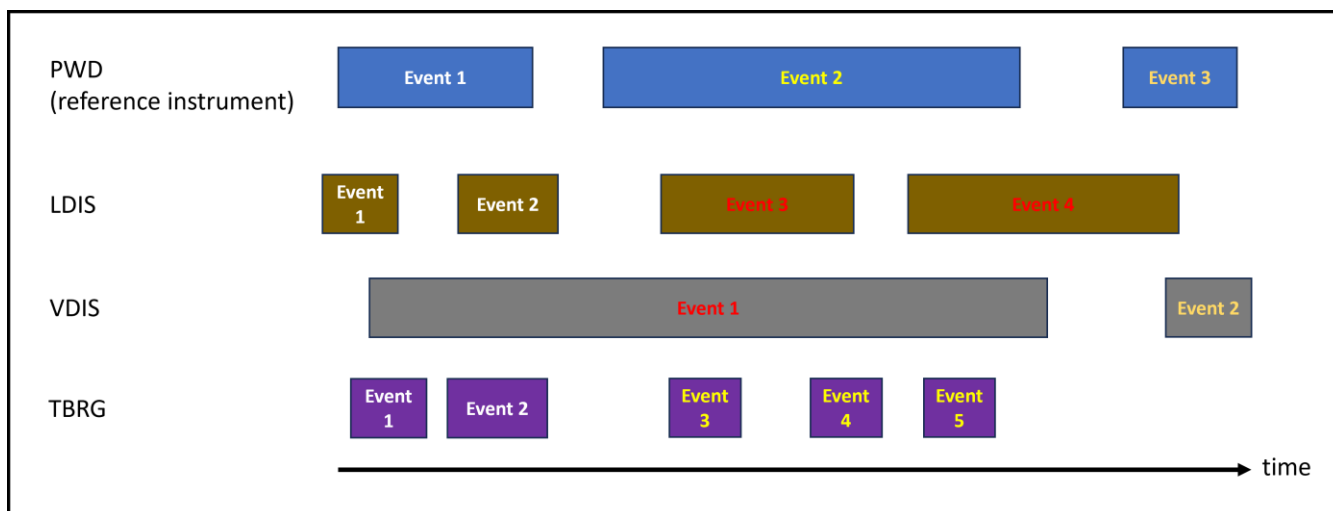
Figure 1: Simplified diagram exemplifying the challenge of associating precipitation events detected using different instruments when a reference instrument is used. Here, the present weather detector (PWD) serves as the reference instrument, and its events are designated using different font colors. Precipitation events detected using the laser disdrometer (LDIS), the video disdrometer (VDIS), or the tipping bucket rain gauge (TBRG) become associated with PWD events only if they are not conflicting with it (event
145 **font colors match the associated PWD events). Conflicting events are designated using the red font color.**

This event association and aggregation exercise results in the removal of some instrument pair events. Removal percentages range from 0.7% of TBRG events to 53% of the PWD event pairs with the Belfort weighing bucket rain gauge (RAINWB; Bartholomew, 2016c). Smaller conflicting percentages, such as in the case of the TBRG or the Pluvio2 weighing bucket

150 (WBPLUVIO2; Bartholomew, 2019) with 4.5% of events being conflicted with the reference instrument, are often the result of the compared instrument tending to record shorter events than the reference (see the TBRG versus PWD example in Figure 1). Larger conflicting percentages, such as in the case of the RAINWB or the Joss-Waldvogel impact disdrometer (DISDROMETER; Bartholomew, 2016a) data, with 47%, often occur when the compared instrument tends to longer events than the reference instrument (see the VDIS events versus PWD example in Figure 1). We note that the filtering of QC-flagged

155 or anomalous reading events prior to the aggregation exercise had minor influence on analysis results (not shown), but it could theoretically be more impactful in other cases.

2.2 Inter-Comparison Results

Figure 2 shows probability density functions (PDFs) of precipitation event total amount deviations of different ARM instruments from the reference (PWD). As suggested by the shape of the PDFs, the vast majority of ARM precipitation

160 instruments tend to be consistent with each other, with mean deviations (μ) smaller than 3 mm in magnitude and variability (represented here by the standard deviation; σ) being smaller than 10 mm. The PWD appears to be consistent to the greatest extent with the TBRG and the WBPLUVIO2 (means of 0.5 mm or less; σ on the order of 5 mm; in Figure 2a and Figure 2h, respectively). Some instruments and data products tend to record larger event totals relative to the PWD (e.g., LDQUANTS

in Figure 2e, AOSMET in Figure 2f) whereas others exhibit a tendency for smaller totals (e.g., VDISQUANTS in Figure 2c,

165   DISDROMETER in Figure 2g). These patterns are robust with the same qualitative results and minor quantitative variations

if only events with totals greater than 1 mm are analyzed, for example, and deviations appear directly susceptible only to the

magnitude of the evaluated variable (i.e., event total) in the reference instrument (examined via linear regression; not shown).

While the RAINWB is statistically consistent on average with the PWD (Figure 2d), its variability is somewhat greater than

the other instruments. However, it is the ORG's deviations that stand out with a much larger variability (~15 mm) and an

170   average overestimation by more than 5 mm (Figure 2b) (see also Kyrouac and Tuftedal, 2024). This overestimation becomes

stark when conditioning on event totals greater than 1 mm with an average deviation from the reference of +36 mm.
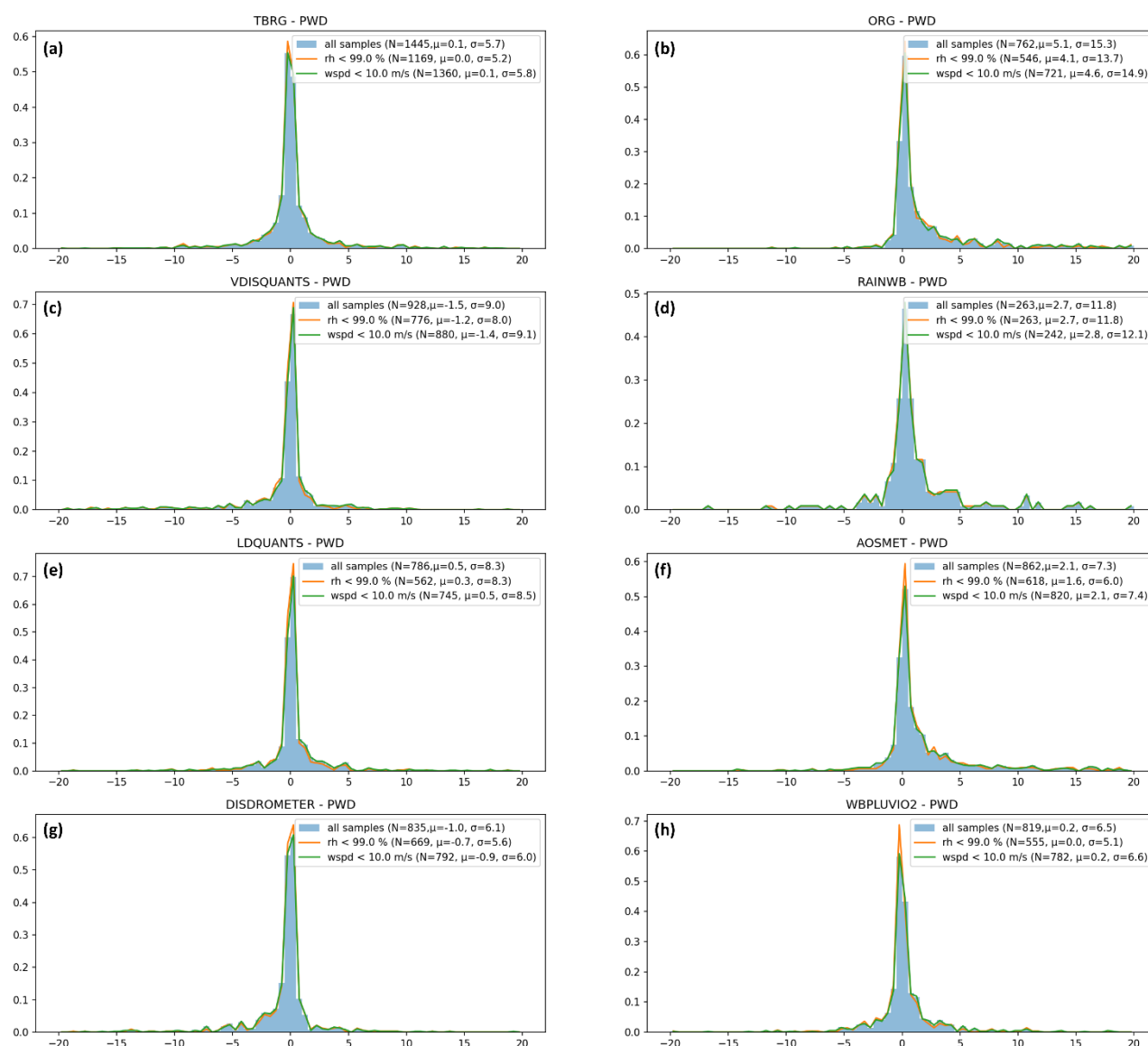
Figure 2: Probability density functions (PDFs) of ARM SGP instrument precipitation event total amount deviations from the PWD, serving in this inter-comparison as the reference instrument (bin size of 0.5 mm). The blue histogram is calculated using all precipitation event samples, while the orange and green curves designate histograms calculated while conditioning on event-mean relative humidity (omitting likely foggy conditions) and wind speed (omitting strong winds), respectively, both of which are derived from MET observations. In each panel, see the legends for the total number of event samples (N), mean deviation (μ), and standard deviation (σ). x-axis and legend quantity units are mm.

The differences between instrument precipitation event measurements are relatively more variable when examining event periods (Figure 3). The ORG measurements suggest precipitation events are positively skewed and longer by more than 50 min, on average, with a considerable variability exceeding 140 min (Figure 3b). (Note that some of the positive PDF skewness is influenced by the aggregation and filtering methodology discussed above). The DISDROMETER shows a stronger tendency, with precipitation events being 145 min longer on average (Figure 3g), whereas the RAINWB exhibits an even more substantial positive bias, exceeding 7 hours (Figure 3d). These long-event tendencies reflect the challenge in aggregating precipitation events, which resulted in the exclusion of a large subset of samples taken by those instruments from this analysis. In fact, the RAINWB event period bias and errors are so large, to an extent that is highly challenging to reconcile in an integrated dataset without introducing significant biases. In this regard, the PWD role as a reference instrument can be justified in the current analysis by the instrument's precipitation measurement properties being "somewhere in the middle" across the ARM precipitation instrument suite. The PWD's event period statistics and general instrument behavior is in good agreement with the VDISQUANTS and LDQUANTS VAPs (Figure 3c and Figure 3e, respectively), with average deviations of up to several minutes, as well as with the AOSMET with average deviations of 14 min (Figure 3f). The TBRG (Figure 3a) and WBPLUVIO2 (Figure 3h) display negatively skewed deviation relative to the reference instrument, with mirror-like patterns compared to the ORG and DISDROMETER.
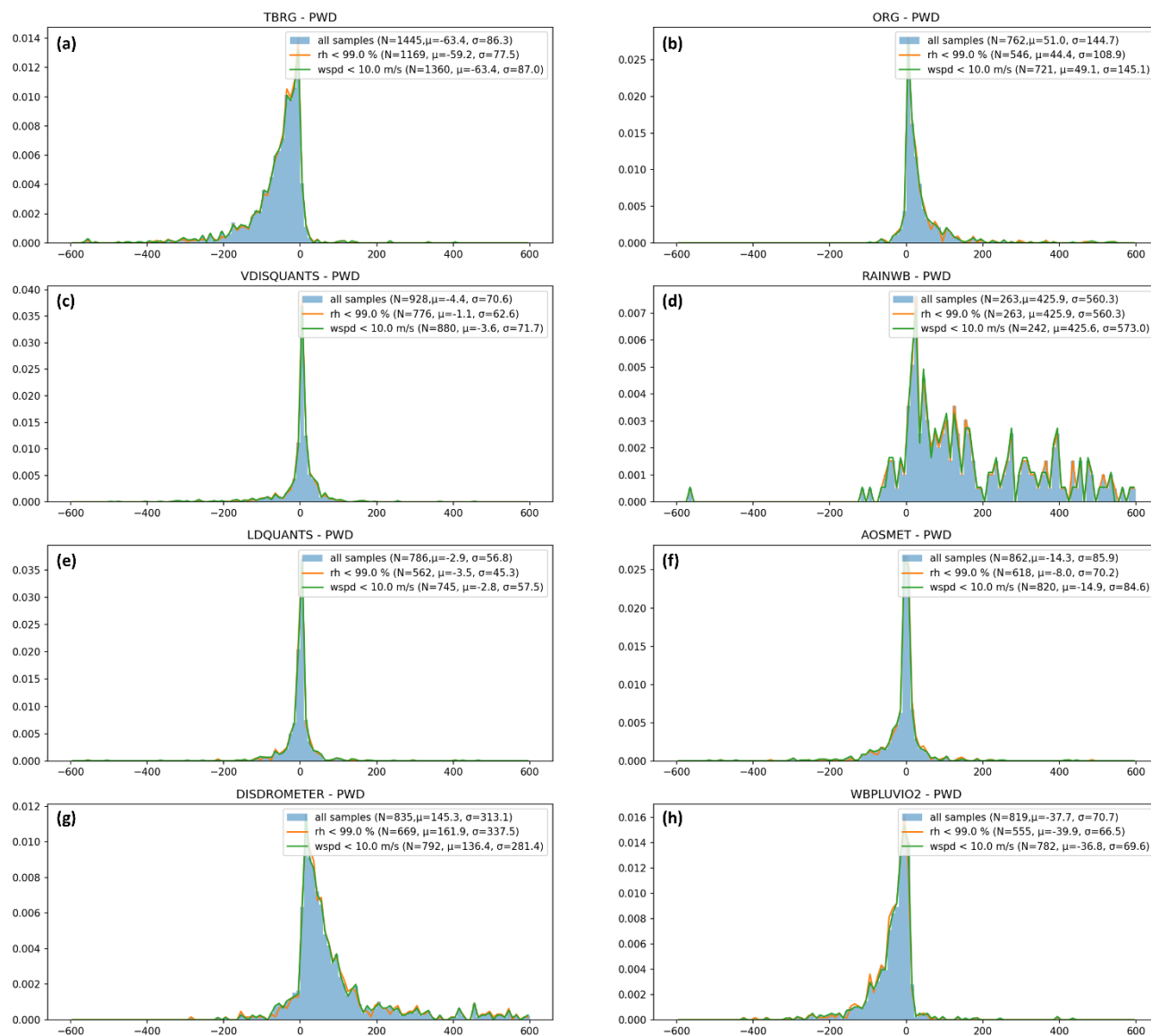
Figure 3: As in Figure 2, but for the event period (bin size of 10 min; x-axis and legend quantities are given in units of min).

The event 1-min-average maximum and event-mean precipitation rate comparisons (Figure 4 and Figure 5, respectively) suggest that most instruments are generally consistent with each other, especially in the case of mean rates, with the VIDSQUANTS (Figure 5c), LDQUANTS (Figure 5e), WBPLUVIO2 (Figure 5h), and AOSMET (Figure 5f) having average differences from the reference of 1 mm/h or less, and standard deviations of 3 mm/h or less. In both the mean and 1-min maximum precipitation rates, the TBRG (Figure 4a and Figure 5a) exhibits a distinct bi-modal PDF shape, which originates

in its coarse minimum least count of 0.254 mm. Accounting for this instrument limitation by omitting precipitation events with total amounts smaller than 1 mm results in a behavior consistent with the above-mentioned instruments and the disappearance of the bi-modal PDF artifact, suggesting that higher event total thresholds should be used for the TBRG in an integrated data product. The negative (positive) event period tendency of the TBRG in Figure 3a (DISDROMETER in Figure 3g) are compensated by the positive (negative) event-mean precipitation rates observed in Figure 5a (Figure 5g), resulting in a net event amount that is in agreement with other instruments, as indicated in Figure 2. In this regard, the DISDROMETER compensates for its relatively large event-mean rate variability with a restrained event 1-min maximum precipitation (Figure 4g). While this event characteristic is generally the most variable across the various instruments (Figure 4) due to the irregular, potentially tempestuous nature of precipitation over the commonly-used 1-min precipitation instrument averaging period, combined with sensitivity limitations of different instruments, the ORG and RAINWB exhibit a much more erratic behavior. Specifically, The RAINWB significantly underestimates both the event 1-min maximum precipitation rate (Figure 4d) and event-mean precipitation rate (Figure 5d) by several tens of mm/h, which provides an extreme case of error compensation resulting in a moderate bias, as seen in the event total amount PDF (Figure 2d). The ORG displays less variable, yet inconsistent response, overestimating event precipitation rate peaks (Figure 4b) but underestimating average precipitation rates (Figure 5b), a behavior that becomes even more pronounced when focusing on events with precipitation totals exceeding 1 mm (not shown).
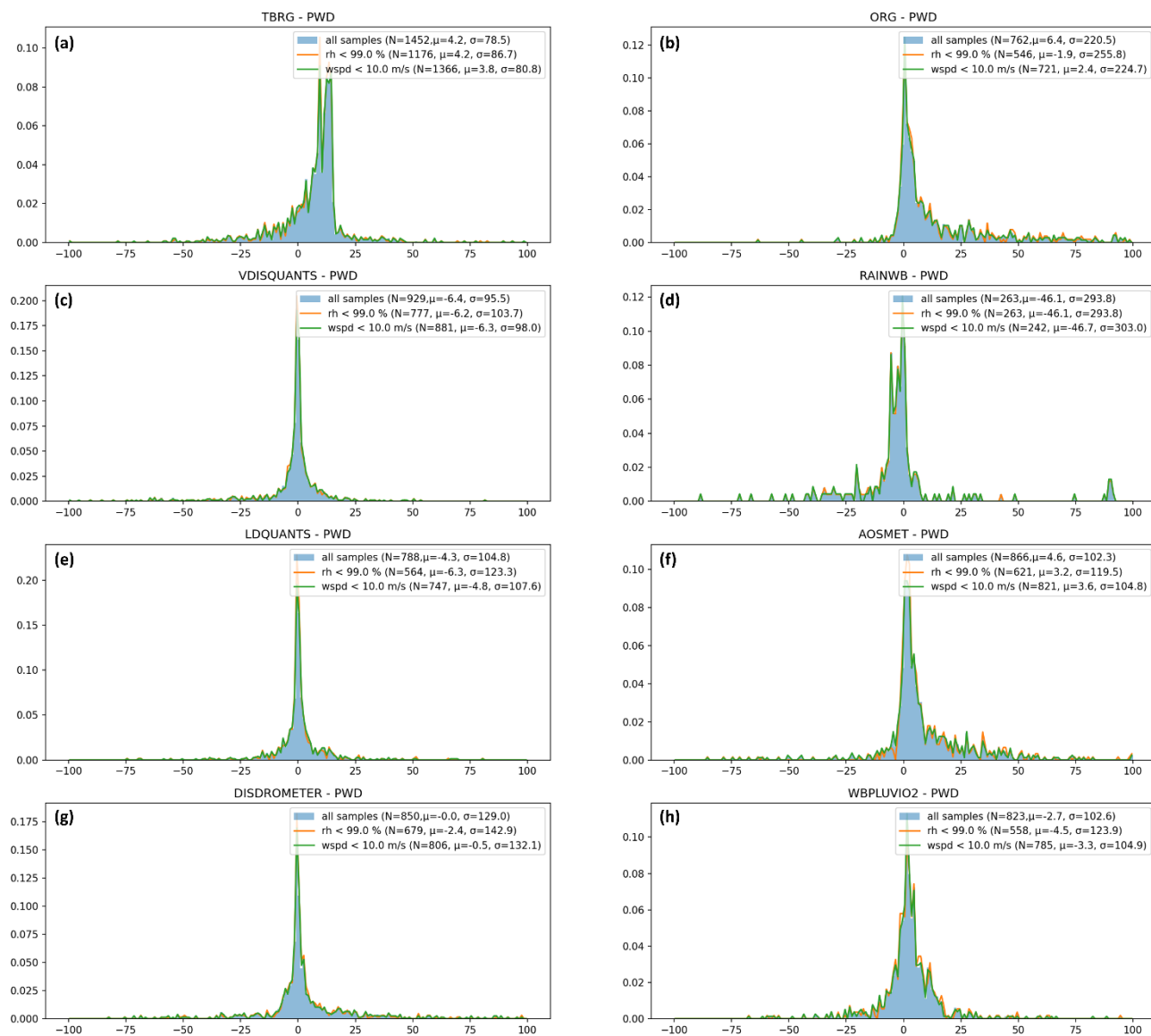
Figure 4: As in Figure 2, but for event 1-min-averaged maximum precipitation rate (bin size of 1 mm/h; x-axis and legend quantities are given in mm/h units).
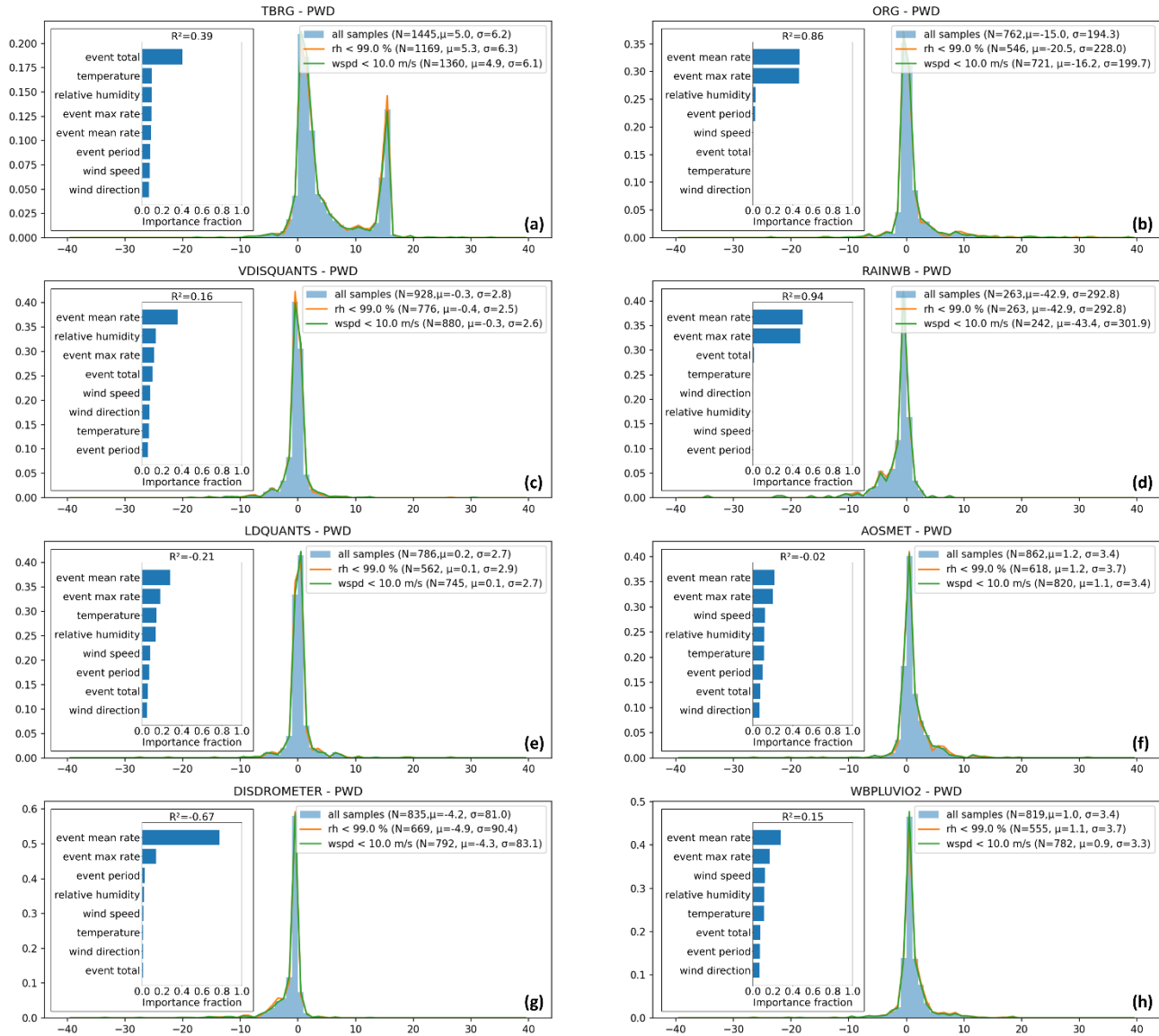
Figure 5: as in Figure 2 but for event-mean precipitation rate (in mm/h; bin size of 1 mm/h). The inset panels show feature importance analysis of various PWD event properties and event-mean atmospheric state variables derived from MET observations. The feature importance results are derived from a random forest regression model fit (see text) with the coefficient of determination specified at the top of the inset.

Ambient conditions can influence deviations between instrument measurements, often as a function of the instrument operation mechanism (e.g., Bartholomew, 2016c, 2020a; Kyrouac and Tuftedal, 2024; Montero-Martínez et al., 2016; Wang et al., 2021). In the secondary PDFs illustrated in Figure 2, Figure 3, Figure 4, and Figure 5, we examined the impact of some of these influencing factors. Specifically, by excluding events with event-mean relative humidity exceeding 99% (likely foggy

conditions; accounting for the MET system's uncertainty) or events with event-mean wind speeds higher than 10 m/s (high winds), some of the potential impact of these forcings on event statistics can be evaluated. The effect of wind speed appears somewhat limited, with mixed behavior of increasing or decreasing the deviation magnitude or standard deviations. This mixed
235   and weak behavior could be the result of the wind speed threshold used, or simply that examined instruments are robust to wind effect, which is likely not the case. (We note that surface winds stronger than 10 m/s are likely more impactful on certain instrument samples, but such events are too rare in the SGP dataset to form a robust statistical dataset). The relative humidity (RH) PDFs, however, indicate that conditioning for foggy conditions reduces the instrument deviation mean and standard deviations in precipitation event totals across all instruments (Figure 2) and most instruments in the event period (Figure 3),
240   but the results are less conclusive regarding precipitation rates (Figure 4 and Figure 5).

To further demonstrate the challenge in disentangling the influence of different parameters on the difference in precipitation event measurements and statistics, we conduct a feature importance analysis by using the Random Forest (RF) regressor available in the Scikit-Learn Python package (Pedregosa et al., 2011). The feature importance analysis enables ranking the factors (features) that are most influential on the fitted RF model; i.e., features that have the most impact on the prediction of
245   the model's target variable (in this case, inter-instrument deviations). Using the default algorithm's hyperparameters (100 estimators/trees, unlimited tree depth, etc.), we input as features the four precipitation event properties from the reference instrument (total, period, mean, and 1-min maximum precipitation rates) as well as the event-mean temperature, RH, pressure, wind speed, and wind direction measured by the MET system. We run the algorithm separately for instrument pairs and event properties; that is, a single RF algorithm run examines the deviations of an instrument pair in one of the four event properties
250   (the target variable). Because the purpose of this ML exercise is qualitative, for brevity, we only present the results for the run using the event-mean precipitation rate, depicted in the insets shown in Figure 5. We present but overlook RF fits in which the resultant coefficient of determination (R^2) is negative, indicating a fit with no predictive skill.

All instrument pairs with the PWD are most influenced by the event-mean precipitation rate, suggesting the existence of some proportionality between the deviations and variable itself. This proportionality was also indicated in joint distributions we
255   tested, and the general dominance of the examined target variable with its deviation feature (estimated relative errors for that matter) was seen in the vast majority of cases (not shown). In those analyses, the other examined features typically showed very weak, if any, proportionality (not shown). The ORG (Figure 5b) and RAINWB (Figure 5d), while having the event-mean rate as the dominating feature, have the event 1-min maximum rate at 2nd place in importance, reflecting their inconsistent behavior discussed above. The TBRG (Figure 5a) is the only instrument for which the event total feature dominates the
260   deviations from the reference, interpreted as the impact of its coarse least count on precipitation event properties (this event total feature dominance is no longer seen when the exercise is applied to events with total amount greater than 1 mm; not shown). Interestingly, the event-mean temperature (2nd place) and RH (3rd place) features appear to have some influence over the event-mean precipitation rate. The event-mean RH seems to influence the VDISQUANTS (Figure 5c) to a relatively greater extent, and to some extent, the ORG (3rd place) and WBPLUVIO2 (4th place). The influence of the RH on comparisons using
265   different instruments with different operation mechanisms suggests that the PWD is susceptible to RH variability, numerous

precipitation instruments are susceptible to RH variability, or both of these options. (Note that this RH variability is likely related to the foggy conditions demonstrated above and potentially also to other conditions such as low RH — dry conditions). The WBPLUVIO2 has the wind speed in 3rd place, not far from the precipitation rate features in importance fraction (Figure 5h). This unique importance of wind speed (overlooking instruments with negative R^2) insinuates that wind effects are still

270 influencing the rain gauge, even though it is deployed together with a wind shield. We note that equivalent RF analyses using data from other ARM deployments, such as the Tracking Aerosol Convection Interactions Experiment (TRACER; Jensen et al., 2023; not shown), indicated a stronger influence of both wind direction and wind speed on multiple instruments. While additional quantitative characterization of instrument susceptibilities to deployment properties and conditions is essential, a deployment-dependent effort of this type requires a significant amount of work. The outcomes of such extensive efforts are

275 highly challenging to predict in advance and are considered less accurate. Therefore, these data characterization efforts often take place post-deployment, when the collected dataset is sufficiently large to produce substantial results (beyond the scope of this study).

Suppose one wishes to develop an operational, unbiased (or at least, bias-mitigated) precipitation best-estimate data product. In that case, given that they do not have a *true* benchmark, they need to be aware of all the factors described above by

280 performing robust characterization, which would ideally require a best-estimate product — this presents a conundrum. A first step towards resolving this conundrum would be to assume, given the evidence from this inter-comparison about instrument consistency, that the suite of ARM instruments measure some perturbations from the *true* value, such that their mean could serve as a best-estimate of the *actual* precipitation value, and that other statistics (e.g., minimum, maximum, and standard deviation) could be used to estimate confidence intervals. (Note that because the number of available instruments is typically

285 limited, the traditional 10th and 90th percentiles of a quantity as confidence intervals are of little meaning in this case). This approach serves as the basis for PrecipBE, ARM's best-estimate precipitation data product described and demonstrated in the section below.

3 The PrecipBE Algorithm

The PrecipBE VAP processing is performed on a per-precipitation-event basis, leveraging ARM measurement capabilities,

290 depending on instrument data availability per deployment, while considering QC samples and ARM data quality reports (DQRs). The processing workflow is described in the flowchart shown in Figure 6. Processing begins separately for each instrument. However, because given precipitation events can persist for more than a day or through 23:59:59 UTC of a given day, data from all available instruments are loaded for up to 7 days following the currently processed day, depending on whether a continuing event is indicated by one or more of the available instruments. This buffer data loading prevents

295 precipitation event biases driven by day-transition artifacts. Consistent with the inter-comparison discussed above, a continuing event suspect is identified if precipitation instances (precipitation amount sample greater than 0 mm) are detected by a given instrument less than 30 minutes from the end of the given day, i.e., after 23:30 UTC.
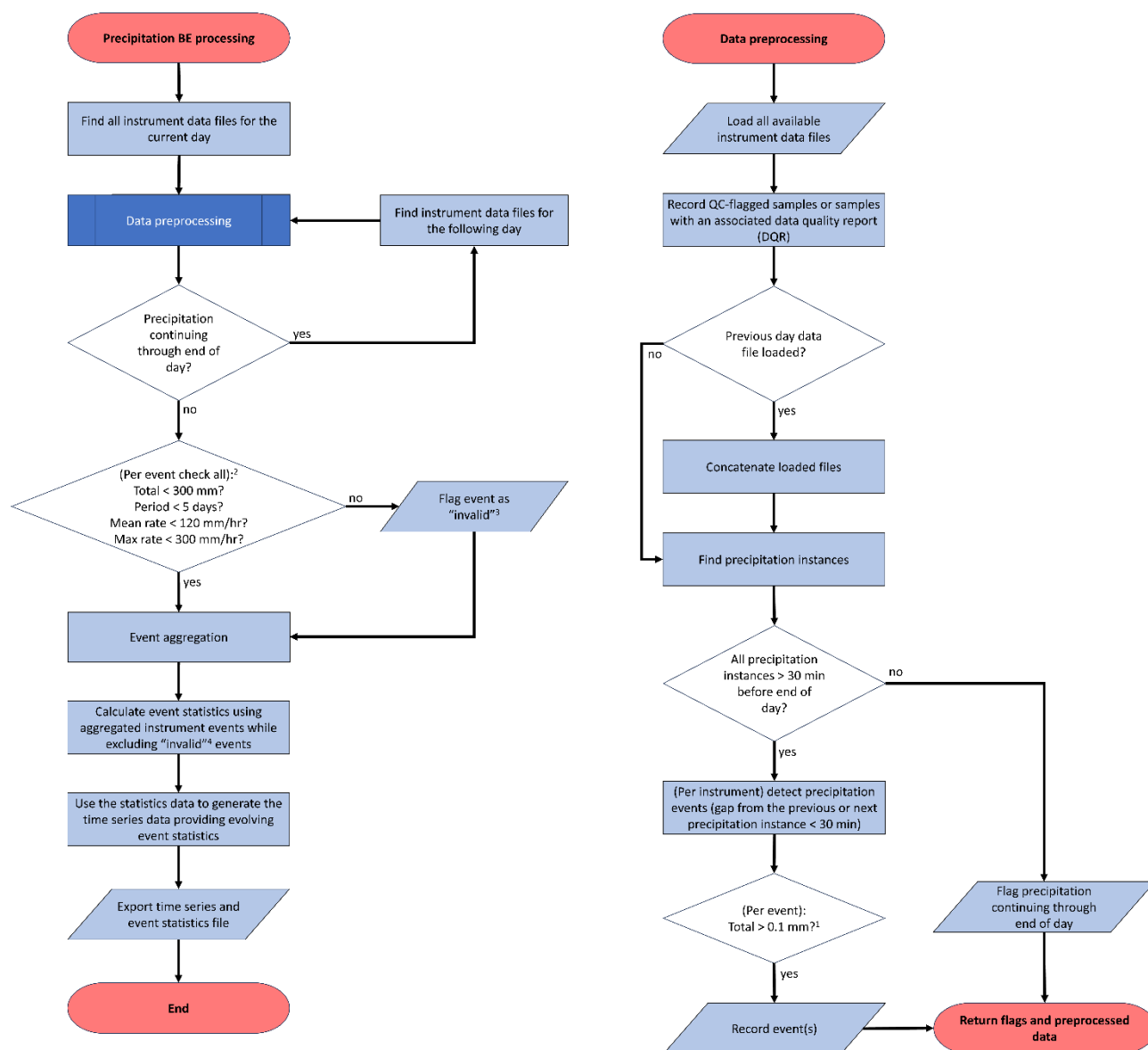
Figure 6: PrecipBE processing flowchart

300

Precipitation event processing is generally similar to the methodology discussed in Section 2.1 until the event aggregation step of the flowchart. Following the inter-comparison results, the 0.1 mm cumulative precipitation event minimum is applied to all instruments, except for the TBRG, in which case a 1.5 mm threshold is used due to its coarse measurement (equivalent to a

minimum effective error of ~8.5%), and because it is more prone to sporadic counts (not shown). In addition, given the
305   RAINWB and ORG biases demonstrated above, those two instruments are entirely omitted from the PrecipBE algorithm (see
Table 1).

During the aggregation stage, all valid instrument events are aggregated together while following the 30-min no-precipitation
logic discussed above. As such, if during the aggregation stage, a continuing event suspect (precipitation instances after 23:30
UTC by one or more instruments) is ultimately gapped by more than 30 min from the closest precipitation instance(s) during
310   the following day, the event is not a continuing event, and the loaded buffer day data are discarded.

The PrecipBE algorithm robustly addresses potential issues stemming from problematic data. Here, flagged events (events
with one or more QC samples or anomalous readings) or events with associated DQRs are not omitted before aggregation, as
in the comparison above. Instead, all events detected by a given instrument are still included in the aggregation stage to resolve
a PrecipBE event but are excluded from the PrecipBE event statistics calculations if one or more of them have one or more
315   problematic samples. For example, in the case of the diagram shown in Figure 1, the commonly occurring interlaced event
configuration will end in a single resolved PrecipBE event incorporating all four instruments (PWD, LDIS, VDIS, and TBRG),
regardless of whether one or more instrument events have problematic samples or an associated DQR. Assuming that all
instrument events are valid, all four of them will be included in the statistics calculation. However, assuming an issue with
TBRG event 3, for example, all TBRG events will be removed from the resolved PrecipBE event statistics, which will only
320   incorporate three instruments (PWD, LDIS, and VDIS). Assuming instead that LDIS event 4 has problematic samples,
PrecipBE will still resolve a single event, even if the period between the end of VDIS event 1 and the onset of PWD event 3
is greater than 30 min. In that case, statistics will be based on the PWD, VDIS, and TBRG events. We note that other
approaches, such as omitting those problematic events from the aggregation stage as well, were extensively tested and resulted
in significant PrecipBE event biases driven by the sporadic nature of anomalous samples across instruments (not shown). The
325   currently implemented approach, therefore, prevents event onset and ending inconsistency issues at the expense of fewer
incorporated instruments. This approach also served as the main incentive for excluding the RAINWB instrument from the
algorithm due to its substantial positive event period biases (see Figure 3d).

As suggested by the flowchart in Figure 6, once the resolved PrecipBE event statistics are calculated, they are used to generate
time series data, followed by the export of daily PrecipBE files, which are described below.

4 PrecipBE Dataset Structure and SGP Output Demonstration

PrecipBE includes two datastreams (data set types) streamlining both process understanding and model evaluation studies
using ARM surface precipitation data. The first datastream provides time series (evolving) precipitation data, whereas the
second includes per-event statistics in an easy-to-use one-dimensional (tabular) format. (PrecipBE data file structure and the
utilization of each of these datastreams is demonstrated in a Juypter notebook available on the ARM Notebooks Github
335   repository at: https://github.com/ARM-Development/ARM-Notebooks/blob/main/VAPs/precipbe/precipbe_intro.ipynb).

The time series datastream (`precipbetseries`) provides the temporally-evolving instrument-mean, minimum, maximum, and standard deviation of event-cumulative precipitation and 1-min precipitation rates. Each timestamp indicates the number of instruments used, and flags are provided for events detected using only a single instrument. The time series files also include bitwise flag arrays for instrument availability, invalid instrument samples, and instrument DQRs. Figure 7 shows an example

340   of the PrecipBE time series output for two events that started at the SGP site on November 8, 2024, with the second event ending just after 04:00 UTC of the following day. Note that the cumulative precipitation (top panel) zeros out after the end of the first event until the beginning of the second event, enabling straightforward, low overhead, analysis. For example, in the first depicted event, cumulative precipitation increases at a varying rate with a short burst around 09:45 UTC, during which a 1-min averaged precipitation rate exceeding 200 mm/h is observed by one of the instruments (lower panel), with very weak

345   and intermittent precipitation in the final 4.5 hours of this event.

The PrecipBE time series data suggest that none of the 7 available instruments were omitted from the statistics calculations of these two events due to flags, bogus samples, or existing DQRs. The time series data file provides information about which instruments were available via its bitwise `available_instruments` field — in this case, the SGP C1 facility's VDISQUANTS and LDQUANTS VAPs, DISDROMETER, and the WBPLUVIO2, and the SGP E13 facility's PWD,

350   AOSMET, and TBRG. However, while this datastream provides all available precipitation data converted to accumulated totals in 1-min increments (in units of mm/min), examining statistics of particular events, such as the two depicted in Figure 7 would require additional processing. Alternatively, one could use the PrecipBE statistics datastream (`precipbestats`) files, which are only generated for days with precipitation event onsets, having the number of timestamps equal to the number of precipitation events that started on a given day. In case of November 8, 2024, illustrated below, the corresponding statistics

355   data file includes two timestamps.

In each timestamp, `precipbestats` informs about statistics of the given event such as the instrument-mean, minimum, maximum, and standard deviation of onset, end time, period, total amount, mean precipitation rate, 1-min-averaged maximum precipitation rate, and precipitation rate standard deviation, as well as various flags and information such as which instrument recorded the highest precipitation rate or smallest total amount for that event. For example, the major precipitation event

360   depicted in Figure 7 resulted in a cumulative amount of ~54 ± 9 mm with an instrument minimum and maximum of 43 and 75 mm, respectively. The statistics data file indicates that the DISDROMETER recorded the maximum precipitation rate of 218 mm/h during that event. In comparison, the instrument-mean maximum precipitation rate was more moderate yet still rather intense at 85 mm/h. Finally, the statistics dataset contains statistical information about the surface-level atmospheric state during precipitation events, with data harvested from (in order of preference) the MET, the automatic weather station (MAWS;

365   Holdridge and Kyrouac, 2017), or one of the Vaisala WXT systems operated by ARM (see Table 1), as well as drop distribution moment data derived using the VDISQUANTS or LDQUANTS, depending on availability. For example, the surface temperature during the major November 8, 2024, event ranged between 9.3° and 13.6° C, with an average of 10.0° C, while the event-mean RH was 97.5%. The even-mean liquid water content derived by the VDISQUANTS VAP was 0.3 g/m^3, and the average mass-weighted mean drop diameter was ~1.5 mm.

370
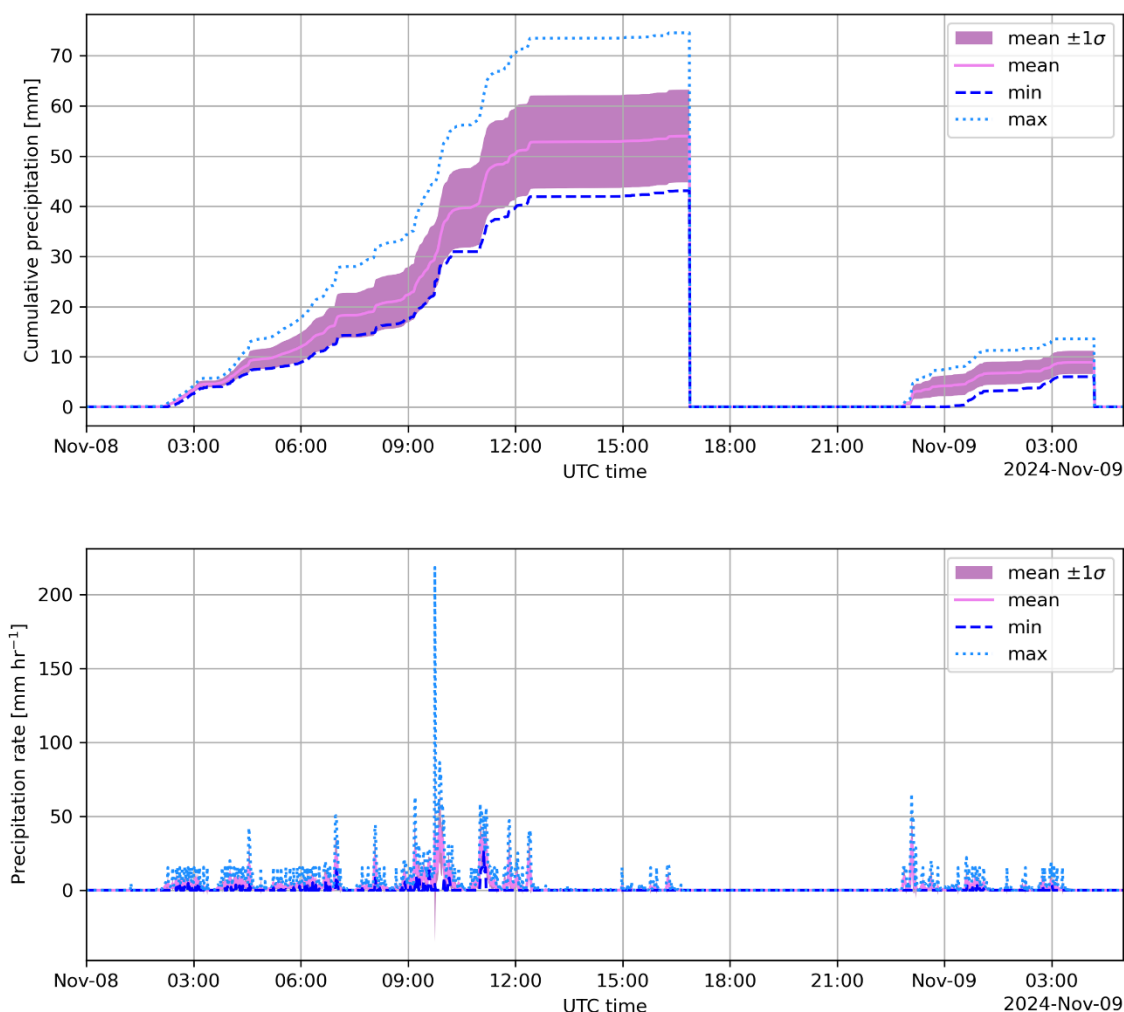
Figure 7: PrecipBE time series of two precipitation events that occurred on November 8, 2024, at the ARM SGP site, the second of which continued into November 9. (Top) Per-event cumulative precipitation, (bottom) precipitation rates. The plot illustrates the instrument-minimum, maximum, mean, and mean ± standard deviation (σ) (see legend).

375

5 Long-term Trend Analysis of PrecipBE Output for the ARM SGP Site

Using PrecipBE statistics data files generated for the SGP site, spanning September 2, 1993, to March 4, 2025 (~31.5 years), we can easily examine precipitation event trends at the ARM site. Figure 8 shows running-mean time series data that facilitates basic trend analysis. We depict both curves calculated using the full dataset and curves calculated using a data subset derived

380   only from multi-instrument events. Ideally, one should be inclined to use precipitation event properties and statistics derived from more than one instrument, as they are considered more robust than those based on a single instrument. However, ARM operated only the TBRG starting in September, 1993, ~14 months after the SGP site launch, until April, 2006, when the DISDROMETER was deployed as the first addition to the growing suite of precipitation instruments ARM operates at the site. The results of the instrument inter-comparison in Section 2 indicated that the TBRG is generally consistent with other advanced

385   precipitation instruments in event totals. It is also consistent with other instruments in event precipitation rates, as long as it is conditioned for event total greater than its least count by some factor (e.g., effective uncertainties of 12.7% and 8.5% at event total of 1 and 1.5 mm, respectively). We follow these conclusions to derive the statistics depicted in Figure 8, which are also part of the motivation to examine 1-year-windows.

The 1-year running sum (annual) precipitation record (Figure 8a) largely shows little difference in annual amount between the

390   full dataset and multi-instrument subset, with annual means of ~800 and ~750 mm (respectively) in agreement with previous studies (cf. Sisterson et al., 2016). The SGP annual rainfall is quite variable, with some years in which the site experienced significant amounts (e.g., 2008 and 2019 exceeding 1100 mm), and others when the site exhibited small amounts (e.g., below 400 mm in 2006 and 2011). Statistically significant linear fits suggest a decadal increase in annual rainfall of more than 36 mm per decade (~5%). Those positive rainfall trends are consistent with studies that examined trends over the south-central

395   US, where the SGP site is located (e.g., Harp and Horton, 2022; Sun et al., 2021). The number of significant precipitation events, referred to here as events with totals exceeding 1 mm, tentatively suggests a statistically significant increasing trend (Figure 8b), commensurate with ~18 min (~7%) decadal reduction in event period (not shown). Here, the higher event total amount threshold mitigates the positive (negative) bias in the number of events (event period) in the earlier years of the SGP site, when the TBRG was the only operating precipitation-measuring instrument, such that event properties are strongly

400   influenced by the TBRG's tendencies discussed in Section 2.2. Yet, between the full dataset and multi-instrument subset during overlapping periods, a limited positive bias is still observed in the case of the number of events (Figure 8b). Therefore, all else being equal, it is more likely that the decadal trend leans towards the multi-instrument subset, with an increasing trend in the number of events on the order of 10 more events per year per decade. Given the definition of precipitation events in PrecipBE (precipitation instances gapped by less than 30 min), these results could indicate a growing tendency to more precipitation

405   from broken cloud systems, which could be related to observed trends and feedbacks (e.g., Goessling et al., 2025; Loeb et al., 2024; Sherwood et al., 2020; Song et al., 2023), yet additional research using PrecipBE and other datasets is required to support this hypothesis.
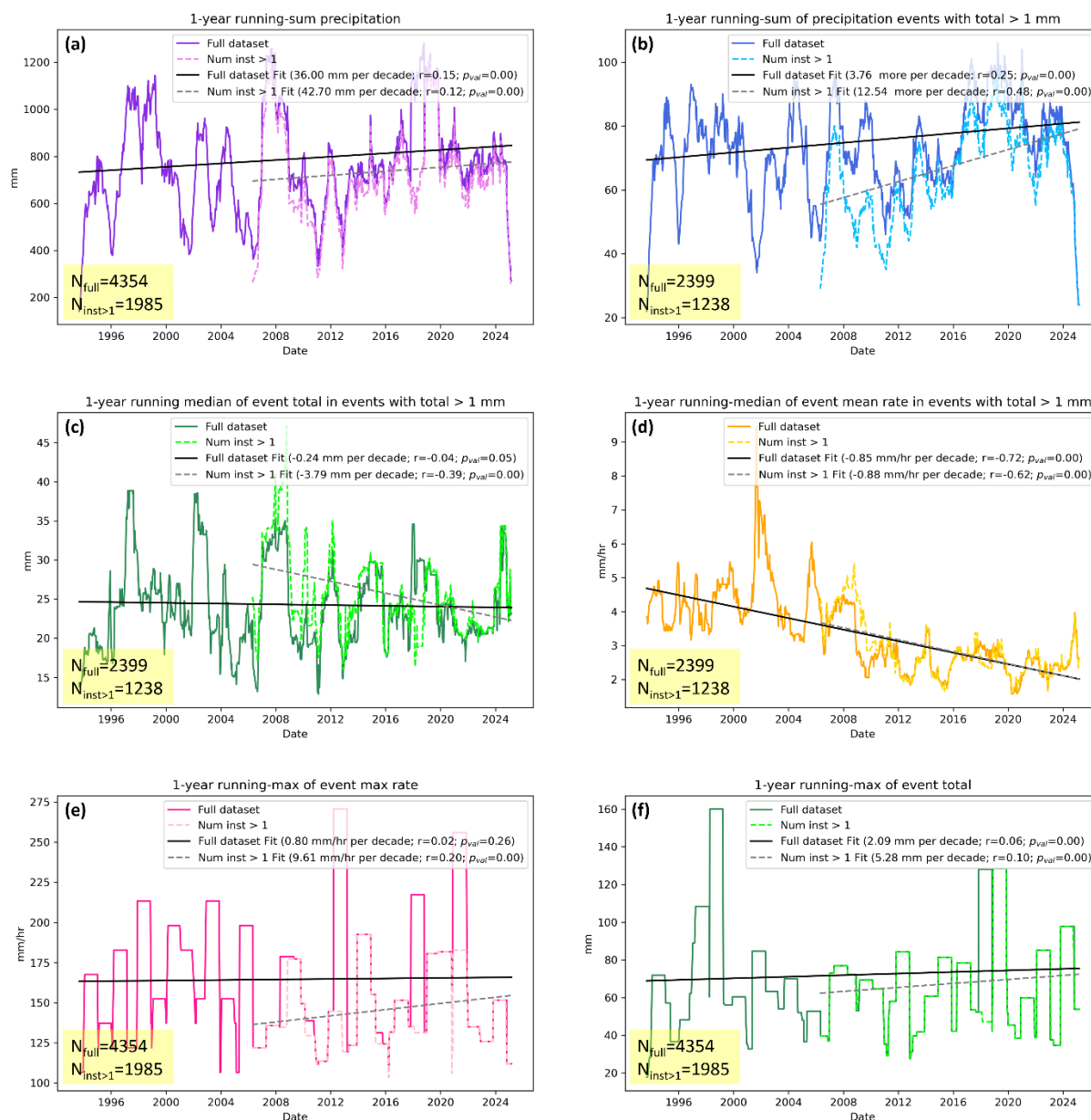
Figure 8: Long-term trends in PrecipBE precipitation event properties for the ARM SGP site between September 2, 1993, and
410 **March 4, 2025. The solid curves were generated using all of the available precipitation event statistics, whereas the dashed curves were generated using precipitation events detected by two or more instruments (first effective sample on April 11, 2006, with the addition of the DISDROMETER). (a) 1-year running-sum (annual) precipitation totals, (b) 1-year running sum (annual) number of precipitation events with total > 1 mm, (c) 1-year running median of precipitation event total in events with total > 1 mm, (d) 1-year running median of event-mean precipitation rate in events with total > 1 mm, (e) 1-year running maximum of event 1-min-averaged**
415 **maximum precipitation rate, and (f) 1-year running maximum of precipitation event total. The solid black and dashed grey lines denote linear fits to the full dataset and the multi-instrument subset, respectively. Decadal trends, correlation coefficients, and P-values are given in the legends. All quantities were calculated using the instrument-mean data. The total number of samples (precipitation events) used in the illustrated curves are given the bottom left corner of each panel.**

From a bulk perspective, all else being equal, the reduction in the precipitation event period can be translated, on average, to a trending decrease in event totals, which is indeed suggested from Figure 8c, consistent in both the full dataset and the multi-instrument subset. Following the same logic, one might expect an increasing average precipitation rate, but the 1-year running median of the event-mean precipitation rates indicates a statistically significant decreasing trend, consistent between the full and multi-instrument datasets (Figure 8d). By examining event period in the multi-instrument subset, thereby mitigating the effect of the TBRG's negative event period bias (e.g., Figure 3a), a 1-year running sum (annual) precipitation time (not shown) indicates a minimal and statistically insignificant reduction. Therefore, these results raise an apparent inconsistency between higher annual rainfall and shorter yet less intense events, on average. However, this inconsistency can be reconciled via examination of precipitation extremes in a 1-year timeframe of event maximum 1min-averaged maximum precipitation rates and maximum event totals (Figure 8e and Figure 8f, respectively). These curves exhibit general consistency between the full dataset and the multi-instrument subset and all except the maximum precipitation rate using the full dataset indicate a statistically significant increasing trend in both metrics: more than 9 mm/h per decade increase in maximum precipitation rate (6.5%) using the multi-instrument subset and between 2-5 mm per decade increase in extreme event totals (~3–7.5%) over a 1-year timeframe. Taken together, this precipitation event trend analysis indicates that the observed increase in annual rainfall could be catalyzed by a few more extreme precipitation events taking place at the SGP site. Examination of the causal sources of these trends via counterfactual exercises and their attribution to potential drivers such as regional natural variability (e.g., Higgins et al., 2007; McKinnon and Deser, 2021) or changes to the local land use (e.g., Krishnamurthy et al., 2025) remain a topic of future studies.

6 Conclusions and Outlook

In this study, we presented an analysis of differences in ARM precipitation instrument measurements from a unique per-event perspective. Supported by a ML application to the instrument differences to examine the importance of various atmospheric state variables and parameters, the analysis indicates that, by and large, most ARM instruments are generally consistent with each other. Yet, deviations, occasionally of significant magnitudes, often occur, and could be driven by specific parameters such as relative humidity and wind properties, which could be deployment-dependent, or by differing instrument response functions to the same parameters those instruments are aimed at measuring (e.g., precipitation rates). Without additional prior knowledge, these results suggest that, on a first-order basis, the best estimate of precipitation properties is ostensibly that which incorporates all available valid data, which motivates the design of the PrecipBE value-added product (VAP). That said, while the analysis showed that specific instruments show some tendency for certain behaviors, such as shorter precipitation event periods in the case of the TBRG and WBPLUVIO2, other instruments, specifically, the RAINWB and ORG exhibit clear and significant biases, which cannot be ameliorated and therefore integrated into PrecipBE. Fortunately, ARM retired the RAINWB several years ago, and the ORG is in the process of being retired in 2025.

450 PrecipBE provides time series and tabular statistics datasets that are easy to use and comprehensive, including precipitation event properties, and are supplemented with ancillary data from various ARM datasets. Therefore, it is likely that this VAP would become the baseline (go-to) precipitation product for the ARM user community, augmenting the derivation of scientific insights and streamlining model evaluation. Those features of this VAP were demonstrated via the examination of a single-day output as well as a long-term trend analysis of precipitation events at the ARM SGP site. The trend analysis tentatively

455 suggests mainly shorter and less intense precipitation events at the SGP site, but also a long-term increase in annual rainfall driven primarily by more extreme event properties (event totals and maximum precipitation rates) of relatively rare, highly intense precipitation events. While we believe that numerous additional insights about surface precipitation at the SGP and other ARM sites can be derived via conditioning on various metrics related to drop size distribution moments, temperatures, diurnal cycle, time of year, etc. provided in the PrecipBE data files, we leave such analyses for the ARM user community.

460 PrecipBE will soon become an operational product with a several-day lag from real-time, and hence, its datasets will be continuously updated and made available via the ARM Data Discovery (https://adc.arm.gov/discovery). Future planned VAP updates include the addition of solid precipitation properties at applicable sites and the potential integration of radar-based low-level precipitation estimates. We invite the ARM user community to leverage PrecipBE and provide feedback to further enhance this new and exciting data product.

465 **Data Availability**

Current and future releases of PrecipBE time series (Silber, 2025c, d) and statistics datasets (Silber, 2025a, b) are and will be available on the ARM Data Discovery (https://adc.arm.gov/discovery/#/results/s::precipbe). A Jupyter notebook demonstrating the structure and application of PrecipBE datasets is available on the ARM Notebooks Github repository at: https://github.com/ARM-Development/ARM-Notebooks/blob/main/VAPs/precipbe/precipbe_intro.ipynb. The precipitation

470 datasets of the PWD (Kyrouac et al., 2021), AOSMET (Kyrouac and Tuftedal, 2010), DISDROMETER (Wang, 2006), VDISQUANTS (Hardin et al., 2021), LDQUANTS (Hardin et al., 2019), TBRG (Kyrouac et al., 2006, 2021), WBPLUVIO2 (Zhu et al., 2016), RAINWB (Shi et al., 2010), and ORG (Kyrouac et al., 2021) from the ARM SGP site are available on the ARM Data Discovery (https://adc.arm.gov/discovery/; last access: 10 March 2025).

Author Contribution

475 Conceptualization: IS, JMC, and AT

Formal analysis, investigation, methodology, visualization, and original draft preparation: IS

Project administration: JMC

Data curation and validation: IS and MRK

Manuscript review and editing: all authors

Competing Interests

The authors declare that they have no competing interests

Acknowledgments

The authors thank Scott Collis and Scott Giangrande for valuable comments and feedback. Data were obtained from the ARM user facility, a U.S. Department of Energy (DOE) Office of Science user facility managed by the Biological and Environmental Research (BER) program.

Financial Support

This research was supported by the ARM user facility, a U.S. Department of Energy (DOE) Office of Science user facility managed by the Biological and Environmental Research (BER) program. Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle under contract DE-AC05-76RL01830.

References

Bartholomew, M. J.: Impact Disdrometers Instrument Handbook, ARM Climate Research Facility, Pacific Northwest National Laboratory, Richland, WA, United States, https://doi.org/10.2172/1251384, 2016a.

Bartholomew, M. J.: Optical Rain Gauge Instrument Handbook, ARM Climate Research Facility, Pacific Northwest National Laboratory, Richland, WA, United States, https://doi.org/10.2172/1251388, 2016b.

Bartholomew, M. J.: Rain Gauges Handbook, https://doi.org/10.2172/1245982, 2016c.

Bartholomew, M. J.: Weighing Bucket Rain Gauge Instrument Handbook, United States, https://doi.org/10.2172/1572341, 2019.

Bartholomew, M. J.: Laser Disdrometer Instrument Handbook, DOE ARM Climate Research Facility, Pacific Northwest National Laboratory, Richland, Washington, United States, https://doi.org/10.2172/1226796, 2020a.

Bartholomew, M. J.: Two-Dimensional Video Disdrometer (VDIS) Instrument Handbook, ARM Climate Research Facility, Pacific Northwest National Laboratory, Richland, WA, United States, https://doi.org/10.2172/1251384, 2020b.

Battaglia, A., Rustemeier, E., Tokay, A., Blahak, U., and Simmer, C.: PARSIVEL Snow Observations: A Critical Assessment, J. Atmospheric Ocean. Technol., 27, 333–344, https://doi.org/10.1175/2009JTECHA1332.1, 2010.

Bretherton, C. S., Uchida, J., and Blossey, P. N.: Slow Manifolds and Multiple Equilibria in Stratocumulus-Capped Boundary Layers, J. Adv. Model. Earth Syst., 2, https://doi.org/10.3894/JAMES.2010.2.14, 2010.

Christensen, M. W., Wu, P., Varble, A. C., Xiao, H., and Fast, J. D.: Aerosol-induced closure of marine cloud cells: enhanced effects in the presence of precipitation, Atmospheric Chem. Phys., 24, 6455–6476, https://doi.org/10.5194/acp-24-6455-2024, 2024.

Ciach, G. J.: Local Random Errors in Tipping-Bucket Rain Gauge Measurements, J. Atmospheric Ocean. Technol., 20, 752–759, https://doi.org/10.1175/1520-0426(2003)20<752:LREITB>2.0.CO;2, 2003.

Colli, M., Lanza, L. G., and La Barbera, P.: Performance of a weighing rain gauge under laboratory simulated time-varying reference rainfall rates, Perspect. Precip. Sci. - Part I, 131, 3–12, https://doi.org/10.1016/j.atmosres.2013.04.006, 2013.

Emmenegger, T., Kuo, Y.-H., Xie, S., Zhang, C., Tao, C., and Neelin, J. D.: Evaluating Tropical Precipitation Relations in CMIP6 Models with ARM Data, J. Clim., 35, 6343–6360, https://doi.org/10.1175/JCLI-D-21-0386.1, 2022.

Fehlmann, M., Rohrer, M., von Lerber, A., and Stoffel, M.: Automated precipitation monitoring with the Thies disdrometer: biases and ways for improvement, Atmospheric Meas. Tech., 13, 4683–4698, https://doi.org/10.5194/amt-13-4683-2020, 2020.

Goessling, H. F., Rackow, T., and Jung, T.: Recent global temperature surge intensified by record-low planetary albedo, Science, 387, 68–73, https://doi.org/10.1126/science.adq7280, 2025.

Hardin, J., Giangrande, S., and Zhou, A.: ldquants, https://doi.org/10.5439/1432694, 2019.

Hardin, J., Giangrande, S., and Zhou, A.: Laser Disdrometer Quantities (LDQUANTS) and Video Disdrometer Quantities (VDISQUANTS) Value-Added Products Report, DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States), https://doi.org/10.2172/1808573, 2020.

Hardin, J., Giangrande, S., Fairless, T., and Zhou, A.: vdisquants: Video Distrometer derived radar equivalent quantities. Retrievals from the VDIS instrument providing radar equivalent quantities, including dual polarization radar quantities (e.g., Z, Differential Reflectivity ZDR)., https://doi.org/10.5439/1592683, 2021.

Harp, R. D. and Horton, D. E.: Observed Changes in Daily Precipitation Intensity in the United States, Geophys. Res. Lett., 49, e2022GL099955, https://doi.org/10.1029/2022GL099955, 2022.

Higgins, R. W., Silva, V. B. S., Shi, W., and Larson, J.: Relationships between Climate Variability and Fluctuations in Daily Precipitation over the United States, J. Clim., 20, 3561–3579, https://doi.org/10.1175/JCLI4196.1, 2007.

Holdridge, D. and Kyrouac, J.: ARM: ARM-standard Meteorological Instrumentation, Marine, https://doi.org/10.5439/1095605, 2012.

Holdridge, D. J. and Kyrouac, J. A.: Meteorological Automatic Weather Station (MAWS) Instrument Handbook, https://doi.org/10.2172/1373930, 2017.

Howie, J., Protat, A., Kyrouac, J., and Tuftedal, M.: Australian Bureau of Meteorology surface meteorology data, https://doi.org/10.5439/1597382, 2016.

Jensen, M. P., Flynn, J. H., Kollias, P., Kuang, C., McFarquhar, G., Powers, H., Brooks, S., Bruning, E., Collins, D., Collis, S. M., Fan, J., Fridlind, A., Giangrande, S. E., Griffin, R., Hu, J., Jackson, R. C., Kumjian, M., Logan, T., Matsui, T., Nowotarski, C. J., Oue, M., Rapp, A. D., Rosenfeld, D., Ryzhkov, A., Sheesley, R., Snyder, J., Stier, P., Usenko, S., van den Heever, S. C., van Lier-Walqui, M., Varble, A., Wang, Y., Aiken, A., Deng, M., Dexheimer, D., Dubey, M., Feng, Y., Ghate,

V. P., Johnson, K. L., Lamer, K., Saleeby, S. M., Wang, D., Zawadowicz, M. A., and Zhou, A.: Tracking Aerosol Convection Interactions Experiment (TRACER) Field Campaign Report, United States, https://doi.org/10.2172/2202672, 2023.

Koralegedara, S. B., Lin, C.-Y., and Sheng, Y.-F.: Numerical Analysis of the Mesoscale Dynamics of an Extreme Rainfall and Flood Event in Sri Lanka in May 2016, J. Meteorol. Soc. Jpn. Ser II, 97, 821–839, https://doi.org/10.2151/jmsj.2019-046, 2019.

Krishnamurthy, R., Newsom, R. K., Kaul, C. M., Letizia, S., Pekour, M., Hamilton, N., Chand, D., Flynn, D., Bodini, N., and Moriarty, P.: Observations of wind farm wake recovery at an operating wind farm, Wind Energy Sci., 10, 361–380, https://doi.org/10.5194/wes-10-361-2025, 2025.

Kyrouac, J.: Aerosol Observing System Surface Meteorology (AOSMET) Instrument Handbook, United States, https://doi.org/10.2172/1573797, 2019a.

Kyrouac, J.: Precipitation Meteorological Instruments (PRECIPMET) Handbook, https://doi.org/10.2172/2575103, 2019b.

Kyrouac, J. and Shi, Y.: metwxt, https://doi.org/10.5439/1455447, 2018.

Kyrouac, J. and Tuftedal, M.: pws (b1), https://doi.org/10.5439/1992050, 2001.

Kyrouac, J. and Tuftedal, M.: aosmet (a1), https://doi.org/10.5439/1984920, 2010.

Kyrouac, J. and Tuftedal, M.: Surface Meteorological System (MET) Instrument Handbook, Atmospheric Radiation Measurement user facility, Richland, WA., United States, https://doi.org/10.2172/1007926, 2024.

Kyrouac, J., Shi, Y., Jane, M., and Wang, D.: raintb.b1, https://doi.org/10.5439/1224827, 2006.

Kyrouac, J., Shi, Y., and Tuftedal, M.: precipmet, https://doi.org/10.5439/1353192, 2017.

Kyrouac, J., Shi, Y., and Tuftedal, M.: met.b1, https://doi.org/10.5439/1786358, 2021.

Lagouvardos, K., Kotroni, V., Defer, E., and Bousquet, O.: Study of a heavy precipitation event over southern France, in the frame of HYMEX project: Observational analysis and model results using assimilation of lightning, Atmospheric Res., 134, 45–55, https://doi.org/10.1016/j.atmosres.2013.07.003, 2013.

Loeb, N. G., Ham, S.-H., Allan, R. P., Thorsen, T. J., Meyssignac, B., Kato, S., Johnson, G. C., and Lyman, J. M.: Observational Assessment of Changes in Earth's Energy Imbalance Since 2000, Surv. Geophys., 45, 1757–1783, https://doi.org/10.1007/s10712-024-09838-8, 2024.

Martin, S. T., Artaxo, P., Machado, L., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J., Biscaro, T., Brito, J., Calheiros, A., Jardine, K., Medeiros, A., Portela, B., Sá, S. S. de, Adachi, K., Aiken, A. C., Albrecht, R., Alexander, L., Andreae, M. O., Barbosa, H. M. J., Buseck, P., Chand, D., Comstock, J. M., Day, D. A., Dubey, M., Fan, J., Fast, J., Fisch, G., Fortner, E., Giangrande, S., Gilles, M., Goldstein, A. H., Guenther, A., Hubbe, J., Jensen, M., Jimenez, J. L., Keutsch, F. N., Kim, S., Kuang, C., Laskin, A., McKinney, K., Mei, F., Miller, M., Nascimento, R., Pauliquevis, T., Pekour, M., Peres, J., Petäjä, T., Pöhlker, C., Pöschl, U., Rizzo, L., Schmid, B., Shilling, J. E., Dias, M. A. S., Smith, J. N., Tomlinson, J. M., Tóta, J., and Wendisch, M.: The Green Ocean Amazon Experiment (GoAmazon2014/5) Observes Pollution Affecting Gases, Aerosols, Clouds, and Rainfall over the Rain Forest, Bull. Am. Meteorol. Soc., 98, 981–997, https://doi.org/10.1175/BAMS-D-15-00221.1, 2017.

575 Mather, J.: Atmospheric Radiation Measurement (ARM) Management Plan, DOE ARM User Facility, Pacific Northwest National Laboratory, Richland, WA, United States, https://doi.org/10.2172/1253897, 2024.

Mather, J. H., Turner, D. D., and Ackerman, T. P.: Scientific Maturation of the ARM Program, Meteorol. Monogr., 57, 4.1-4.19, https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0053.1, 2016.

McKinnon, K. A. and Deser, C.: The inherent uncertainty of precipitation variability, trends, and extremes due to internal
580 variability, with implications for Western US water resources, J. Clim., 1–46, https://doi.org/10.1175/JCLI-D-21-0251.1, 2021.

Mikkelsen, A., McCoy, D. T., Eidhammer, T., Gettelman, A., Song, C., Gordon, H., and McCoy, I. L.: Constraining Aerosol-Cloud Adjustments by Uniting Surface Observations with a Perturbed Parameter Ensemble, https://doi.org/10.5194/egusphere-2024-2158, 30 August 2024.

Milewska, E. J., Vincent, L. A., Hartwell, M. M., Charlesworth, K., and Mekis, É.: Adjusting precipitation amounts from
585 Geonor and Pluvio automated weighing gauges to preserve continuity of observations in Canada, Can. Water Resour. J. Rev. Can. Ressour. Hydr., 44, 127–145, https://doi.org/10.1080/07011784.2018.1530611, 2019.

Montero-Martínez, G. and García-García, F.: On the behaviour of raindrop fall speed due to wind, Q. J. R. Meteorol. Soc., 142, 2013–2020, https://doi.org/10.1002/qj.2794, 2016.

Montero-Martínez, G., Torres-Pérez, E. F., and García-García, F.: A comparison of two optical precipitation sensors with
590 different operating principles: The PWS100 and the OAP-2DP, Atmospheric Res., 178–179, 550–558, https://doi.org/10.1016/j.atmosres.2016.05.007, 2016.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

595 Raupach, T. H. and Berne, A.: Correction of raindrop size distributions measured by Parsivel disdrometers, using a two-dimensional video disdrometer as a reference, Atmospheric Meas. Tech., 8, 343–365, https://doi.org/10.5194/amt-8-343-2015, 2015.

Reynolds, M., Ermold, B., and Shi, Y.: Marine Precipitation Measurements, https://doi.org/10.5439/1372168, 2017.

Ritsche, M. T.: Surface and Tower Meteorological Instrumentation at Barrow (METTWR4H) Handbook,
600 https://doi.org/10.2172/1020564, 2008.

Ro, Y., Chang, K.-H., Hwang, H., Kim, M., Cha, J.-W., and Lee, C.: Comparative study of rainfall measurement by optical disdrometer, tipping-bucket rain gauge, and weighing precipitation gauge, Nat. Hazards, 120, 2829–2845, https://doi.org/10.1007/s11069-023-06308-z, 2024.

Saha, R., Testik, F. Y., and Testik, M. C.: Assessment of OTT Pluvio2 Rain Intensity Measurements, J. Atmospheric Ocean.
605 Technol., 38, 897–908, https://doi.org/10.1175/JTECH-D-19-0219.1, 2021.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., Von Der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Rev. Geophys., 58,
610 e2019RG000678, https://doi.org/10.1029/2019RG000678, 2020.

Shi, Y., Jane, M., and Wang, D.: rainwb.b1, https://doi.org/10.5439/1224830, 2010.

Silber, I.: precipbestats (c0), https://doi.org/10.5439/2523641, 2025a.

Silber, I.: precipbestats (c1), https://doi.org/10.5439/2523640, 2025b.

Silber, I.: precipbetseries (c0), https://doi.org/10.5439/2523643, 2025c.

615 Silber, I.: precipbetseries (c1), https://doi.org/10.5439/2523642, 2025d.

Sisterson, D. L., Peppler, R. A., Cress, T. S., Lamb, P. J., and Turner, D. D.: The ARM Southern Great Plains (SGP) Site, Meteorol. Monogr., 57, 6.1-6.14, https://doi.org/10.1175/AMSMONOGRAPHS-D-16-0004.1, 2016.

Song, H., Choi, Y.-S., and Kang, H.: Global change in cloud radiative effect revealed in CERES observations, https://doi.org/10.21203/rs.3.rs-3781529/v1, 28 December 2023.

620 Sun, Q., Zhang, X., Zwiers, F., Westra, S., and Alexander, L. V.: A Global, Continental, and Regional Analysis of Changes in Extreme Precipitation, J. Clim., 34, 243–258, https://doi.org/10.1175/JCLI-D-19-0892.1, 2021.

Tokay, A., Petersen, W. A., Gatlin, P., and Wingo, M.: Comparison of Raindrop Size Distribution Measurements by Collocated Disdrometers, J. Atmospheric Ocean. Technol., 30, 1672–1690, https://doi.org/10.1175/JTECH-D-12-00163.1, 2013.

Tokay, A., Wolff, D. B., and Petersen, W. A.: Evaluation of the New Version of the Laser-Optical Disdrometer, OTT Parsivel2,
625 J. Atmospheric Ocean. Technol., 31, 1276–1288, https://doi.org/10.1175/JTECH-D-13-00174.1, 2014.

Tokay, A., D'Adderio, L. P., Marks, D. A., Pippitt, J. L., Wolff, D. B., and Petersen, W. A.: Comparison of Raindrop Size Distribution between NASA's S-Band Polarimetric Radar and Two-Dimensional Video Disdrometers, J. Appl. Meteorol. Climatol., 59, 517–533, https://doi.org/10.1175/JAMC-D-18-0339.1, 2020.

Wang, D.: disdrometer (b1), https://doi.org/10.5439/1987821, 2006.

630 Wang, D., Bartholomew, M. J., Giangrande, S. E., and Hardin, J. C.: Analysis of Three Types of Collocated Disdrometer Measurements at the ARM Southern Great Plains Observatory, United States, https://doi.org/10.2172/1828172, 2021.

Yuter, S. E., Kingsmill, D. E., Nance, L. B., and Löffler-Mang, M.: Observations of Precipitation Size and Fall Speed Characteristics within Coexisting Rain and Wet Snow, J. Appl. Meteorol. Climatol., 45, 1450–1464, https://doi.org/10.1175/JAM2406.1, 2006.

635 Zhang, L., Dong, X., Kennedy, A., Xi, B., and Li, Z.: Evaluation of NASA GISS post-CMIP5 single column model simulated clouds and precipitation using ARM Southern Great Plains observations, Adv. Atmospheric Sci., 34, 306–320, https://doi.org/10.1007/s00376-016-5254-4, 2017.

Zhu, Z., Wang, D., Jane, M., Cromwell, E., Sturm, M., Irving, K., and Delamere, J.: wbpluvio2.a1, https://doi.org/10.5439/1338194, 2016.

640