Assessment of OTT Pluvio^2 Rain Intensity Measurements

RUPAYAN SAHA,[a] FIRAT Y. TESTIK,[a] AND MURAT C. TESTIK[b]

[a] *Civil and Environmental Engineering Department, University of Texas at San Antonio, San Antonio, Texas*
[b] *Industrial Engineering Department, Hacettepe University, Ankara, Turkey*

(Manuscript received 23 December 2019, in final form 4 February 2021)

ABSTRACT: This study investigates the OTT Pluvio^2 weighing precipitation gauge's random and systematic error components as well as stabilization of the measurements on time-varying rainfall intensities (RI) under laboratory conditions. A highly precise programmable peristaltic pump that provided both constant and time-varying RI was utilized in the experiments. Abrupt, gradual step, and cyclic step changes in the RI values were evaluated. RI readings were taken in real time (RT) at different time resolutions (6–60 s) for the RI range of 6–70 mm h^{-1}. Our analysis indicates that the lower threshold for the OTT Pluvio^2's real-time RI measurements should be redefined as 7 mm h^{-1} at a 1-min resolution. Tolerance intervals containing 95% of the repeated measurements with a probability of 0.95 are given. It is shown that the measurement variances are unequal over the range of RI and the measurement repeatability is not uniform. A statistically significant negative bias was observed for the RI values of 7 and 8 mm h^{-1}, while there was not a statistically significant linearity problem. Through the use of statistical control limits, it is shown that means of the RI measurements stabilized on the actual RI value. A detailed investigation on RT bucket weight measurements revealed a time delay in bucket weight measurements, which causes notable errors in reported RI measurements under dynamic rainfall conditions. To demonstrate the potentiality of large errors in Pluvio^2's real-time RI measurements, a set of equations was developed that faithfully reproduced the Pluvio^2's internal (hidden) algorithm, and results from dynamic laboratory and in situ rainfall scenarios were simulated. The results of this investigation show the necessity of modifying the present Pluvio^2 RI algorithm to enhance its performance and show the possibility of postprocessing the existing Pluvio^2 RI datasets for improved measurement accuracies.

KEYWORDS: Atmosphere; Precipitation; In situ atmospheric observations; Instrumentation/sensors; Measurements; Error analysis

1. Introduction

Accurate and precise measurement of rainfall intensity (RI) is of significant importance for various meteorological and hydrological endeavors such as weather pattern change predictions, quantitative precipitation forecasts (QPFs), derivation of rainfall intensity–duration–frequency (IDF) curves, and rainwater harvesting planning [see the comprehensive volume on rainfall by Testik and Gebremichael (2010)]. Several RI measurement studies focused on extreme events of high intensity rainfall (Hou et al. 2014; Lanza and Vuerich 2009; Larsen et al. 2009; Westra et al. 2014; Willems 2000). However, studies (Karl and Knight 1998; Kidd and Joe 2007) have also highlighted the contributions of light (RI <2.5 mm h^{-1}) and moderate ($2.5 \leq$ RI ≤ 7.5 mm h^{-1}) rainfall to overall rainfall occurrence and accumulations. Kidd and Joe (2007) showed that light rainfall occurs for about 50% and 80% of the total rain duration in the tropics and in Europe, accounting for 10% and less than 45% of the accumulated rainfall therein, respectively. Moreover, a series of data from four Micro Rain Radars (24 GHz) for a 10-month observation period showed that nearly 80% of the total rain duration over the midlatitude regions is light and moderate rainfall. While contributions from light and moderate rainfall to the total rainfall accumulation of any particular event is often less than the contribution of heavy rainfall, light and moderate rainfall can be significant in terms of the total rain duration and frequency (Karl and Knight 1998). Hence, the identification and measurement of light and moderate rainfall are essential to ensure that the full range of intensities is retrieved and changes in rain intensities are accurately assessed.

The OTT Pluvio^2, developed in 2008 as the successor of Pluvio^1, which was launched in 1994, has become one of the most popular weighing rain gauges due to its ease of application and reliability in all weather conditions (Nitu et al. 2019). It has been installed throughout Europe, including remote locations for which frequent visits would be impractical, and it has been replacing older technologies as the standard instrument for precipitation measurements in the United States and Canada. Over the past couple of decades, several government agencies, such as the National Weather Service (NWS) and the U.S. Geological Survey (USGS), have completed extensive testing programs on several rain gauges, including Pluvio^1, to assist the National Atmospheric Deposition Program (NADP) in determining the appropriate precipitation gauge for the National Trends Network (NTN). The detailed outlines of NADP and NTN are provided by Lamb and Bowersox (2000) and Dossett and Bowersox (1999), respectively. The testing programs consisted of three test phases. Phase 1 consisted of indoor bench testing with known amounts of simulated precipitation, and phase 2 consisted of 26 weeks of actual precipitation collection at a single outdoor test site near Bay Saint Louis, Mississippi (Gordon 2003). In phase 3, rain

Corresponding author: Dr. Firat Y. Testik, firat.testik@utsa.edu

DOI: 10.1175/JTECH-D-19-0219.1

gauges were evaluated for accuracy, ease of operation, and reliability. For this phase, six NTN sites across the United States were chosen based on different climate conditions, equipment configurations, topographic settings, and power availability. Pluvio[1] performed consistently well in all of the test phases. The evaluations demonstrated that the Pluvio[1] gauges at all of the test sites were in good agreement with other rain gauges, and the paired t tests showed that the gauge measurements were not significantly different when the false positive data (a zero response from a gauge concurrent with a recorded response from other precipitation gauges) were removed from the dataset (Tumbusch 2003). Since then, several agencies have replaced their Pluvio[1] gauges and upgraded to Pluvio[2] gauges (NADP 2015) since Pluvio[2] precipitation gauges are low cost and lower gauge network maintenance costs. The expanding popularity of Pluvio[2] increases the necessity to assess its measurements errors.

According to the manufacturer's guidelines (OTT HydroMet 2014), Pluvio[2] can measure 1 min real-time rain intensities that are equal to or larger than 6 mm h^{-1} within ±1% accuracy. However, findings of Colli et al. (2014) indicated a low degree of Pluvio[2] measurement accuracy in the RI range of 6–12 mm h^{-1}. Previously, Lanza et al. (2010) and Lanza and Stagi (2009) had also documented the underperformance of Pluvio[2]'s predecessor (Pluvio[1]) for low intensity RI measurements, providing motivation for our detailed investigation on this ongoing issue despite the instrument upgrades. Colli et al. (2014) performed a series of controlled laboratory tests to assess the Pluvio[2]'s accuracy under dynamic RI conditions at 1-min intervals. Their results indicated that Pluvio[2] tends to underestimate the RI measurements in the 6–12 mm h^{-1} range with occurrences of −100% relative errors that indicate no rain was detected by Pluvio[2] when low precipitation rates were still generated by the systems. When filtering out all of the zero RI measurements from their dataset, the average relative errors of 6–12 mm h^{-1} class were improved at the expense of excluding nearly 40% of the data of this range. Moreover, when they analyzed the entire dataset for a longer time interval (>1 min resolution), the accuracy and precision of RI measurements deteriorated with the increase of the time interval compared to the filtered dataset. This was caused by the contributions of the large negative relative errors for the zero RI measurements (−100%) of filtered-out data, which dragged the average relative errors down to negative values and increased the standard deviation of the entire sample.

In a previous study on error sources of precipitation measurements using electronic weighing systems, Sevruk and Chvíla (2005) found that electronic weighing gauges, including the Pluvio[1], measured less liquid precipitation than the standard nonrecording Hellmann gauge due to inaccurate measurements of small precipitation amounts, and the relative error of measurements increased particularly when the measuring intervals exceed 1 min. Colli et al. (2013) performed a dynamic two-step RI response test to determine the contribution of the RI magnitude from the first step on the Pluvio[2] RI measurements to the second step. They identified that the measurement error of the second step was a function of the RI ratios from the first and second steps, which increased linearly with the increase of this ratio. This error, which is caused by a delay in Pluvio[1]'s response time (Lanza et al. 2005) and/or due to partial reading from the previous minute, induces a large measurement uncertainty that is often beyond the World Meteorological Organization (WMO) recommended limit of ±5% (Nitu et al. 2019). However, Pluvio[2], which is the successor of Pluvio[1], implements faster data acquisition techniques while using the same data processing algorithm. This resulted in faster data processing and nearly nondelayed response time of Pluvio[2] (Nemeth 2008).

The main purpose of this study is to assess Pluvio[2] [operational system (OS) version 1.03.0] gauge's random (repeatability and uniformity) and systematic (bias and linearity) error components, as well as stabilization of the measurements on time-varying rainfall intensities. The paper is organized as follows. Section 2 describes the methodology of the study. Measurement system analysis of the Pluvio[2] with repeated measurements is provided in section 3. Stabilization of the measurements on time-varying rainfall intensities is discussed in section 4. RI measurement errors induced by real-time bucket weight response time are quantified according to the Pluvio[2] algorithm and are demonstrated in a dynamic scenario in section 5. Conclusions are provided in section 6.

2. Methodology

Considered RI values were in the range of 6–70 mm h^{-1}. A more detailed investigation was performed close to 6 mm h^{-1} (=0.1 mm min^{-1}), which is the RI measurement threshold value of the Pluvio[2]. There are two available Pluvio[2] models: Pluvio[2] 200 (bucket type: 200 cm^2) and Pluvio[2] 400 (bucket type: 400 cm^2). In this study, Pluvio[2] 400 was utilized. While the selection of Pluvio[2] 400 over Pluvio[2] 200 was solely based upon the availability of Pluvio[2] 400 in our laboratory, the results presented in this study are equally applicable to both of the Pluvio[2] models as explained as follows. According to OTT HydroMet technical reports (OTT HydroMet 2014), the new-generation OTT Pluvio[2] 200 and OTT Pluvio[2] 400 both use the same weight-balance technology with the same accuracy and electronics. This information was verified through communication with the manufacturer. The only differences between OTT Pluvio[2] 200 and 400 are the size of the orifice opening and the recordable precipitation amount that each holds. Pluvio[2] 400 (used in the present study) has a 400 cm^2 orifice area, and Pluvio[2] 200 has a 200 cm^2 orifice area; they have precipitation collection capacities of 750 and 1500 mm, respectively. The field performances of the two Pluvio[2] versions may differ due to parameters such as evaporative losses and sampling volume. Since our experiments were conducted in a controlled laboratory environment, potential sources for the field performance differences between the two Pluvio[2] versions are not applicable to our study. Therefore, the results presented in this study based on the experiments with Pluvio[2] 400 are also applicable to Pluvio[2] 200. Moreover, to remove the possible dependency on the OS, internal filtering, and other installation characteristics of Pluvio[2], a guided accuracy test was carried out as per Pluvio[2] guidelines. The results of the test showed that the deviation in weight measurements due to the

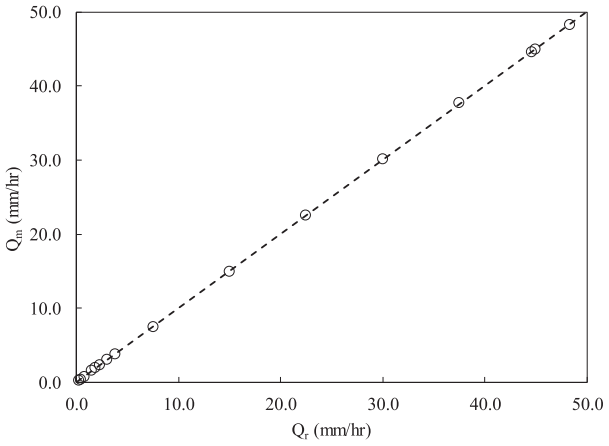


FIG. 1. Comparison of pump-generated (Q_r) and measured (Q_m) flow rates using the precise balance and densitometer to evaluate the pump precision and accuracy. Open circles denote measurements and the dashed line denotes a 1:1 relationship.

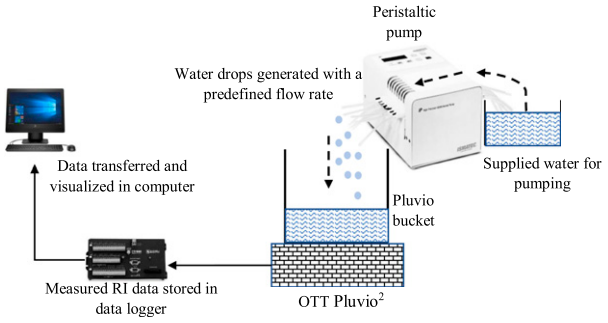


FIG. 2. Schematic diagram of the laboratory experimental setup.

installation was within the manufacturer's specified limit (±4.0 g). Note that Pluvio^2 was updated to the latest firmware (OS version 1.03.0) before starting this study. Moreover, we also investigated Pluvio^2's algorithm for the reported real-time (RT) rain intensity, RI_m, calculations and discussed the findings in section 5. In the Pluvio^2 weighing precipitation gauge, precipitation is collected and instantaneously weighed. The volumetric amount of precipitation is then directly calculated from the measured weight of the collected precipitation by considering the water density value and ambient temperature. For the open bucket designs, the decrease in mass due to evaporation has to be considered. Colli (2014) approximated the magnitude of precipitation depth measurement fluctuations due to evaporation as 0.3 mm for temperature changes of approximately 20°C. In general, inaccuracy in measurements can originate from instrumental and environmental variables. Typically, environmental variables associated with measurement inaccuracies include gradients of atmospheric temperature, wind speed, and humidity, and instrumental variables associated with measurement inaccuracies include instrument sensitivity, sampling characteristics, measuring range, and mechanical errors (Colli 2014).

Experimental tests in this study were performed under laboratory conditions without notable changes in the environmental variables. A fully controlled automatic peristaltic pump (Cole-Parmer Ismatec IPC ISM931C) of variable flow rates was used to perform laboratory simulations of different RI patterns. Before the start of each experiment, the pump flow rate was calibrated to provide flow rates within an accuracy of ±0.05 mm h^{-1}. Moreover, a precision balance (±0.01 g) and a high accuracy (density: ± 0.001 g cm^{-3}; temperature: ±0.1°C) densitometer (brand and model: Anton Paar DM-35) were used to validate the water volume flow rates provided by the pump. Figure 1 demonstrates a comparison of the configured pump flow rates (Q_r) and measured flow rates (Q_m). Each configured pump flow rate is tested 15 times, and the measured flow rate values presented in this figure are the averages of these tests. The average of the relative errors, $\overline{E}[E$ = (Predicted value − Reference value)/Reference value], and the associated standard deviation of the errors were 0.12% and 0.83%, respectively. In these calculations, Q_r was considered as the reference value. Water drops generated by using the pump for the selected flow rates (i.e., the reference RI, which is denoted by RI_r) dropped vertically in still air (i.e., no wind effects) and were collected by the Pluvio^2 weighing precipitation gauge. The Pluvio^2 RI readings (i.e., RT precipitation intensity from the Pluvio^2, which is denoted by RI_m) were recorded by a datalogger at different time intervals (6–60 s) throughout the experiments. A schematic diagram of the experimental setup is presented in Fig. 2.

The pump characteristics and the experimental setup were suitable to investigate the measurement system under both static and dynamic experimental conditions. Four types of experimental simulations were conducted, and these are named as Constant, Abrupt Change, Gradual Step Change, and Cyclic Step Change, based on the settings of the pump over time. The experimental conditions for the simulated patterns are tabulated in Table 1 and described as follows.

a. Constant RI_r

In the Constant RI_r experiments, the pump provided a specific flow rate over time to simulate a constant RI_r value. There were three sets of experiments corresponding to the Constant RI_r study. The main goal of the first set of experiments was to determine the random and systematic error components of RI_m for a wide range of RI_r (section 3). This set included experiments with constant RI_r values of 6, 6.25, 6.5, 6.75, 7, 8, 10, 20, 30, 50, and 70 mm h^{-1} for a period of 12 min each (runs 1–11, Table 1). The second set of experiments was performed to determine the effectiveness of the measurement system in producing a measurement when RI is close to the manufacturer-provided lower threshold value of 6 mm h^{-1} (section 3). This set included experiments with constant RI_r values of 6, 6.25, 6.5, 6.75 and 7 mm h^{-1}, but for a period of 52 min each (runs 12–16, Table 1). Note that the initial and last minutes of Pluvio^2 RI readings in the first and second sets were omitted from the analyses due to the possibility of partial measurements for those minutes. Also, both sets of experiments had a time resolution of 1 min. The third set of experiments was conducted to investigate the time delay on bucket

TABLE 1. Summary of the experimental conditions.

Run No.	Pump setting	RI_r (mm h^{-1})	Duration (min)	Assessments
1–11	Constant[a]	6, 6.25, 6.5, 6.75, 7, 8, 10, 20, 30, 50, 70	12	Bias, linearity, repeatability, uniformity, tolerance intervals
12–16	Constant[a]	6, 6.25, 6.5, 6.75, 7	52	Bias, linearity, repeatability, uniformity, tolerance intervals
17	Constant[a]	10	5	Time delay
18–21	Abrupt[b]	(70→7), (70→8), (70→10), (70→20)	12	Stabilization
22	Gradual[c]	50→30→20→10→7	60	Stabilization
23	Cyclic	7→10 →20→30 →50→30→20→10→7[d]	120	Stabilization, hysteresis
24	Cyclic	30→15→10→15→30[e]	5	Time delay, algorithm validation
25	Cyclic	7→10→20→30→50→30→20→10→7[f]	9	Error demonstration

[a] First (runs 1–11), second (runs 12–16), and third (run 17) test sets included experiments with 12, 52, and 5 min, respectively, for each RI_r.
[b] RI_r suddenly dropped from 70 mm h^{-1} to the selected RI_r value for 12 min of measurement for each of the four test cases.
[c] RI_r was gradually stepped down from 50 to 7 mm h^{-1} for 12 min of measurement at each step.
[d] RI_r was gradually increased from 7 to 50 mm h^{-1} and decreased back to 7 mm h^{-1} for 12 min of measurement at each step except 50 mm h^{-1} where 24 min measurement recorded.
[e] RI_r was gradually decreased from 30 to 10 mm h^{-1} and increased back to 30 mm h^{-1} for 1 min of measurement at each step.
[f] As in run 23, except 1 min of measurement was taken at each step.

weight measurements (section 5). In this experiment, RI was set to 10 mm h^{-1} for a duration of 5 min (run 17, Table 1) and data were recorded at a high time resolution (6 s). This experiment was different from other constant tests due to the different time resolutions, and the initial and last minutes of this experiment were not omitted.

b. *Abrupt change in RI_r*

In the Abrupt Change experiments, RI_r was set to 70 mm h^{-1} initially and then suddenly dropped to 7, 8, 10, or 20 mm h^{-1} in different experimental runs (i.e., four separate experimental runs, runs 18–21, Table 1). Each experimental run continued for at least 12 min after the abrupt change of the RI_r value to get 10 RI_m (excluding the RI_m in the first and last minutes that are considered as intermittent values). Both the Abrupt Change and the Gradual Step Changes (explained next) simulations assessed the stabilization of the RI_m on the RI_r under dynamic conditions (section 4).

c. *Gradual step changes in RI_r*

In the Gradual Step Changes experiment, RI_r was gradually stepped down from 50 mm h^{-1} to 30, 20, 10, and finally to 7 mm h^{-1} (run 22, Table 1). In this simulation, the pump ran continuously throughout the gradual step changes from higher to lower RI values. At each step change level, the experiment covered 12 min to obtain 10 RI_m after excluding the first and last measurements.

d. *Cyclic step changes in RI_r*

In the Cyclic Step Changes experiment, the pump flow rate was set to generate RI_r values in sequential steps, starting with 7 mm h^{-1}, which was successively increased to 10, 20, 30, and 50 mm h^{-1}, and then successively decreased to 30, 20, 10, and finally back to 7 mm h^{-1} (run 23, Table 1). Each RI was visited two times separately from an increasing and decreasing sequential steps, except the peak RI value of 50 mm h^{-1}. At each step change level, 12 min of data were collected during both increasing and decreasing step changes, except at 50 mm h^{-1} level, where 24 min of data were collected. This pattern resulted in 24 min of measured data for each RI level. However, in the analyses of the experimental data, the intermittent minute between the changes in RI_r and also the first minute of the cyclic experiment were removed to avoid partial readings. The exclusion of those minutes resulted in 11 min of available data for analysis for each RI_r step, except for 50 mm h^{-1} where 23 min of data were available. In addition to observing the stabilization of the RI_m values around the set RI_r value, the Cyclic Step Changes simulation investigated hysteresis effects (section 4). The hysteresis effect can be evaluated as the maximum difference between the Pluvio^2 reported RI_m output and the reference RI_r when the specified RI_r is approached first from an increasing and then a decreasing way in a cyclic pattern. A comparison of the relative errors for RI_m with respect to the specified RI_r for the increasing and decreasing phases in a cyclic test would indicate the impact of hysteresis in RI measurements.

An additional Cyclic Step Changes experiment was conducted where RI_r was changed more frequently to explore the dynamic response of Pluvio^2 measurement (section 5). In this experiment, different pump flow rates were set to generate different RI_r values at each minute of the experiment, and data were collected at a 6 s time resolution. The RI_r was initially set to 30 mm h^{-1} and then consecutively stepped down to 15 and 10 mm h^{-1}, and then consecutively stepped up to 15 and 30 mm h^{-1} (run 24, Table 1). In addition to investigating the response delay in Pluvio^2 measurement, this type of cyclic test assisted in evaluating the RI measuring algorithm using the real-time bucket weight measurement. Furthermore, another Cyclic Step Changes experiment was performed, in a similar setting as run 23, except 1 min of measurement was taken at each step (run 25, Table 1) to demonstrate the potentiality of large error under frequent RI changes (section 5, Case I).

TABLE 2. Summary of the repeatability test statistics. $\overline{RI_m}$ indicates the average of the RI_m in the sample of corresponding RI_r. The repeatability standard deviation S estimates are based on Eq. (1).

RI_r (mm h^{-1})	S	95% measurement range ($\overline{RI_m}$ ± indicated value)	Half width of the intervals as RI_r (%)
7	0.19	0.64	9.2
8	0.2	0.68	8.4
10	0.24	0.81	8.1
20	0.37	1.25	6.3
30	0.6	2.03	6.8
50	0.74	2.50	5.0
70	1	3.38	4.8

3. Measurement system analysis of the Pluvio^2 gauge

Error in the Pluvio^2 measurements with respect to time and over the RI range of 6–70 mm h^{-1} is assessed through statistical analysis of the measurement variability. Minitab statistical software was used in the study. In particular, bias, linearity, repeatability, and uniformity of the measurement system were studied under the Constant RI_r setup described in the previous section. Measurement variability is often described by the normal distribution, and this is an assumption of the standard methods of measurement system analysis. Hence, goodness of fit of the normal distribution to the measurements was checked in the analysis. Unless otherwise stated, samples of 10 RI_m taken at 1 min resolutions by the Pluvio^2, after omitting the initial and last measurements, were considered for each of the experimental settings.

A number of experiments were performed to test the Pluvio^2 under Constant RI_r with values close to the lower RI threshold value of 6 mm h^{-1}. The initial set of experiments (runs 1–11, Table 1) revealed that many RI_m were zero when RI_r was set to less than 6.5 mm h^{-1}, indicating that the gauge may not be effective in measuring the RI_r close to the lower threshold. To confirm the findings of these experiments and to estimate the proportion of RI_m that are nonzero (in fact, nonzero measurements were close to and around the corresponding RI_r values), a new set of experiments was conducted to obtain 50 RI_m for each of the selected RI_r that are less than or equal to 7 mm h^{-1} (runs 12–16, Table 1). The estimated proportions of nonzero measurements with a sample size of 50 were 0.06, 0.1, 0.74, 0.96, and 1 for the RI_r settings of 6, 6.25, 6.5, 6.75, and 7.00 mm h^{-1}, respectively. This finding indicates that there is a high probability for missed RI_m when RI is less than 6.5 mm h^{-1} and the effectiveness of the gauge increases as RI_r is raised to 7 mm h^{-1}. Consequently, it is suggested that Pluvio^2's lower RI_r threshold value should be redefined as 7 mm h^{-1}. In the following analyses, RI_r values of 7 mm h^{-1} and larger are considered.

The reproducibility of the experimental conditions was checked by collecting two sets of RI_m samples during two different experimental runs with the same RI_r value of 7 mm h^{-1}. Two-sample t test and F test for two variances were utilized for the evaluation. Normal probability plots of the two RI_m samples indicated that normal distribution is a good fit to the data. The two-sample t test was used to test the hypothesis of equal means of the RI_m under the two experimental conditions against the alternative that these means are not equal. The p value of the test was 0.747 and the 95% confidence interval for the differences between the means is (−0.1307, 0.1787). This result indicates that there is not enough evidence to claim that the RI_m means are different for reasonable choices of significance level α. The F test for the two variances was used to test the null hypothesis that the ratio of the variances of the RI_m under the two experimental conditions is equal to 1 (homogeneity of the variances) against the alternative that the ratio is different from 1. The p value of the F test was 0.556 and the 95% confidence interval for the standard deviation ratio was (0.610, 2.457). Hence, the test fails to reject that the variances were equal for reasonable choices of significance level α. Overall, there is no evidence to claim that the samples came from different populations and we conclude that it is reasonable to assume that the experimental conditions in the study were reproducible.

A repeatability study to assess the random error component of the measurement system (i.e., the variation in the measurements obtained when the same RI_r is measured repeatedly under the same experimental conditions) was conducted. To quantify the repeatability component of the measurement error, the standard deviation estimate,

$$S = \sqrt{\frac{\sum_{i=1}^{n} (RI_{m,i} - \overline{RI}_m)^2}{n-1}}, \tag{1}$$

is used, where n is the sample size, $RI_{m,i}$ is the ith RI measurement, and $\overline{RI}_m$ is the average of the RI_m in the sample. A summary of repeatability test statistics is presented in Table 2 including the standard deviation for corresponding RI_r. Under the normal distribution model of the measurements and for the sample size of 10, at least 95% of the repeated measurements will lie within the tolerance interval $\overline{RI}_m \pm 3.379S$ with a probability of 0.95. Table 2 presents the half width of the 95% repeated measurements intervals as a percentage of the corresponding RI_r, and it can be seen from this table that the uncertainties in the measurements decrease as a percentage of RI_r with an increase of the RI_r value.

To assess the uniformity of the measurement system (i.e., the change in repeatability over the measurement range considered) Bartlett's test (assuming normal distribution) and Levene's test (for any continuous distribution) were used to test the equality of the variances at different RI_r values. The p values (0.000 for the Bartlett's test and 0.000 for the Levene's test) indicated that the data provide enough evidence to claim that the variances of the repeated measurements are unequal at different RI_r values. In particular, variation in the measurements by the Pluvio^2 increases as the RI_r is increased. The effect of the random error on the measurements increases and a problematic uniformity is observed over the range of RI. The 95% Bonferroni confidence intervals for the standard deviations are provided in Fig. 3. Also seen in this figure, the standard deviation estimates increase, and their confidence intervals widen with an increase of RI_r.

The systematic error component of the measurement system is often evaluated by a study of the bias ($\overline{RI}_m - RI_r$) and linearity. It is clear from this equation that we did not perform normalization on bias estimation with respect to RI_r. Therefore,

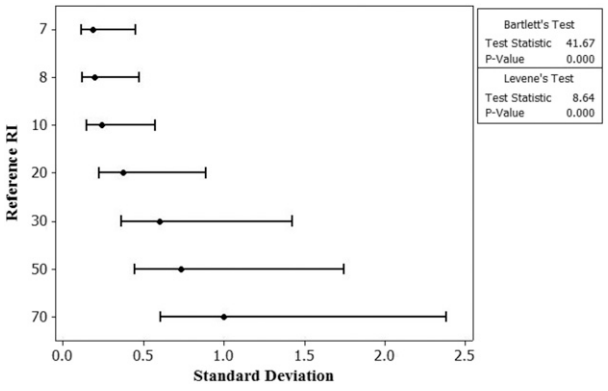


FIG. 3. Results of tests for equal variances for repeated RI_m at the selected RI_r values and the 95% Bonferroni confidence intervals for the standard deviations.

bias may seem to increase with increases of RI_r. Bias is the combined effect of all sources of variability, which consistently offsets the results of repeated applications of the measurement system. Linearity is the change in the bias over the measurement range. It is assessed by using the slope of a linear regression, where the relationship between the average bias (response variable) and the RI_r values (explanatory variable) is described by a straight line. A slope of 0 for the regression line indicates that linearity is not present, that is the measurement system has the same bias for all reference values. Consider the gauge bias results provided in Fig. 4; *t* tests of the null hypothesis that the bias is zero against the alternative that it is different from zero indicate that the biases are statistically significant for the RI_r values of 7 and 8 mm h^{-1} at a level of $\alpha = 0.05$. The bias estimates are -0.26 and -0.21, showing that the RI_r values are underestimated and on average the RI_m are less than the corresponding RI_r values. For the rest of the RI_r values considered, a statistically significant bias is not observed. Now consider the regression line, its 95% confidence interval, and the gauge linearity result in Fig. 4. The *p* value is 0.264 for the *t* test of the null hypothesis that the regression line slope is zero against the alternative that it is different from zero. This suggests that the measurement system does not have a statistically significant linearity problem across different RI_r values.

4. Stabilization of the OTT Pluvio^2 RI measurements on the RI reference value under dynamic conditions

Experimental runs were also conducted with time-varying RI_r values as simulations of RI patterns. The RI patterns simulated are Abrupt Change, Gradual Step Changes, and Cyclic Step Changes as described in section 2. The RI_m values from the experiments were used to observe the stabilization of the measurement system on the time-varying RI_r values. In this regard, RI_m are plotted over time, together with statistical control limits (UCL—upper control limit; and LCL—lower control limit; see, e.g., Montgomery 2009)

$$\text{UCL} = \text{RI}_r + 3\frac{\overline{R}}{1.128} \quad \text{and} \quad \text{LCL} = \text{RI}_r - 3\frac{\overline{R}}{1.128} \tag{2}$$

for each RI_r. Here, $\overline{R}/1.128$ is a standard deviation (σ) estimate of the measurement process, and

$$\overline{R} = \frac{\sum_{i=2}^{n} |\text{RI}_{m,i} - \text{RI}_{m,i-1}|}{n-1} \tag{3}$$

for a sample of *n* RI_m corresponding to a value of RI_r. Under the assumption of normal distribution and when there is no measurement bias, 99.73% of the measurements are confined by RI $\pm 3\sigma$. Hence, an RI_m exceeding the control limits can be taken as an indication that the mean of the measurement process did not stabilize on the corresponding RI_r.

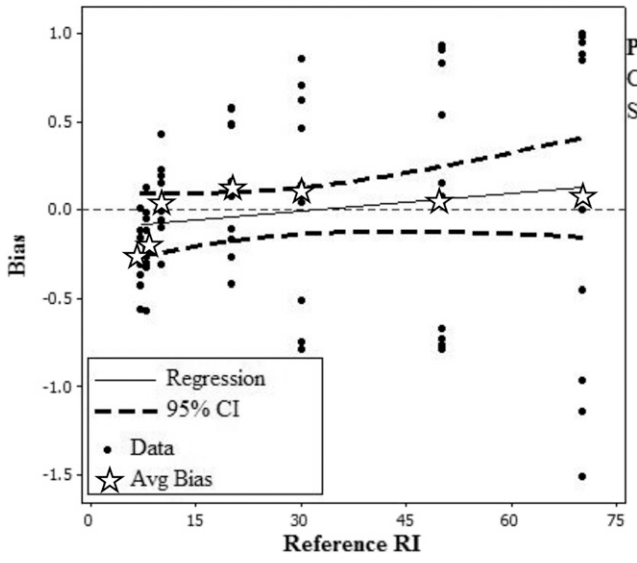
RI_m and their corresponding statistical control limits under the experimental simulations of the time-varying RI_r are provided in the graphs in Figs. 5a–c, respectively, for the Abrupt Change, Gradual Step Changes, and Cyclic Step Changes experiments. As can be seen from the graphs, none of the RI_m exceed the control limits. Therefore, it is concluded that there is no evidence to claim that the means of the measurements are not stabilized on the corresponding time-varying RI_r values. Moreover, the relative errors for RI_m with respect to the specified RI_r for the increasing and decreasing phases was within -0.07% to $+0.51\%$. This low error percentage indicates that there was no significant effect of hysteresis in the RI measurements.

Homogeneity of variances of RI_m under the considered experimental scenarios was also evaluated for some of the RI_r values. From the results of tests for equal variances and the 95% confidence intervals given in Figs. 6a and 6b, respectively, for the RI_r values of 7 and 20 mm h^{-1}, it is concluded that there is not enough evidence at a significance level of $\alpha = 0.05$ to claim that the measurement variances are different under dynamic conditions for a given RI_r.

5. Pluvio^2 rain intensity measurement errors induced by bucket weight response time

In this section, we assessed the RT rainfall intensity output of Pluvio^2 in relation to Pluvio^2's RT bucket weight measurements. Pluvio^2's algorithm calculates RI values based on RT bucket weight measurements. Therefore, we first assessed Pluvio^2's RT bucket weight measurements using the highest possible temporal resolution output produced by Pluvio^2, which led to the identification of a time delay in RT bucket weight measurements. We then faithfully reproduced Pluvio^2's RT rainfall intensity calculation algorithm using RT bucket weight measurements with relevant validation and provided the pertinent mathematical relations. We finally used this faithfully reproduced algorithm to evaluate Pluvio^2's RT rainfall intensity measurement capability under two different dynamic rainfall scenarios.

To identify the time delay (t_d) in RT bucket weight time series measurements, two sets of experiments were conducted with constant and cyclic step changes of the reference rainfall intensity in Pluvio^2 measurements with the highest temporal resolution, $\Delta t = 6$ s (experimental runs 17 and 24 in Table 1). In Fig. 7, measured bucket weight increment time series ΔB for experimental runs 17 (constant) and 24 (cyclic) are shown.



Gage Linearity

Predictor	Coef	SE Coef	P
Constant	-0.1063	0.1054	0.317
Slope	0.003329	0.002954	0.264

S 0.551155 R-Sq 1.8%

Gage Bias

Reference	Bias	P
Average	-0.013571	0.783
7	-0.261000	0.002
8	-0.209000	0.018
10	0.040000	0.602
20	0.139000	0.201
30	0.087000	0.611
50	0.049000	0.782
70	0.060000	0.816

FIG. 4. Bias and linearity study results for the Pluvio^2 gauge. Solid circles represent RI_m, stars represent average bias, thick dashed lines represent confidence intervals (CI), and the solid line represents the linear regression line. Coef is the regression coefficient, SE Coef is the standard error of the regression coefficient, and P is the p value.

This figure indicates that Pluvio^2 starts reporting bucket weight measurements between 13 and 17 s from the start of the rainfall. Therefore, the measurement time delay t_d value is in between 13 and 17 s, which cannot be identified with a higher accuracy due to the limitation in Pluvio^2's highest temporal measurement resolution of 6 s. In our analyses below, we used the t_d value of 17 s as it will introduce the maximum possibility of error in RI calculation and will produce a conservative estimation during the analysis.

To demonstrate the errors induced by the identified time delay t_d in bucket weight measurements, a relationship between the Pluvio^2 bucket weight time series and the known/reference rainfall intensity time series $RI_{r,t}$ is necessary. Based upon our experiments, we derived the following equation to predict the reported bucket weight time series, $B_{p,t}$, by Pluvio^2 for a given rainfall event with the known/reference rainfall intensity time series $RI_{r,t}$:

$$B_{p,t} = B_{p,t-\Delta t} + (RI_{r,t-t_d}\Delta t). \tag{4}$$

Here, the p, r, and m indices are used to indicate predicted, reference, and measured parameters, respectively. $B_{p,t-\Delta t}$ is the predicted bucket weight at time instant $t - \Delta t$, $RI_{r,t-t_d}$ is the reference rain intensity at time instant $t - t_d$, and Δt is the time resolution of the Pluvio^2 bucket weight measurements which can be set as multiples of 6 s. The predictions of the reported bucket weight time series can then be used to predict the reported rain intensity time series by Pluvio^2 as follows.

We identified that the internal (hidden) algorithm of Pluvio^2 calculates the reported $RI_{m,t}$ values by implementing a moving precipitation accumulation over the last minute in the bucket weight measurements. Based upon this finding, we developed Eq. (5) to predict the reported rain intensity values by Pluvio^2 using the predicted bucket weight values calculated from Eq. (4). Note that the relationship between the reported real-time rainfall intensity $RI_{m,t}$ and the measured bucket weight $B_{m,t}$ at a given time t has the same mathematical form of Eq. (5). Depending on the purpose, the predicted and measured quantities can be used interchangeably in Eq. (5):

$$RI_{p,t} = \begin{cases} 60 \times \sum_{i=t-t_a+\Delta t}^{t} (B_{p,i} - B_{p,i-\Delta t}), & RI_{p,t} \geq 6\,\text{mm h}^{-1} \\ 0, & RI_{p,t} < 6\,\text{mm h}^{-1} \end{cases}. \tag{5}$$

Here, $RI_{p,t}$ is the predicted rain intensity for the reported value by Pluvio^2 at a given time, t_a is the rain accumulation time step of 60 s that is used for rain intensity calculations by Pluvio^2, and the multiplication factor of 60 min h^{-1} is included in Eq. (5) to provide $RI_{p,t}$ values per hour following the convention. Pluvio^2 has a reported minimum threshold for rain intensity measurements of 6 mm h^{-1}. This means that whenever $RI_{p,t} < 6$ mm h^{-1}, the $RI_{p,t}$ value is reported as 0 by Pluvio^2. Therefore, a conditional statement is implemented in $RI_{p,t}$ calculations in Eq. (5).

Figure 8 provides a comparison between the real-time rain intensity measurements, $RI_{m,t}$, reported by Pluvio^2 and predicted rain intensity time series, $RI_{p,t}$, using Eq. (5) alone for the experimental run 24. Here, Pluvio^2-measured bucket weight values, rather than the values predicted by Eq. (4), were used in calculations of Eq. (5) to evaluate the predictive capabilities of Eq. (5) alone. Note that predictive capability of Eq. (4) alone and predictive capabilities of Eqs. (4) and (5) combined are assessed and demonstrated later in Figs. 9 and 10, respectively. The comparison presented in Fig. 8 shows that the average relative error $\overline{E}$ $[E = (RI_{p,t} - RI_{m,t})/RI_{m,t}]$ is −0.085%. Ideally, $RI_{m,t}$ and $RI_{p,t}$ values should overlap with $\overline{E}$ value of 0.00%. The observed deviation is caused by the

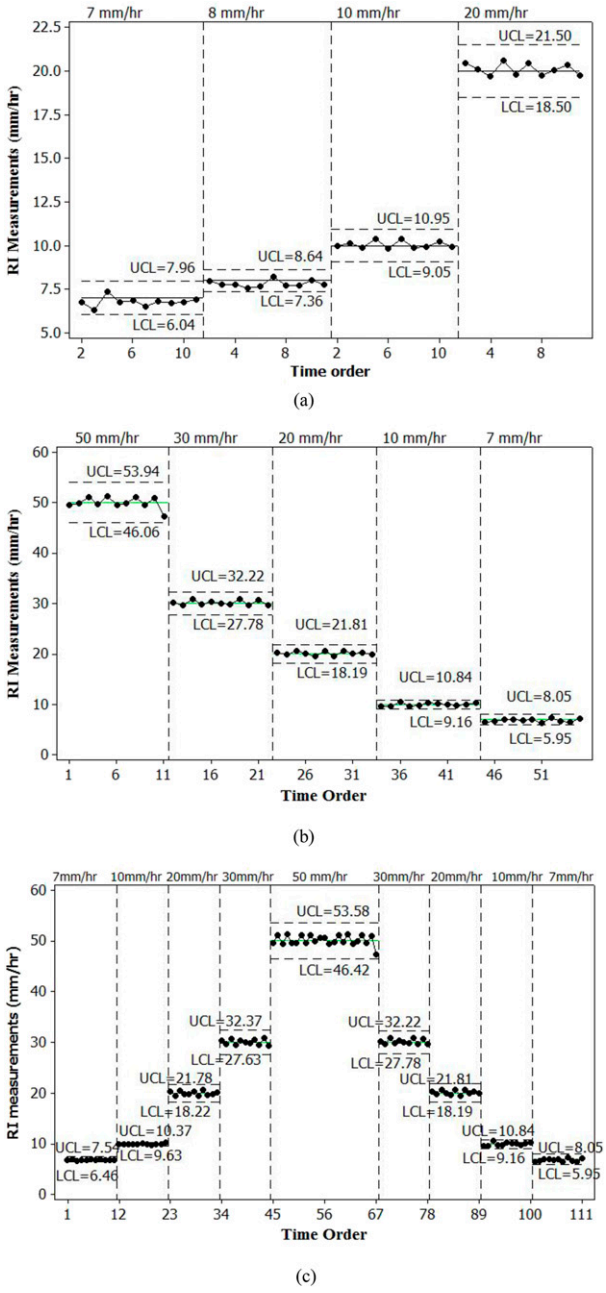


FIG. 5. Statistical control limits (UCL and LCL) and stabilization of the RI_m on the RI_r after (a) Abrupt Change, (b) Gradual Step Change, and (c) Cyclic Step Changes. The x axis provides the time order, which can be considered as the experimental time in minutes.

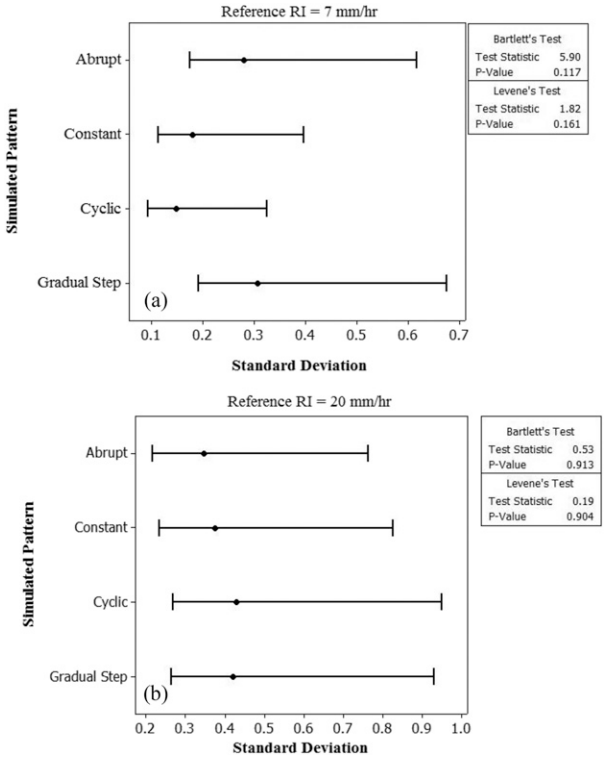


FIG. 6. Results of tests for equal variances of RI_m and 95% Bonferroni confidence intervals for standard deviations under the experimental simulations considered for the RI_r settings (a) 7 and (b) 20 mm h^{-1}.

resolution difference between the providing bucket weight measurement output (with resolution of 0.01 mm) and the collecting bucket weight by the precipitation gauge known as internal raw data (with resolution of 0.001 mm). Nevertheless, the insignificant relative error value clearly indicates that our suggested algorithm faithfully reproduced the one used by Pluvio^2 with sufficient accuracy for our analyses presented in this section. A similar approach was followed to evaluate the predictive capability of Eq. (4) separately by comparing the predicted bucket weight time series ($B_{p,t}$) which was calculated using Eq. (4) with the measured real-time bucket weight ($B_{m,t}$) output from Pluvio^2 for the reference rainfall intensity time series ($RI_{r,t}$) of experimental run 24 (Fig. 9). This comparison shows that the average relative error, $\overline{E}$, value between $B_{p,t}$ and $B_{m,t}$ values for the entire time series is −0.0006% with a standard deviation of 0.004%, validating Eq. (4) for highly accurate bucket weight predictions. Once the predictive capabilities of Eqs. (4) and (5) were validated individually, we assessed the predictive capabilities of Eqs. (4) and (5) combined. Figure 10 shows a comparison between the measured $RI_{m,t}$ and predicted $RI_{p,t}$ time series of experimental run 24 and demonstrates the predictive capability of Eqs. (4) and (5) combined. Here, the predicted $RI_{p,t}$ time series was obtained by, first, calculating the predicted bucket weight time series ($B_{p,t}$) for the reference rainfall intensity time series ($RI_{r,t}$) of experimental run 24 using Eq. (4), and then computing the predicted rainfall intensity time series using Eq. (5). The average relative error $\overline{E}$ between $RI_{m,t}$ and $RI_{p,t}$ values for the entire time series is −0.286%, demonstrating the good predictive capability of Eqs. (4) and (5) combined.

To demonstrate the potential errors in Pluvio^2 RI measurements, two different dynamic rainfall scenarios were simulated, and results were evaluated based on comparisons

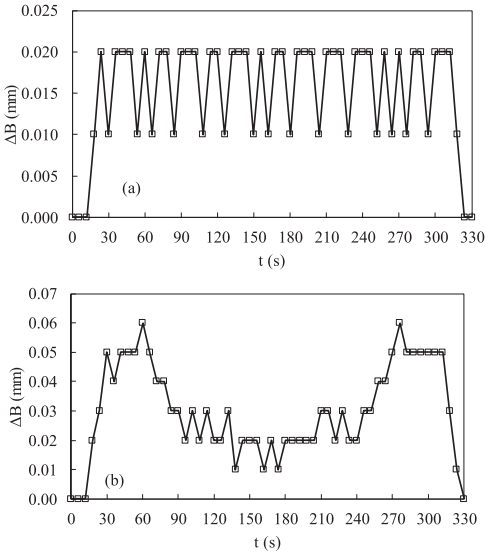


FIG. 7. Measured bucket weight increment (ΔB) as a function of experimental time t during experiments with RI_r settings as (a) constant at 10 mm h^{-1} with a 5 min duration (refer to the experimental run 17 in Table 1) and (b) cyclic step changes as 30 → 15 → 10 → 15 → 30 mm h^{-1} with a 1 min duration for each RI_r step (refer to the experimental run 24 in Table 1). ΔB is measured every 6 s. Both plots exhibit that the first measured bucket weight increment was around 18 s after the beginning of the experiments.

between the reference ($RI_{r,t}$) and predicted ($RI_{p,t}$) rainfall intensities for both simulated scenarios (see Figs. 11 and 12, discussed later). These two scenarios are Case I—an additional laboratory experiment with Cyclic Step Changes in RI_r with an RI_r duration of 1 min at each step (run 25 in Table 1) to validate the developed equations and demonstrate the error under frequent RI changes; and Case II—in situ rainfall intensity measurements from one of our field experiments to incorporate a typical real-world scenario for the time evolution of rainfall intensity in our demonstration. As presented next, results from these two cases demonstrate that Pluvio^2 RI measurements include large errors, which can be improved through improvements in Pluvio^2 algorithms for RI calculations.

a. Case I

A laboratory experiment with cyclic changes of RI was conducted with a pump setting to generate RI_r values in sequential steps, starting with 7 mm h^{-1}, which was successively increased to 10, 20, 30, and 50 mm h^{-1}, and then decreased successively to 30, 20, 10, and finally back to 7 mm h^{-1} with 1 min time duration for each of the RI steps (refer to run 25, Table 1). For the modeling, the bucket weight time series, $B_{p,t}$, was predicted using Eq. (4) and the predicted rain intensity time series, $RI_{p,t}$, was calculated from $B_{p,t}$ values using Eq. (5). Figure 11 demonstrates the combined predictive capability of Eqs. (4) and (5) by comparing predicted, $RI_{p,t}$, and measured, $RI_{m,t}$, rainfall intensity time series for experimental run 25. This comparison served two purposes: (i) verification of the reliable reproduction of the Pluvio^2 RI algorithm and (ii) estimation of the Pluvio^2 RI measurement errors under a simplified hypothetical rainfall scenario that simulates the dynamic features of rainfall (Case I). The average relative error, $\overline{E}$, for the comparison between $RI_{p,t}$ and $RI_{m,t}$ values was 0.8%, validating the predictive capability of the developed equations under frequent RI changes. This validation also gives us the opportunity to use predicted, $RI_{p,t}$, instead of measured, $RI_{m,t}$, time series for Pluvio^2 RI error estimation. Comparison of reference, $RI_{r,t}$, with both measured, $RI_{m,t}$, and predicted, $RI_{p,t}$, rainfall intensity time series individually exhibit the same significant average absolute relative error, $|\overline{E}|$, value of 41%. This significant error called for an investigation of the potential errors in Pluvio^2 RI measurements under field conditions (Case II) where RI changes can be rather frequent. Note that according to WMO (1996) guidelines, laboratory test results of catchment type rainfall intensity gauges should be presented in terms of relative error to exhibit the measurement

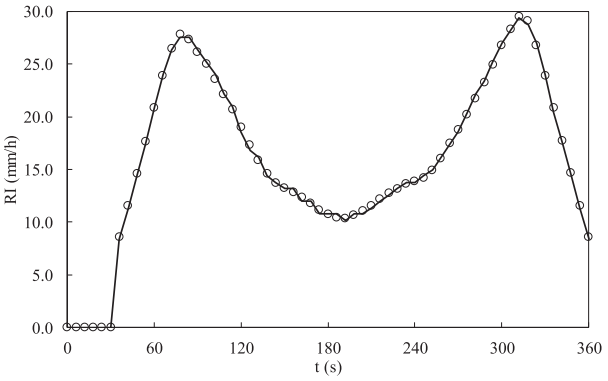


FIG. 8. Comparison of Pluvio^2-measured $RI_{m,t}$ and calculated $RI_{p,t}$ from the measured rain accumulation over the last minute in the bucket for the experimental run 24 (refer to Table 1). The result indicated an average relative error ($\overline{E}$) of −0.085% for the whole time series. Open circles denote the measured $RI_{m,t}$ and the solid line denotes the calculated $RI_{p,t}$.

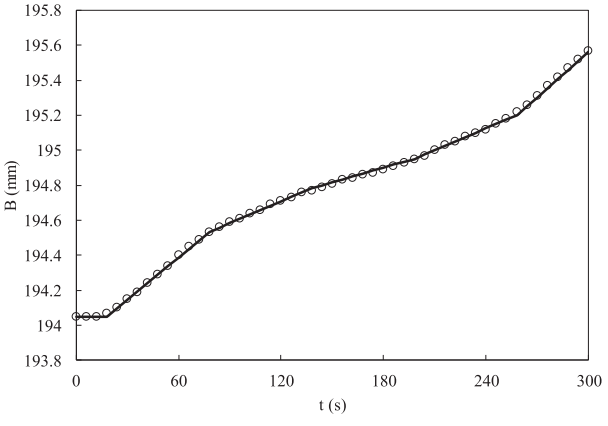


FIG. 9. Validation of Eq. (4) based on the comparison of predicted bucket weight ($B_{p,t}$) and measured bucket weight ($B_{m,t}$) for experimental run 24 (refer to Table 1). The average relative error ($\overline{E}$) for the entire time series is −0.0006% with a standard deviation of 0.004%. Open circles denote $B_{m,t}$ and the solid line denotes $B_{p,t}$.

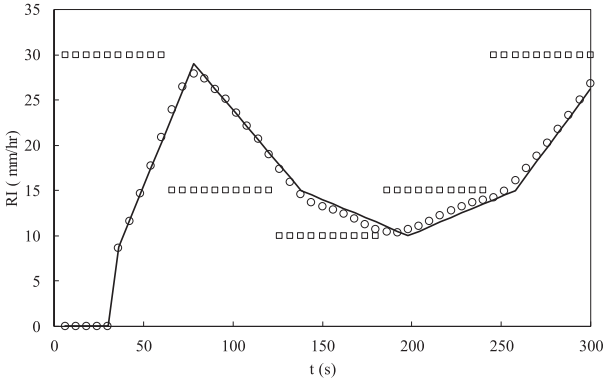


FIG. 10. Comparison of predicted $RI_{p,t}$ using Eq. (5), Pluvio^2-measured $RI_{m,t}$, and the reference $RI_{r,t}$ for experimental run 24 (refer to Table 1). Open circles denote $RI_{m,t}$, the solid line represents $RI_{p,t}$, and open squares denote $RI_{r,t}$.

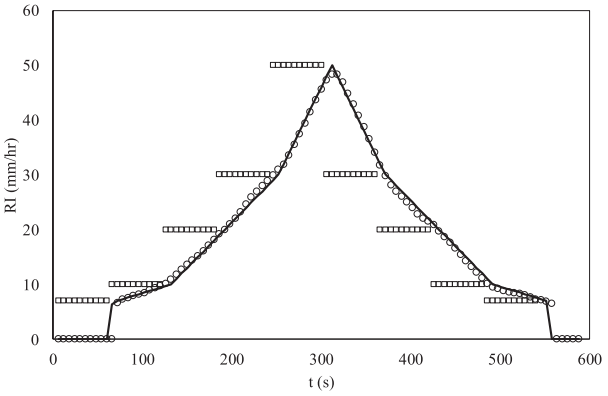


FIG. 11. Demonstration of errors in Pluvio^2 RI measurements for Case I. The time series for the predicted $RI_{p,t}$, measured $RI_{m,t}$, and reference $RI_{r,t}$ for experimental run 25 (refer to Table 1) are compared. The average relative error ($\overline{E}$) between $RI_{m,t}$ and $RI_{p,t}$ is 0.8% whereas the average absolute relative error ($|\overline{E}|$) between $RI_{p,t}$ and $RI_{r,t}$ is 41%. Open circles denote $RI_{m,t}$, the solid line represents $RI_{p,t}$, and open squares denote $RI_{r,t}$.

accuracy, and the relative error value should be within ±5% of the reference RI. In this study, we referred to the error as significant when the relative error value is larger than ±5% of the reference RI. However, here in Case I and Case II, we presented the error statistics in rain intensity measurement as absolute relative error, which is a good way to demonstrate not only the accuracy but also the precision of measurements.

b. *Case II*

For this case, an in situ rainfall intensity time series measured using an OTT Parsivel^2 Laser Weather Sensor (firmware V 2.11.6) (Löffler-Mang and Joss 2000) from one of our field experiments is utilized. Parsivel^2 provides measured RI values for each minute. In this analysis, the Parsivel^2 measured RI time series is used only to provide a typical RI change pattern during typical rainfall events. This measured RI time series was used to simulate and demonstrate Pluvio^2 RI algorithm's performance under a realistic dynamic scenario. Note that our goal was not a comparative evaluation of the measurements by Pluvio^2 and Parsivel^2, and collocated Pluvio^2 and Parsivel^2 measurements were not conducted. Therefore, accuracy of Parsivel^2 measurements is not of importance for our analysis. For the following analysis, Parsivel^2-measured RI data with a 1-min time resolution are converted to a 6-s time resolution RI time series using linear interpolation. Figure 12a shows the Parsivel^2 RI time series with both 1-min and 6-s time resolutions. The RI time series with the 6-s time resolution is used in our subsequent analysis and referred to as the reference RI, $RI_{r,t}$. Based on the linearly interpolated time series $RI_{r,t}$, the predicted Pluvio^2 RI time series $RI_{p,t}$ was calculated using Eqs. (4) and (5) and is shown in Fig. 12b. The average absolute relative error $|\overline{E}|$ between $RI_{p,t}$ and $RI_{r,t}$ was 39% for the entire time series, and when the minimum RI measurement threshold criterion as mentioned in Pluvio^2 guidelines was imposed in $RI_{p,t}$ time series calculations [recall Eq. (5)], the average absolute relative error, $|\overline{E}|$, between $RI_{p,t}$ and $RI_{r,t}$ was 19% after excluding the time steps whenever $RI_{p,t} < 6\,\text{mm h}^{-1}$. Such significant measurement errors under realistic dynamic conditions indicate high uncertainty associated with Pluvio^2 RI measurements. A closer look into Fig. 12b indicates that there is a time shift between the two time series, $RI_{p,t}$ and $RI_{r,t}$. This shift occurred due to both the time-delay in Pluvio^2's bucket weight measurement and also Pluvio^2's RI calculation algorithm. Since Pluvio^2 RI calculations at each time step incorporate precipitation accumulation over the preceding minute, modifications of the Pluvio^2 RI calculation algorithm can potentially improve the accuracy of Pluvio^2's RI measurements. To demonstrate this, the $RI_{p,t}$ time series was shifted 42 s backward in time in Fig. 12c. As can be seen from this graph, the time-shifted $RI_{p,t}$ and $RI_{r,t}$ matched closely with an average absolute relative error $|\overline{E}|$ of 2% after imposing the minimum RI threshold criterion. These results reveal valuable opportunities to significantly improve the field performance of Pluvio^2's RI measurements.

6. Conclusions

In this study, we statistically analyzed the Pluvio^2 gauge's measurement error components and the stabilization of the gauge's measurements on the corresponding reference value under constant, abrupt, gradual step, and cyclic step change simulations of the RI. The range of RI values considered in the experiments was 6 to 70 mm h^{-1}, with a higher granularity in the range of 6 to 7 mm h^{-1} and with 6 mm h^{-1} being the manufacturer provided lower threshold value for the measurements. Moreover, we examined the RT bucket weight measurements and the algorithm of Pluvio^2 RI calculations from the bucket weight measurements through experiments with high time resolution.

All experiments were conducted by using a highly precise and automated peristaltic pump that provided predefined flow rates in simulating constant RI and its changes. Repeated measurements of constant RI indicated that there is a high probability of false zero readings when the RI_r is in the range $6 \le \text{RI} < 6.5\,\text{mm h}^{-1}$. Furthermore, statistically significant negative biases were observed for the RI_r settings of 7 and

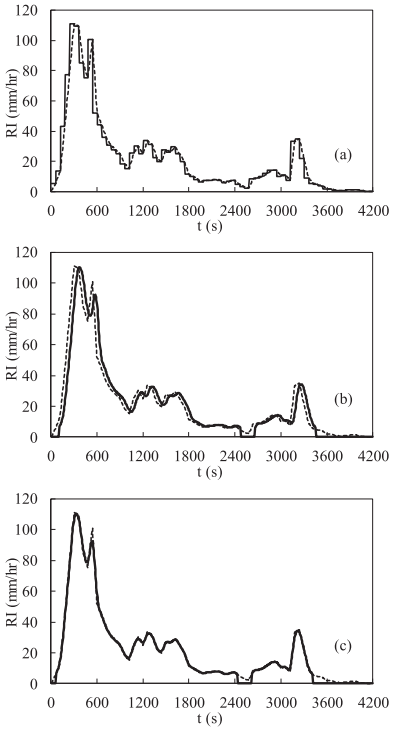


FIG. 12. Demonstration of Pluvio^2 RI measurement analysis for Case II. (a) In situ Parsivel^2 RI measurements with a 1 min time resolution (solid line) and the reference $RI_{r,t}$ with a 6 s time resolution (dashed line) obtained by linear interpolation of the 1 min Parsivel^2 RI measurements. (b) A comparison of the reference ($RI_{r,t}$, dashed line) and predicted Pluvio^2 [$RI_{p,t}$ calculated using Eqs. (4) and (5)—solid line] RI time series. (c) A comparison of the reference RI time series ($RI_{r,t}$, dashed line) and predicted Pluvio^2 RI time series ($RI_{p,t}$, solid line) that is shifted 42 s backward in time for a closer match with the reference RI time series.

8 mm h^{-1}. That is, on average, the RI_m values are less than the corresponding RI_r values. As a percentage of the RI_r, the estimated systematic biases are −3.7% and −2.6% for the 7 and 8 mm h^{-1} RI_r, respectively. Uncertainty of the measurements, characterized by the half widths of the tolerance intervals as a percentage of the corresponding RI_r values, are larger for the smaller RI_r settings and decrease with an increase of the RI_r. The study for the uniformity of the measurement system indicated that variances of repeated measurements are unequal over the range of RI_r studied. A statistically significant linearity problem, that is a change of bias over the range of RI studied, is not observed. However, a bias test for the individual reference RI_r exhibits that the biases are statistically significant for the RI_r values of 7 and 8 mm h^{-1} at a level of $\alpha = 0.05$. Analysis of the measurements under simulated abrupt, gradual step, and cyclic step changes of the RI_r indicated that there is not sufficient evidence to claim that the means of the measurements do not stabilize rapidly on the corresponding RI_r setting after the changes, as RI_m never exceeded the control limits after disregarding the initial and intermittent minutes between the changes in RI_r. Moreover, there was no significant evidence of hysteresis effects on the measurements during the RI simulation with cyclic step changes. Our findings suggest that the lower threshold value for the OTT Pluvio^2 RI measurements should be redefined as 7 mm h^{-1}, and the bias should be considered close to the lower threshold value for RI measurements.

Investigation of the time delay in bucket weight measurements and identification of the internal algorithm in Pluvio^2 real-time RI calculation using real-time bucket weight measurement led us to develop a set of equations that demonstrate error for dynamic rainfall scenarios that are typical in the field. Implementation of those equations in two cases (Cases I and II) revealed the potentiality of large errors in Pluvio^2's RI measurements. Identification of the RI calculation algorithm and potential error sources present valuable opportunities to improve Pluvio^2 RI measurement capabilities through improvements in the algorithm; and hence, Pluvio^2's wider use for RI measurements. Furthermore, considering the time delay in bucket weight measurements and adequate relevant improvements in the RI calculation algorithm, a simple postprocessing of the available Pluvio^2 datasets may lead to valuable RI datasets with desirable accuracy for various applications.

This study clearly shows that understanding of the Pluvio^2 weighing gauge behavior and functionality with special focus on its proprietary internal algorithm is essential for all applications that utilize Pluvio^2 measurements. In regard to Pluvio^2 RI output, our results grant new studies on improving RI calculation algorithms to reduce the measurement errors. Such an algorithm development effort may consider, instead of using a constant last-minute precipitation accumulation, a dynamic time frame based on the changes in the frequency of RI to improve Pluvio^2's RI measurement performance under the dynamic rainfall conditions. An improved algorithm using Pluvio^2 datasets would enable postprocessing of the existing and new datasets of Pluvio^2 bucket weight measurements for RI calculations. This would have importance in applications where the use of RI is required, for example, verification of radar or lidar outputs where precise ground RI measurements are essential. Overall, this study provides insights on Pluvio^2 weighing gauge behavior under different RI conditions and reveals the instrument's proprietary internal calculation procedure for RI measurement, which is indispensable for RI applications. Moreover, similar studies to this one that consider a wide range of RI values, larger sample sizes, and extended time periods can be recommended for future research. Furthermore, effects of potential sources of variability in the measurement system can be studied through experimental design methods.

Acknowledgments. This research was supported by the funds provided by the National Science Foundation under Grant AGS-1741250 to the second author (FYT). The first author is a graduate student under the guidance of FYT.

REFERENCES

Colli, M., 2014: Assessing the accuracy of precipitation gauges: A CFD approach to model wind induced errors. Ph.D. thesis, School in Science and Technology for Engineering, University of Genova, 237 pp.

——, L. G. Lanza, and P. La Barbera, 2013: Performance of a weighing rain gauge under laboratory simulated time-varying reference rainfall rates. *Atmos. Res.*, **131**, 3–12, https://doi.org/10.1016/j.atmosres.2013.04.006.

——, ——, ——, and P. W. Chan, 2014: Measurement accuracy of weighing and tipping-bucket rainfall intensity gauges under dynamic laboratory testing. *Atmos. Res.*, **144**, 186–194, https://doi.org/10.1016/j.atmosres.2013.08.007.

Dossett, S. R., and V. C. Bowersox, 1999: National Trends Network site operation manual. NADP Manual 1999-01, 99 pp.

Gordon, J. D., 2003: Evaluation of candidate rain gauges for upgrading precipitation measurement tools for the National Atmospheric Deposition Program. U.S. Geological Survey Water-Resources Investigations Rep. 2002-4302, 37 pp., https://doi.org/10.3133/wri024302.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, https://doi.org/10.1175/BAMS-D-13-00164.1.

Karl, T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the United States. *Bull. Amer. Meteor. Soc.*, **79**, 231–242, https://doi.org/10.1175/1520-0477(1998)079<0231:STOPAF>2.0.CO;2.

Kidd, C., and P. Joe, 2007: Importance, identification and measurement of light precipitation at mid- to high-latitudes. *Proc. Joint EUMETSAT Meteorological Satellite Conf. and 15th Satellite Meteorology and Oceanography Conf.*, Amsterdam, Netherlands, EUMETSAT and Amer. Meteor. Soc.

Lamb, D., and V. Bowersox, 2000: The National Atmospheric Deposition Program: An overview. *Atmos. Environ.*, **34**, 1661–1663, https://doi.org/10.1016/S1352-2310(99)00425-2.

Lanza, L. G., and L. Stagi, 2009: High resolution performance of catching type rain gauges from the laboratory phase of the WMO Field Intercomparison of Rain Intensity Gauges. *Atmos. Res.*, **94**, 555–563, https://doi.org/10.1016/j.atmosres.2009.04.012.

——, and E. Vuerich, 2009: The WMO Field Intercomparison of Rain Intensity Gauges. *Atmos. Res.*, **94**, 534–543, https://doi.org/10.1016/j.atmosres.2009.06.012.

——, M. Leroy, J. van der Meulen, and M. Ondras, 2005: WMO Laboratory Intercomparison of Rainfall Intensity Gauges. WMO Rep. WMO/TD-1265, 8 pp., https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-82-TECO_2005/Papers/3(09)_Italy_Lanza.pdf.

——, E. Vuerich, and I. Gnecco, 2010: Analysis of highly accurate rain intensity measurements from a field test site. *Adv. Geosci.*, **25**, 37–44, https://doi.org/10.5194/adgeo-25-37-2010.

Larsen, A. N., I. B. Gregersen, O. B. Christensen, J. J. Linde, and P. S. Mikkelsen, 2009: Potential future increase in extreme one-hour precipitation events over Europe due to climate change. *Water Sci. Technol.*, **60**, 2205–2216, https://doi.org/10.2166/wst.2009.650.

Löffler-Mang, M., and J. Joss, 2000: An optical disdrometer for measuring size and velocity of hydrometeors. *J. Atmos. Oceanic Technol.*, **17**, 130–139, https://doi.org/10.1175/1520-0426(2000)017<0130:AODFMS>2.0.CO;2.

Montgomery, D. C., 2009: *Statistical Quality Control*. 7th ed. Wiley, 768 pp.

NADP, 2015: National Trends Network site operations manual. NADP NTN Operations Manual 2015-03, version 2.3, 19 pp., https://nadp.slh.wisc.edu/lib/manuals/NTN_Operations_Manual_v_2-3.pdf.

Nemeth, K., 2008: OTT Pluvio^2: Weighing precipitation gauge and advances in precipitation measurement technology. *WMO Technical Conf. on Meteorological and Environmental Instruments and Methods of Observation*, Saint Petersburg, Russia, WMO.

Nitu, R., and Coauthors, 2019: WMO Solid Precipitation Intercomparison Experiment (SPICE) (2012-2015). WMO Instruments and Observing Methods Rep. 131, 1445 pp.

OTT HydroMet, 2014: Pluvio precipitation gauge—Operating instructions. OTT HydroMet GmbH Doc., 52 pp., https://www.ott.com/download/operating-instructions-precipitation-gauge-ott-pluvio2-l.pdf.

Sevruk, B., and B. Chvíla, 2005: Error sources of precipitation measurements using electronic weight systems. *Atmos. Res.*, **77**, 39–47, https://doi.org/10.1016/j.atmosres.2004.10.026.

Testik, F. Y., and M. Gebremichael, 2010: *Rainfall: State of the Science*. *Geophys. Monogr.*, Vol. 191, Amer. Geophys. Union, 287 pp.

Tumbusch, M. L., 2003: Evaluation of OTT Pluvio Precipitation Gauge versus Belfort Universal Precipitation Gauge 5-780 for the National Atmospheric Deposition Program. U.S. Geological Survey Water-Resources Investigations Rep. 03-4167, 31 pp., https://pubs.usgs.gov/wri/wrir034167/wrir034167.pdf.

Westra, S., and Coauthors, 2014: Future changes to the intensity and frequency of short-duration extreme rainfall. *Rev. Geophys.*, **52**, 522–555, https://doi.org/10.1002/2014RG000464.

Willems, P., 2000: Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types. *J. Hydrol.*, **233**, 189–205, https://doi.org/10.1016/S0022-1694(00)00233-X.

WMO, 1996: Guide to meteorological instruments and methods of observation. WMO Rep. 8, 501 pp.