[©]Evaluation of Multiple Precipitation Sensor Designs for Precipitation Rate and Depth, Drop Size and Velocity Distribution, and Precipitation Type

BEN S. PICKERING,^a RYAN R. NEELY III,^b JUDITH JEFFERY,^c DAVID DUFTON,^b AND MARYNA LUKACH^b

^a Institute for Climate and Atmospheric Science, Leeds, United Kingdom ^b National Centre for Atmospheric Science, Leeds, United Kingdom ^c STFC Rutherford Appleton Laboratory, Didcot, United Kingdom

(Manuscript received 10 April 2020, in final form 6 January 2021)

ABSTRACT: Observations of the precipitation rate/depth, drop size distribution, drop velocity distribution, and precipitation type are compared from six in situ precipitation sensor designs over 12 months to assess their performance and provide a benchmark for future design and deployment. The designs considered are tipping bucket (TBR), drop counting (RAL), acoustic (JWD), optical (LPM), single-angle visiometer with capacitor (PWD21), and dual-angle visiometer (PWS100). Precipitation rates are compared for multiple time resolutions over the study period, while drop size and velocity distributions are compared with cases at stable precipitation rates. To examine precipitation type, a new index and a logic algorithm to amalgamate consecutive precipitation type observations consistently is introduced and applied. Overall, the choice of instrument for deployment depends on the usage. For fast response (less than 15 min), the PWD21 and TBR should not be used. As precipitation rate or the duration of a sample increases, the correlation of the TBR with the majority of other instruments increases. However, the PWD21 consistently underestimates precipitation. The RAL, PWS100, and JWD are within $\pm 15\%$ for precipitation depth over 12 months. All instruments are inconsistent in their ability to observe drop size and velocity distributions for differing precipitation rates. There is low agreement between the instruments for precipitation type estimation. The PWD21 and PWS100 rarely report some precipitation types, but the LPM reports more broadly. Meteorological stations should use several instrument designs for redundancy and to more accurately capture precipitation characteristics.

KEYWORDS: Drizzle; Precipitation; Snow; Drop size distribution; Mixed precipitation; Surface observations

1. Introduction

All meteorologists agree that precipitation must be recorded accurately, yet there is no consensus on the best method to do so. There are many ways to measure precipitation, both in situ or remote sensing. For remote sensing techniques, the sample volume of any single remote sensing measurement contains a population of hydrometeors that must be derived statistically from the measurement. As such, spatial variability smaller than the measurement scale is lost and important details may be obscured. Meanwhile, though surface in situ measurements are able to directly measure, they typically only sample a fraction of a square meter, which renders their observations as unrepresentative of the wider area. A combination-using surface precipitation measurements as a "ground truth" to calibrate remote sensing techniques-is commonly used in operational meteorological agencies (Fulton et al. 1998; Harrison et al. 2000; Rubel and Brugger 2009). Precipitation depth (mm) and rate $(mm h^{-1})$, distributions of drop size (mm) and velocity $(m s^{-1})$, and dominant precipitation type are five fundamental variables used to describe precipitation. While WMO intercomparison experiments have focused on the rate of liquid (Lanza and Vuerich 2009) or solid (Kochendorfer et al. 2017) precipitation, this study examines all of these variables for a single location over 12 months.

a. Precipitation depth and precipitation rate

This study will consider measured precipitation depth (PD) and precipitation rate (PR) performance from very weak to very intense precipitation $(0.05-50.0 \text{ mm h}^{-1})$. Instruments must be capable of accurately detecting the heaviest precipitation since these events have the highest impact on society through flooding. At the other end of the scale, weak precipitation can also be important. Oppenheim and Shinar (2012) showed that drivers do not reduce their speed sufficiently on wet roads. Warning signage based on automatic rainfall sensors must be able to detect the weakest precipitation events that could otherwise endanger motorists. When studying the accuracy of precipitation gauges, a ground truth reference value is difficult to obtain. This is because multiple factors can either increase (splashing, nonhydrometeors, convergent airflow, condensation) or decrease (divergent airflow, shadowing, unfilled buckets, edge hits, leaves, insects, evaporation) measured rainfall rates. This study examines 12 months of real-world data and uses methods to compare the instruments which reveals strengths and weaknesses in different use applications.

b. Drop size distribution

The drop size distribution (DSD) is a distribution of the number of drops as a function of diameter per unit volume (Jameson and Kostinski 2001), written as N(D). Instruments that are able to measure this variable are called disdrometers;

DOI: 10.1175/JHM-D-20-0094.1

^o Denotes content that is immediately available upon publication as open access.

Corresponding author: Ben S. Pickering, eebp@leeds.ac.uk

	Observing design	Sample region	Resolution	Accuracy	Rain rate range
LPM	Optical occlusion disdrometer	$46.5 {\rm cm}^2$	$1 \times 10^{-3} \mathrm{mm}$	≤15% ^d	$<0.005 \dots > 250 \text{ mm h}^{-1}$
PWS100	Dual-angle visiometer	$40.0{\rm cm}^2$	$1 imes 10^{-4} \ \mathrm{mm}$	±10% ^e	$0 \dots 400 \text{ mm h}^{-1 \text{ i}}$
PWD21	Visiometer + capacitor + temperature	$0.1 L^{a} + 7.2 cm^{2 b}$	0.01	$\pm 30\%^{f}$	$0.5 \dots 20 \text{ mm h}^{-1 \text{ j}}$
JWD	Acoustic disdrometer	$50.0\mathrm{cm}^2$	$1.3 imes 10^{-7} \mathrm{~mm^{c}}$	±16% ^g	Undefined
RAL	Drop counting	$324\mathrm{cm}^2$	$1.89 imes 10^{-3} \text{ mm}$	$<\pm 10\%^{h}$	$5 \dots 200 \text{ mm h}^{-1}$
TBR	Tipping bucket	$324\mathrm{cm}^2$	0.2 mm	Undefined	Undefined

TABLE 1. Summary of the study instrumentation specifications.

^a PWD21 visiometer sample volume in liters.

^b PWD21 capacitive plate area.

^c Calculated from 0.313 minimum drop diameter (Distromet Ltd. 2002).

^d With a range 0.5–20 mm h^{-1} (Adolf Thies GmbH 2011).

^e Rain and $0.5 \dots 20 \text{ mm h}^{-1}$ (Campbell Scientific 2012).

f "Accuracy will be degraded in windy conditions, for frozen precipitation and very high rainfall rates" (Vaisala Oyj 2001).

 $^{g}\pm5\%$ of measured drop diameter (Distromet Ltd. 2002), assuming spherical raindrops for \pm 16% volume.

^h At >20 mm h⁻¹ rain rate and at 10-s sampling period (Agnew 2014).

ⁱ Marshall–Palmer distributed (Campbell Scientific 2012).

^j For liquid precipitation (Vaisala Oyj 2001).

for disdrometers, the DSD defines every subsequent calculated variable such as rainfall rate, equivalent radar reflectivity factor and precipitation type. The first DSDs were measured with either ink-dusted paper or trays of flour (Wiesner 1895; Bentley 1904; Laws and Parsons 1943) which led to the Marshall–Palmer relations of DSD to rainfall rate R and radar reflectivity Z (Marshall and Palmer 1948). DSD-derived Z-R relations are commonly used to convert radar reflectivity factor into quantitative precipitation estimates (Joss and Waldvogel 1970; Battan 1973; Jorgensen and Willis 1982; Maeso et al. 2005). Therefore, the importance of the accuracy of in situ surface observations of DSD cannot be understated.

c. Drop velocity distribution

The drop velocity distribution (DVD), written as N(V), is the number of drops as a function of drop velocity. DVD is important because it allows the total kinetic energy of precipitation to be considered. Our understanding of soil erosion (Kinnell 1981; Rosewell 1986), building erosion (Tang and Davidson 2004; Erkal et al. 2012), vertical winds and downbursts caused by mass loading (Feingold et al. 1991), and the erosion of aerofoils like planes and wind turbine blades (Keegan et al. 2013; Slot et al. 2015; Eisenberg et al. 2018) benefits from more accurate observations of DVD. Some instruments also rely on the fall velocity of the particle to distinguish its precipitation type, since different hydrometeors have different aerodynamic properties (Gunn and Kinzer 1949; Locatelli and Hobbs 1974).

d. Precipitation type

Precipitation type (henceforth PT) describes the dominant phase, shape, and density of hydrometeors within a volume of the atmosphere that fall to the surface. The present weather (PW) code is used to represent PT, which presents problems for bulk statistical analysis. The codes are a qualitative description of the PW type. Table 4680 from the World Meteorological Organization is the standard for automatic sensors (WMO 1988, 2017), which has 99 entries. Some hydrometeor types are represented by several PW codes—typically variations of intensity and longevity—whereas this investigation only concerns the type of hydrometeor detected. The motivation for the PW code's existence was to reduce the bandwidth of descriptive information, only upheld today for consistency with existing data.

PT can have great impact on transportation, agriculture and infrastructure but is poorly forecasted (Ralph et al. 2005; Reeves 2016). PT has become more prominent in the field of operational meteorology in the last decade as operational radar networks have gained dual-polarization capabilities (Park et al. 2009; Saltikoff and Nevvonen 2011; Al-Sakka et al. 2013). PT information within a radar sample can also be used to provide more accurate quantitative precipitation estimates (QPE) because Z–*R* relationships vary with hydrometeor class (Atlas and Ludlam 1961; Harimaya 1978; Fujiyoshi et al. 1990).

2. Experimental conditions

The following sections briefly describe the instruments as summarized in Table 1, and recorded variables in Table 2. This is followed by a summary of the experimental conditions onsite, and finally the broader structure of this study.

a. Instruments

1) THIES LASER PRECIPITATION MONITORS

Two identical Thies Laser Precipitation Monitors (LPMs) are used in this study (henceforth Thies1, Thies2, or the Thies LPMs). The Thies LPMs are both part of a wider network of 14 Thies LPM instruments called the Disdrometer Verification Network (DiVeN), described by Pickering et al. (2019a). The Thies LPM emits an infrared beam which is received by a photodiode. As precipitation or other particles cross the path of the beam, the electrical signal produced by the photodiode is reduced (Adolf Thies GmbH 2011). The amplitude and duration of signal reduction is analogous to the size and speed of the particle. Löffler-Mang and Joss (2000) describe this optical occlusion method of observation. Observed particles are sorted into 20 diameter bins

TABLE 2. Summary of the instruments, variables, and time aggregations used in this study. Precipitation data are scrutinized from six different instrument designs, five precipitation variables, and over several time scales. The precipitation variables each have their own methodologies for data analysis and are thus presented separately in each section. For the DSD and DVD variables, cases of 5 min, 15 min, and 1 h are presented. The 1-min distributions are rarely used because of known sampling issues (Tapiador et al. 2016; Smith 2016), and precipitation events longer than 1 h typically do not retain consistent distribution characteristics. The 1-min, 5-min, 15-min, 1-h, and daily time durations of PR are considered, and PT analyses are performed for 1-min, 5-min, and 1-h intervals (bulk statistics, not case studies).

	Thies1	Thies2	PWS100	PWD21	JWD	RAL	TBR	Total	Durations
PR and PD	\checkmark	\checkmark	\checkmark	~	1	\checkmark	1	7	1 min, 5 min, 15 min, 1 h, 1 day
DSD	1	\checkmark	\checkmark		\checkmark			4	Cases of 5 min, 15 min, 1 h
DVD	1	\checkmark	\checkmark					3	Cases of 5 min, 15 min, 1 h
PT	\checkmark	\checkmark	\checkmark	\checkmark				4	1 min, 5 min, 1 h

(from ≥ 0.125 to >8 mm) and 22 velocity bins (from >0.0 to >20.0 m s⁻¹), and 21 out of 93 total PW codes are supported for six PTs (drizzle, rain, mixed rain/snow, snow grains, snow aggregates, hail). Nonhydrometeors (insects, debris) can appear as precipitation and the housing of the instrument is a surface on which precipitation can rebound into the beam. Partial beam hits are accounted for with internal processing.

2) CAMPBELL SCIENTIFIC PWS100

The Campbell Scientific Present Weather Sensor 100 (PWS100) uses a forward-scattering technique using four light beams and two receiving diodes: one diode at a vertical angle only and one diode at a combined horizontal and vertical angle from the emitted beams. Drop size distributions are measured in 0.1-mm bins from >0.0 to 30.0 mm but the manual states that the "proportion of particles detected will fall off significantly below about 0.5 mm diameter." Drop velocity distributions are recorded in nonlinearly spaced bins from >0.0 to >25.6 m s⁻¹. Both diameter and velocity measurements have a quoted accuracy of $\pm 5\%$ (for liquid particles >0.3 mm). Optical scattering characteristics differ between solid and liquid hydrometeors and integrated temperature and humidity sensors assist the determination of PT. Hydrometeor types are reported explicitly: drizzle, rain, snow grains, snowflakes, hail, ice pellets, graupel, and 58 out of a total 93 PW codes (Table 4680) are supported. The limitations of the PWS100 are similar to those of the Thies LPM.

3) VAISALA PWD21

The Vaisala Present Weather Detector PWD21 (Vaisala Oyj 2001) combines an optical forward-scattering sensor (875-nm peak, single emit and receive diode) with a resistive capacitive plate indicating water content, and a temperature sensor. Heating elements evaporate condensation and melt solid hydrometeors. PT and current weather are reported using Table 4680 (WMO 1988), of which 42 out of 93 descriptors are supported—seven types of precipitation (drizzle, rain, freezing drizzle, freezing rain, snow aggregates, mixed rain/snow, ice pellets). The capacitive plate is only used to discriminate between rain and snow between $0^{\circ}C < T < 6^{\circ}C$. Outside of this range, the types are overruled by temperature ($T \le 0^{\circ}C = \operatorname{snow}$, $T \ge 6^{\circ}C = \operatorname{rain}$).

4) JOSS-WALDVOGEL RD-80

The Joss–Waldvogel Disdrometer (JWD) has an exposed Styrofoam cone atop a spring-loaded transducer. The kinetic energy applied to the transducer is related to the size of a hydrometeor (Joss and Waldvogel 1967), and is recorded into 127 nonlinearly spaced size bins (from ≥ 0.313 to 5.145 mm with an accuracy of $\pm 5\%$). The terminal fall velocity of the particle species must be assumed to convert kinetic energy into mass and then diameter, so there are no DVD measurements from the JWD. Different hydrometeors have different diameter– velocity relationships, the JWD only functions accurately for one type of hydrometeor because it has no way to distinguish PT to then change the internal processing. By default, the JWD assumes all particles to be liquid. Snowflakes (slower terminal velocities) appear as small raindrops, and hail (faster terminal velocities) appear as large raindrops.

5) RAL DROP COUNTING GAUGE

The Rutherford Appleton Laboratory drop counting rain gauge (RAL) funnels liquid precipitation into a reservoir, from which drops overflow through a precision tube, occluding an optical infrared sensor (Norbury and White 1971). Occlusions are tallied over time to generate a rain rate. A smaller minimum resolution than the TBR should provide greater sensitivity in light rainfall, but the funnel can still disturb the airflow and become blocked. Above 50 mm h⁻¹ droplets begin to merge, so occlusions become nonlinear with increasing rainfall rate. A continuous stream of fluid cannot be measured.

6) MUNRO 0.2-MM TIPPING-BUCKET RAIN GAUGE

The Munro 0.2-mm tipping-bucket rain gauge (TBR) funnels liquid precipitation into a pivoting double bucket, which limits the temporal representation of PR. A single tip over 1 min represents a 12 mm h⁻¹ PR; over 5 min 2.4 mm h⁻¹. Due to this, TBR PRs are typically only used for \geq 15-min durations (1 tip = 0.8 mm h⁻¹). In addition, TBRs are only able to funnel liquid precipitation (when sufficiently wet) into the tipping bucket; snow, hail, and graupel must melt to be detected with delay. A heated funnel consumes significantly more energy and limits where such sensors may be deployed. The TBR alone cannot distinguish between PTs. Furthermore, the TBR can be easily rendered unusable if debris blocks the funnel. The Munro is similar in design to gauges used throughout the fields of hydrology and meteorology and thus represents this category of design.

b. Study location

All instruments were installed at the Chilbolton Atmospheric Observatory (CAO), Hampshire, United Kingdom. The climate



FIG. 1. Aerial view of the Chilbolton Atmospheric Observatory (CAO), showing the instruments used in this study and their proximity. (a) Vaisala PWD21, (b) Campbell Scientific PWS100, (c) Thies Clima LPM 1, (d) Thies Clima LPM 2, (e) Munro 0.2-mm tipping-bucket gauge, (f) Joss–Waldvogel RD-80, and (g) RAL drop counting gauge. The furthest distance between any two instruments is 80 m.

at CAO is temperate and maritime where the weather can fluctuate on a daily basis. Proximity to the polar front jet stream ensures that midlatitude depressions impact the site regularly. Precipitation is therefore frequent, and intermittent rainfall is the modal type. Air temperature extremes are uncommon due to the U.K. being a small landmass, surrounded by ocean and situated on the northeastern Atlantic, influenced by the warmth of the North Atlantic Current. Met Office (2016) provides a climate summary for Southern England and shows that on average (1981-2010) 12 days per year have falling snow and 109.5 days have precipitation accumulation of \geq 1.0 mm. Ventouras et al. (2006) contains more information on the climatology of PRs observed at CAO. The frequency and variation of precipitation events as well as the occurrence of stratiform, convective, and different PTs make CAO a suitable location to conduct this study because it covers a wide range of precipitation conditions seen worldwide.

c. Instrument installations

Figure 1 shows the installation of each instrument at CAO. The JWD is situated inside a circular pit slightly below ground level to reduce turbulence over the instrument whereas the TBR and RAL are on the ground. Both Thies LPM beams are at 1.5 m, whereas the PWD21 and PWS100 are at 8 and 9 m, respectively, mounted on 1- and 2-m poles above a 7-m-tall building. All of the instruments are spatially within 80 m, are installed within manufacturer recommendations and meet

WMO standards where practical. The agricultural land surrounding CAO is flat for at least 500 m in all directions. The Chilbolton Advanced Meteorological Radar (CAMRa) dish (30 m tall) is 80 m east of the instrument enclosure; visible in Fig. 1. The radar should not lead to bias of the observations since prevailing wind direction at the site is westerly and the structure is sufficiently distant.

Both the climatic conditions and the physical mounting conditions of the instruments at CAO makes the results of this study applicable to a wide range of precipitation and monitoring conditions at other midlatitude locations. Therefore the performance of the instruments in this study should be comparable to similar instrumentation located at analogous locations.

d. Data quality

All instruments used in this study have been calibrated as per the manufacturer standards. The TBR and RAL are calibrated dynamically [as described in Humphrey et al. (1997)] with a pump at different flow rates. The TBR was calibrated by the manufacturer in 2011, verified on site in 2018 and in 2020 both with no corrections to the manufacturer settings needed (within $\pm 10\%$). The RAL was last calibrated in 2013 and was checked in 2020 with no corrections to the manufacturer settings needed (within $\pm 10\%$). We therefore conclude that the TBR and RAL instruments are reasonably calibrated over the study period. Note that calibration methods do not take into account the effect of the gauge funnel. The PWD21 and



FIG. 2. Availability of usable, quality-controlled data during the 12-month period under examination (10 Feb 2017–9 Feb 2018, totaling 365 days). Each day is a vertical strip with the opacity representative of the percentage of time successfully recorded. The lower bar (ALL) represents the combination of all instruments above, and represents the data under study since all instruments must be operational for a fair comparison. Any times with one or more instruments nonoperational are excluded. The Thies LPMs had a period of downtime in September 2017 resulting in approximately 20 days without data. The missing data from the TBR and RAL gauges were due to the gauge becoming temporarily blocked with debris. For the PWD21, the fault at the start of February 2017 was due to a failure in datalogging. The remaining missing files were due to temporary, intermittent logging, or file formatting problems. The minutes where all instruments were simultaneously recording during the study sum up to 309.8 days, which is equivalent to 84.9% of the study period.

PWS100 have no standardized calibration method for the optical scattering technique, and therefore the manufacturer calibration cannot be verified on site. The capacitive plate on the PWD21 also has no standardized calibration method. The manufacturer calibration is therefore relied upon for the PWD21 and PWS100-which is also true of any deployment of these sensors and is not unique to this study. Therefore, this study is using a calibration comparable to typical installations of these two instruments. Similarly with the JWD, there is no standardized calibration method for on-site verification. The JWD manufacturer recommends that calibration is repeated every 6 years (Distromet Ltd. 2012), which this study period is within, so the instrument is said to be calibrated. The Thies LPMs were calibrated by the manufacturer in 2011 and again no standardized method exists for verification on-site. The Thies LPM manual also states that a calibration is then only necessary when a component is changed (Adolf Thies GmbH 2011), so the Thies LPMs are said to be calibrated during this study period. In addition, all rain gauges at CAO are monitored on an ongoing basis using all rain data recorded at the site. Any faults or any relative changes in sensitivity of the gauges are investigated and suspect measurements are removed from the datasets before publication to the Centre for Environmental Data Analysis (CEDA). Figure 2 shows the availability of each instrument during the study period.

3. Methodology

The following sections explain the methodologies employed in this study, split by each precipitation variable, since each variable is unique and requires a unique approach. The overall goal is to compare the CAO instruments for all precipitation variables that are important to operational and research meteorology today.

a. Precipitation depth and precipitation rate

PR measurement techniques are subject to random and systematic errors, affecting the exact determination of PR at ground level. As such, there is no measurement at the site that could be considered a reference. The instruments will be compared and discussed in context of their measurement design to compare and contrast the strengths and weaknesses of each. Sample duration is also an important consideration when examining PR measurements. Some uses and applications of PR data need only daily values while others require much shorter PR time scales, so several durations will be considered.

The comparison of PDs uses the full 12 month dataset described above. To examine bulk behaviors and general differences in the observations of the instruments, the cumulative precipitation measured over the study period is compared first. We compare both the raw accumulations and the deviation of the accumulations of each instrument from the mean of the instrument ensemble.

We then compare the PR from each instrument, performed using simple linear regression by least squares applied between two instruments. The observations should have a 1:1 relation through the origin if both instruments have similar characteristics. Gradient and intercept values of the regression reveal differences in behavior of the instruments. Each instrument is compared with every other instrument for five time durations. In each regression, both of the instruments being compared must detect precipitation in order for that datum to be included, which removes major anomalies but favors insensitive instruments.



FIG. 3. PRs from Thies1 for the chosen (a) 5-, (b) 15-, and (c) 60-min cases.

b. DSD and DVD

DSD and DVD measurements are difficult to analyze using bulk statistics because they are nonlinear. To facilitate comparison, DSDs can be parameterized into a gamma or generalizedgamma model (Marshall and Palmer 1948; Ulbrich 1983; Thurai and Bringi 2018) in the form

$$N(D) = N_0 \exp(-\Lambda D) (0 \le D \le D_{\max}), \tag{1}$$

where N_0 is the intercept, Λ the slope parameter, and D the drop diameter. The advantage of parameterization is that the comparison is simplified to a purely numerical one, but in doing so the artifacts or biases in the data are potentially concealed. For example, an overestimation of medium-diameter drops combined with an underestimation of small-diameter drops could result in the same slope parameter as for an instrument without biases. There is also uncertainty on the gamma model performance for drop diameters < 1.0 mm. Chandrasekar et al. (2003) suggest that the reduction in counts of small drops in the gamma model is a consequence of a parametric fit on experimental data that underdiagnoses small drops. Therefore this study compares measured DSDs and DVDs without parameterization.

Here a case study approach is used, similar to that of Tokay et al. (2013) but with some improvements. In Tokay et al. (2013), the cases are all 1 h in duration, the PR varies throughout, and the average PRs range from 1.8 to 12.6 mm h⁻¹, so less common (very weak and very heavy) PR-DSDs are not explicitly represented. Here cases are chosen based on constant PR throughout 5-min, 15-min, and 1-h periods so that the observed DSD/DVD remains consistent. The PRs used to identify cases are: 0.05, 0.3, 1.0, 2.0, 5.0, 20.0, and 50.0 mm h⁻¹.

Thies1 measurements are used to select cases. The absolute deviation from the desired PR is calculated in a moving window of the desired duration and the period with the least absolute deviation is selected as the case from the entire dataset. Figure 3 shows the PR over the cases selected. A 50 mm h^{-1} PR does not occur consistently for 15 min or 1 h, and a 20 mm h^{-1} PR does not occur consistently for 1 h. In total there are 18 cases. When averaged, the 1-h weakest rainfall rate (case 15 in Figs. 5 and 6) is approximately 0.02 mm h^{-1} .

Each instrument uses different DSD and DVD bin widths. The Thies LPMs have the broadest bin widths so the data from the PWS100 and JWD are mean-weight normalized into the Thies LPM bins. The upper and lower limits of the JWD are narrower than the Thies LPMs and result in partially filled bins after normalization that are discarded. The effective measurement range of the JWD in this analysis is therefore shortened to 0.375–5.0 mm. The DSD and DVD data have also been normalized by instrument sampling area.

c. Precipitation type

PT observations are recorded as PW codes, which are difficult to compare because the data are categorical and their interpretation is ambiguous. PT is often not explicitly described by a PW code. To simplify the analysis, a new PT scheme has been created (shown in Table 3) with translations from PW codes, and is broadly ranked by impactfulness to society.

Wexler (1955) notes that empirical evidence by Langleben (1954) demonstrates that a snow aggregate retains its velocity characteristics until >90% of the mass has melted. Lumb (1963) further showed empirically that snowflakes are reported by trained observers up to $1.5^{\circ}-3.0^{\circ}$ C, corresponding to approximately 300–500-m penetration below the 0°C dry-bulb isotherm. Wet types, therefore, occur in a wide range of temperatures and are important to include.

All 1-min PW code data from the study period are converted into PT as an initial step. Confusion matrices are produced which tally occurrences of each PT between instruments. For numerical comparison, a statistical score for evaluating instrument agreement is needed. Rather than comparing individual

TABLE 3. A master lookup table (LuT) for hydrometeor type. The addition of a "wet" descriptor (defined as the presence of additional liquid water with solid PTs) allows for solid PTs to be in the melting phase, where some hydrometeors have begun to melt but others have not, or may be a solid center with a liquid coating. Not all hydrometeor types exist in the present weather (PW) code Table 4680 (WMO 1988, 2017); the master LuT is designed for compatibility with radar hydrometeor classification algorithms.

ABC	РТ	Description	PW Codes	Comments
Er	-2	Instrument error	Undefined	Instrument offline/data corrupt
Un	-1	Unidentified	40-48,80	Hydrometeor detected but type unknown
No	00	No hydrometeor	0, 4, 5, 10, 20-25, 30-34	Includes fog, mist, haze, smoke
Dr	01	Drizzle	50-56	
DrRa	02	Drizzle and rain	57, 58	
Ra	03	Rain	60–66	
Ic	04	Ice	77, 78	Pristine crystals/needles; no aggregation
WIc	05	Wet ice	Undefined	Presence of liquid water
Sn	06	Snow	70–73	Aggregated ice crystals
WSn	07	Wet snow	67, 68	Presence of liquid water
Gr	08	Graupel	Undefined	A.K.A. soft hail or snow pellets
WGr	09	Wet graupel	Undefined	Presence of liquid water
Ha	10	Hail	74–76, 89	A.K.A. ice pellets (density defined)

PTs between instruments, a multiclass score is desirable; the multiclass adaptation of the Heidke skill score (HSS) is used:

$$HSS = \frac{\sum_{i=1}^{l} p(y_i, x_i) - \sum_{i=1}^{l} p(y_i) p(x_i)}{1 - \sum_{i=1}^{l} p(y_i) p(x_i)},$$
(2)

where $\sum_{i=1}^{I} p(y_i, x_i)$ is the proportion correct, $\sum_{i=1}^{I} p(y_i) p(x_i)$ is the random proportion correct, 1 is the perfect forecast score, I is the length of the confusion matrix, y_i is the *i*th row, and x_i is the *i*th column (Heidke 1926). HSS indicates the fractional improvement in agreement over a randomized observation set, which would score zero. The worst possible score is $-\infty$, and the best score is 1; negative values indicate that a random guess would have been more skillful. The HSS is symmetric, e.g., Thies1 versus PWD21 yields the same score as PWD21 versus Thies1. HSS tends to decrease for more complex PT schemes with more classes. The PT scheme introduced here has 10 classes which are explicit (i.e., more classes would not add value to a user). HSS is also calculated for a simplified scheme: none, liquid, mixed, and solid, to illustrate the range of HSS that can be obtained by using a simplified scheme that is more ambiguous to a user.

The instruments used in this study do not output PW codes at consistent time periods; 5-, 10-, 15-, and 60-min intervals are used by some but not by all instruments, but all instruments do output a 1-min interval. Evaluating longer time scales of PT is useful because radars often operate on a 5-min (or longer) sampling strategy, and some weather reports are conducted hourly. The instruments in this study employ different and poorly documented methods to merge 1-min observations.

For fair comparison between instruments, a new algorithm to merge 1-min PT observations into PT assessments of longer duration is developed. A set of Boolean logic criteria are applied sequentially and are outlined in Fig. 4.

4. Results

a. Precipitation depth and rate

Over the entire 12-month period, Fig. 5 shows the variation in PD recorded by the instruments. The PWD21 reports much less (280.8 mm) precipitation than the average for the period (542.8 mm). The JWD records more accumulation than the TBR and remains closest to the mean of all instruments, but this should not be confused with being the most accurate. The PWS100 has a positive bias in July but this is counteracted with a negative bias at the end of September, resulting in a final accumulation close to the average. The RAL and Thies1 agree until November where the RAL develops a positive bias until the end of the period and records 618.8 mm in total. The Thies LPMs disagree, with Thies1 and Thies2 recording 591.7 and 744.8 mm, respectively (26% difference).

Next the PRs are compared, first with a focus on the 5-min time period (operational radar periodicity) before other time periods are considered. In Fig. 6, every comparison with the PWD21 has a gradient of between 0.28 < m < 0.42 and the lowest coefficients of any instrument. The comparison reveals that the PWD21 is consistently measuring less rainfall than the other instruments. Similarly, the TBR records less precipitation than the PWS100, Thies LPMs, JWD, and RAL but more precipitation than the PWD21, and has the second lowest r^2 values. Furthermore, the TBR versus the PWD21 has the lowest r^2 value of any case, indicating that the TBR and PWD21 are also inconsistent in their underestimations. The TBR has a resolution of 0.2 mm, which in a 5-min sampling period equates to a 2.4 mm h⁻¹ minimum rainfall rate, hence the TBR data are aliased (insufficient sampling frequency for the signal frequency) at this time resolution. The intercept values for the TBR tests have large deviations from the origin which highlights the poor performance of the TBR in weak rainfall.

The remaining instruments (RAL, PWS100, JWD, Thies LPMs) have higher coefficient values of $0.89 < r^2 < 0.95$. The Thies2 generally has slightly higher coefficient values than the



FIG. 4. A Boolean algorithm to take multiple precipitation type data periods and merge them into longer periods. Any hail, graupel, or wet graupel in the input data results in that type as the output. This is because these events are rare, high impact, and short duration, which may not meet the next criterion in the algorithm. Next, if any single PT occupies \geq two-thirds of the input data, that is the output. Third, the minutes with no precipitation are excluded. Of the remaining data, if \geq one-third contain rain or any wet PT, then the highest wet PT code index that exists in the input data is used.

Thies1, which is also observed in Fig. 5. The gradient and intercept of the RAL and JWD consistently show a slight negative bias compared with other instruments. The JWD–RAL comparison shows that both instruments must have similar observational characteristics because the r^2 of their comparison is high (0.947). Since the JWD is shown to underestimate drop counts during heavy PRs (section 4b), the RAL must also be underestimating PR during heavy precipitation events.

In general, Fig. 7 shows an increase in agreement as the sample duration increases. This is to be expected due to the decrease in the influence of random error as the sample duration increases. Though Fig. 7 demonstrates that using longer time periods results in higher r^2 values overall, the difference

between 1 h and 1 day for some instruments does not change or even slightly decreases. The RAL, JWD, and PWS100 all have similar r^2 values throughout the time periods. At a 1-h interval, the TBR has equivalent coefficients to the other instruments (excluding the PWD21). The PWD21 reaches a maximum r^2 value of 0.8 at a 1-day interval, but this is still far below the other instruments. The Thies2 has slightly higher albeit similar coefficients of determination as the RAL, JWD, and PWS100, which is unexpected considering that the Thies LPMs use the same instrument design.

b. Drop size distribution

Overall, Fig. 8 shows an increase in the steepness of the right tail of the DSD with increasing PR, as expected from Marshall and Palmer (1948). For more detailed analysis we split the results into three parts.

1) SMALL DROP SIZES (D < 1.0 mm)

The PWS100 records fewer drops than the other instruments for small drops in all PRs and durations; multiple orders of magnitude less in the smallest sizes (<0.8 mm). In case 15 the PWS100 barely detects the precipitation compared to the other three instruments. Case 15 was intended to represent the 1 h, 0.05 mm h^{-1} PR scenario. However, in Fig. 3 it is noted that the actual PR in case 15 is slightly lower than the desired amount, due to the method of finding stable PR cases. The actual PR in case 15 is 0.02 mm h^{-1} and this slightly lower PR, combined with the PWS100 being unable to count small drop diameters, is likely the cause of the PWS100 barely detecting the precipitation. The Thies LPMs agree with the JWD for small drop diameters up to PRs of 2 mm h^{-1} , where the Thies LPMs record more small droplets than the JWD. Above the 2 mm h^{-1} cases, the JWD count values for small drops increasingly deviate from the Thies LPMs and more toward the PWS100, which is known to underestimate small drops across all PRs.

2) MEDIUM DROP SIZES (1.0 < D < 3.0 mm)

The spread of medium-sized drop counts measured by all instruments is within an order of magnitude, showing broad agreement. For the medium-intensity PRs (1.0, 2.0, 5.0 mm h⁻¹), the PWS100 typically records equal or more drops in the 1–3-mm diameter region of the distribution. The largest deviation is in case 17 where the PWS100 records 50 000 drops, Thies1 and JWD 10 000 drops, and Thies2 11 000 drops at 2 mm. The JWD agrees with the Thies LPMs for medium-sized drops in weak PR (0.3 mm h⁻¹), but in three of the cases (10, 17, 18) records the least drops in the 2.0–3.0-mm size bin (approaching one order of magnitude toward 3.0 mm drops) compared to the Thies LPMs and PWS100. Case 10 has the largest spread in medium-sized drop counts, which is one order of magnitude at 2 mm.

3) LARGE DROP SIZES (D > 3.0 MM)

The Thies LPMs and JWD are in agreement for large drops, except during the more intense rain rates (20, 50 mm h⁻¹), where the JWD records 1–1.5 orders of magnitude fewer large drops than the Thies LPMs and PWS100. The JWD also records the largest drops to be in the 4.0–4.5-mm size bin, and none in the 4.5–5.0-mm size bin in any of the cases. The largest



FIG. 5. Long term precipitation accumulation behavior between the different instruments. (a) The total precipitation accumulation and (b) the bias, using the mean of all sensors as the baseline. The total accumulation for each instrument is listed for each instrument in decreasing order in the legend of (a), and the total deviation from the mean of all instruments (542.8 mm) at the end of the period is listed in order in (b). The accumulation measured after 12 months ranges from 262.0 to 744.8 mm (284% difference), with a range of bias from -280.8 to +202.1 mm.

difference between the Thies LPMs is at 4.25 mm in case 6 (5 min, 20 mm h⁻¹) but only approximately 100 drops. The PWS100 is in agreement with the Thies LPMs in most cases for large drops. In four of the cases (12, 17, 18, 19) the large drop counts are marginally higher, but above 4.5-mm diameter drops, the counts from the PWS100 are less than the Thies LPMs. Out of the 18 cases, the Thies LPMs record the largest drop diameters out of any instrument five times and equal largest drop diameters six times, occurring more often in the higher PRs. The PWS100 records the largest drop diameters in seven of the cases, occurring more often in weak PRs.

c. Drop velocity distribution

Figure 9 shows the DVD results, in the same cases and layout as in Fig. 8, but without the JWD which cannot measure drop velocity. Broadly the PWS100 counts fewer particles in total (the cases are identical to Fig. 5) and has a DVD upper tail which ends at lower velocities in all cases compared to both of the Thies LPMs. The DVD lower tail and peak counts of drop velocities have less than an order of magnitude of spread between instruments, with the notable exception of cases 6, 8, and 15. Cases 8 and 15 are again affected by the PWS100 undercounting small drops. The two Thies LPMs agree to within half an order of magnitude (except in case 15, very weak precipitation) but Thies2 records more drops $> 4 \text{ m s}^{-1}$ than Thies1 in almost every case, and records the highest drop velocities in eight cases (equal to Thies1 in eight cases, less than Thies1 in two cases).

d. Precipitation type

Figure 10 shows the agreement between instruments on coincident observations of PT at 5-min intervals (operational radar periodicity). The most striking result is that the Thies LPMs report more PTs than the PWS100 and PWD21. The PWS100 never reports wet ice, and only observes two periods of ice and one period of hail. The PWD21 never reports drizzle

5 Minute Precipitation Rate	PWD21		RAL		TBR		PWS100		JWD		Thies2	
Thies1	0.581		0.911		0.715		0.907		0.927		0.946	
1111031	0.33 0.	15	0.88	0.20	0.51	1.83	1.08	-0.10	0.79	0.06	1.05	0.09
Thios?	0.628		0.938		0.730		0.900		0.951			
mesz	0.31 0	09	0.83	0.08	0.49	1.67	0.98	-0.20	0.73	0.00		
	0.627		0.947		0.710		0.919					
JWD	0.42 0.	09	1.10	0.08	0.67	1.67	1.33	-0.22				
DW/9100	0.574		0.890		0.730							
FWSTUU	0.28 0.	20	0.75	0.38	0.45	1.94						
трр	0.351		0.7	22								
IDN	0.37 -0	.07	1.18	-1.10								
DAL	0.604				•							r^2
RAL	0.36 0.	09									т	с

FIG. 6. Every instrument compared with another for 5-min-averaged PR. Within each cell, the large upper value is the coefficient of determination r^2 , while the lower left and lower right are the gradient *m* and *y*-axis intercept *c* of the linear least squares regression line, respectively. The cells are shaded based on the r^2 value. The instruments in the columns are the *y*-axis data and the instruments in the rows are the *x*-axis data, such that the uppermost left result of $r^2 = 0.581$ is taken from a scatterplot of Thies1 on the *x* axis and PWD21 on the *y* axis. Hence, the gradient of Thies1 vs PWD21 (m = 0.33) indicates that the PWD21 records a third of the rainfall that the Thies1 recorded during the 12-month observation period.

and rain, ice, wet ice, or hail. Along the diagonal boxes (bordered in white) the highest agreement between instruments occurs in the no precipitation class (>98% across the board), the rain class (>80%) and the snow class (>30%).

The drizzle class has broad results; both of the Thies LPMs count more drizzle than rain. The PWS100 and PWD21 record \sim 20 times fewer drizzle periods than the Thies LPMs, and yet in the lower right matrix (Fig. 10f), both instruments rarely (6%) agree on the times of drizzle despite having similar total counts. The PWS100 frequently identifies rain when the other instruments identify drizzle (8%–15%) or drizzle and rain (80%–90%) classes.

The spuriosity of the Thies LPMs is evidenced by the first two rows and columns in Fig. 10a, where the none and drizzle classes all contain other hydrometeors. Wet snow is equally agreed upon as it is disagreed as rain by both Thies LPMs (~40% each). The PWD21 has less spuriousness because it is unable to resolve many hydrometeor types-evident with the many zero count rows (gray boxes) in Figs. 10c, 10e, and 10f. The PWS100 typically classifies Thies LPM drizzle as no precipitation (73%–84%) or rain (8%– 16%), Thies LPM ice as no precipitation, and Thies LPM hail as snow, rain, or none. The snow has some agreement compared to the other instruments (32%-39%) but often the PWS100 classifies Thies LPM snow as no precipitation (52%-64%). The PWS100 and PWD21 have large disagreements, as do the two LPMs with themselves and the other instruments, highlighting the difficulty in observing PT. The HSS from each comparison for each time period examined is summarized in Fig. 11.

Unsurprisingly, Fig. 11 shows the two identical Thies LPM instruments have the highest agreement. Second highest are the PWS100–PWD21 matrices. For the other comparisons, the Thies-PWD21 agrees less than the Thies-PWS100. The vertical lines on Fig. 11 show the HSS if

the confusion matrices were to be simplified into four classes: none, liquid, mixed and solid, where mixed are any wet-denoted type. Solid includes ice, snow, and hail. HSS increases modestly across the board if a simpler hydrometeor class system is used. However, this would not help in instances of hail and snow where the two classes are merged, nor does it indicate which type of mixed precipitation is occurring.



FIG. 7. Coefficient of determination (r^2) used as an indicator of instrument agreement between PRs measured by the labeled instrument and Thies1. Values are shown for multiple time durations to indicate the dependence of time interval on agreement. Data from the whole study period is binned into time durations labeled, and a coefficient is calculated using least squares linear regression as explained in section 3a. Thies1 is used as the baseline instrument because Fig. 6 showed it to be similar to the Thies2, RAL, PWS100, and JWD, and it allows the Thies LPMs to be compared.



Unauthenticated | Downloaded 04/07/21 10:39 AM UTC





FIG. 10. Confusion matrices of hydrometeor type for 5-min time intervals. Values shown inside the matrix are percentages normalized by the total number of observations of the type in that column, such that the instrument listed on the top of the matrix is considered truth in each plot. For example, for the first matrix, for the "wet snow" precipitation type (denoted WSn), 42.9% of the observations made by Thies2 agreed with Thies1. Looking further up in that same column, it shows that 44.4% of the Thies1 "wet snow" events were classified as "rain" by Thies2. The color intensity of each cell from white to dark red corresponds to the percentage values written inside each cell. Totals observed by an instrument are shown in the white boxes surrounding the matrix, and the total number of observations in the black lower right box, which differs for each matrix since both instruments in a matrix must report a PT (section 3c). The multiclass HSS (Heidke 1926) of each confusion matrix is shown in red, in the lower right corner outside each matrix.



FIG. 11. Multiclass HSS between each instrument capable of PT observations. Scores are shown for 1-, 5-, and 60-min intervals. The vertical lines indicate the improvement in score if a simpler hydrometeor class scheme is used (none, liquid, mixed, solid). The PWS100 is shortened to PWS here, Thies1 to T1, Thies2 to T2, and the PWD21 to PWD for readability.

5. Discussion

The performance of the instruments depends on the user. This section will consider applications in context with the results seen in the study to inform existing and future sensor deployments and design.

a. Precipitation depth and rate

The results confirm that the PWD21 is unable to measure over 20 mm h^{-1} ; no precipitation is observed over this rate, and convective events are significantly undercaptured. Concurrently the sensitivity of weak PRs is poorly captured, evidenced by a positive intercept when compared with all other instruments at 5-min sample duration (except the TBR which has time resolution limitations). The PWD21 should not be used as a PR or PD sensor where possible. Funnel gauges (TBR and RAL) are affected by air temperature and humidity. The funnels will take longer to wet in the summer because precipitation can more rapidly evaporate. There is evidence for this seasonality as Fig. 5 shows that the largest increase in bias in the TBR occurs in June-August, and the RAL bias increases more than the other instruments from November onward. The TBR can also lose liquid within a bucket through evaporation, which the RAL mitigated by employing a less exposed drop collection reservoir design. This may explain the larger summer differences in the TBR versus the RAL. Funnel gauges are also affected by wind (Sevruk 1996).

The Thies LPMs have a large difference in recorded precipitation over 12 months (591.7 and 744.8 mm for Thies1 and Thies2, respectively, or 25% more in Thies2). The Thies1 accumulated total is close to other instruments throughout the year and both Thies LPMs have similar r^2 values to other instruments in Fig. 6, the largest difference being for the PWD21 with a difference of 0.047. However, the gradient *m* between the Thies LPMs is 1.05 which suggests a 5% difference. Here, the limitation of the least squares regression is highlighted; examination of the scatterplot between the Thies LPMs (not shown) reveals several cases where the Thies2 reported 5-min precipitation as much as half that of the Thies1, skewing the gradient of the best fit. The scatterplot shows that the overdiagnosis of PR (at 5-min intervals) in Thies2 compared to Thies1 generally occurs across all PRs and is likely a systematic bias with the laser to diode occultation technique. Lanza and Stagi (2009) noted an overestimation in precipitation from optical disdrometers compared with a reference gauge and Lanzinger (2006) quantified a 5%–20% overestimation from optical disdrometers. Frasson et al. (2011) suggests that the Thies LPM is sensitive to precipitation particles outside of the manufacturer-quoted laser beam area, causing an overcount of particles for the specified area when compared with a TBR gauge. Large differences between Thies LPMs were also noted in Frasson et al. (2011).

A phenomenon observed in Fig. 8 with the JWD is the underestimation of small and large drops in heavy PRs, also observed in Fig. 6 because the gradient of the JWD versus RAL, PWS100, and Thies LPMs all show an underestimation by the JWD. The total PD over 12 months is near the mean of all instruments (bias of -4.3 mm), but the DSD results indicate that the JWD underdetects PR, so the PD value should be lower than truth. The PWS100 has a similar 12 month PD (-3.0 mm from average of all instruments) which is likely to be an underestimate and r^2 values are around 0.9 (if the lower TBR and PWD21 are dismissed). These values are the lowest of the RAL, JWD, and Thies LPMs and the PWS100 also has a lower gradient than the Thies1, JWD and RAL for 5-min intervals seen in Fig. 6, further supporting the conclusion that the PWS100 is underdiagnosing PR and PD.

b. Drop size distribution

The DSD comparisons are split up by PR and accumulation time. There are few differences between the time accumulations showing that 5 min is sufficient for a representative DSD shape. The Thies LPMs and PWS100 measure some drops larger than the JWD can measure (>5.0 mm). It has been shown that the maximum stable diameter of a raindrop in stationary air before breakup occurs is $\sim 6 \text{ mm}$ (Villermaux and Bossa 2009; Marshall and Palmer 1948). This suggests that the largest drops measured by the Thies LPMs and PWS100 are realistic. Therefore the JWD appears to be limited and from the results in Fig. 8 is also underdiagnosing the number of large droplets. This is important for heavy PRs which can cause flash flooding because larger drops contribute more volume of water to the rain rate than smaller drops. Capturing larger drops in disdrometer measurements is also important for comparison with, or calibration of, precipitation radars. Smith (2016) found that the largest 6% of drops in a simulated DSD contributed toward 85% of the value of the reflectivity factor Z_h . The JWD should therefore not be used for applications where high PRs need to be well captured.

The JWD also underdiagnoses the small drops at high PRs $(>2.0 \text{ mm h}^{-1})$, which may be due to the dead-time effect, not accounted for here, where large drops cause oscillations on the JWD plate which take a nonnegligible time to dampen. In this time, motion from small drops is undetectable. However, the impact of the dead-time correction has been questioned in

the literature (Ulbrich and Atlas 2007). The JWD agrees with the Thies LPM for lower PRs ($<2.0 \text{ mm h}^{-1}$) so high sensitivity to drizzle and weak PRs makes the JWD suitable for those applications.

Alternatively, the Thies LPMs have a steeper DSD for small drops during high PRs ($>5.0 \text{ mm h}^{-1}$), which is unrealistic; the Thies LPM housing has been suspected of splashing precipitation into the measuring beam (Pickering et al. 2019a), which would increase the number of small drops seen in Fig. 8. Sensitivity at weak PRs and drizzle is also high in the Thies LPMs, although this does occasionally result in some anomalous measurements of insects or debris. The PWS100 underdiagnosing small drops could be a critical failure for automated precipitation-detection applications. In the PWS100 manual it states "proportion of particles detected will fall off significantly below about 0.5 mm diameter," not below 1.0 mm as seen here and should therefore not be used for high-sensitivity purposes. All instruments have suspect drop count values in the small drop diameter sizes for higher PRs ($>5.0 \text{ mm h}^{-1}$), so it is not clear that any of the instruments tested capture the full DSD correctly for high PRs ($>5.0 \text{ mm h}^{-1}$).

c. Drop velocity distribution

As with DSD, the difference in distribution shape is consistent for a given PR across multiple data collection time samples, indicating that a 5-min sample is sufficiently representative. Pickering et al. (2019a) note that there are often cases where the Thies LPM measures a portion of the smallest particles (0.125 mm) at the highest velocity bin (>10 m s⁻¹) which is incompatible with the empirically derived relationship between drop diameter and terminal fall velocity by Gunn and Kinzer (1949). The cause of these small diameter and high velocity drops is unknown but the behavior exists in both Thies LPMs used in this study.

The Gunn–Kinzer curve is a widely used relationship between drop size and terminal velocity. In the 5.0 mm h⁻¹ cases (5, 12, 19), the maximum drop diameters of around 4.0–4.5 mm for both the PWS100 and Thies LPMs should have a terminal velocity of 8.7–9.0 m s⁻¹ according to the Gunn–Kinzer curve. In Fig. 9, however, the PWS100 has a maximum velocity of 4.5 m s⁻¹ and the Thies LPMs of 10.5 m s⁻¹, which is higher than Gunn–Kinzer but closer than the PWS100. This trend continues across all DVDs. The Thies LPMs are close to the values expected from Gunn–Kinzer, but the PWS100 DVDs appear to be linearly stretched to approximately half the values expected.

Since the PWS100 underdetects small drops, the DVD will also be affected. Despite this the counts of low velocity drops are the same or higher than the Thies LPMs in 13 out of the 18 cases, which supports the observation that the Thies LPMs incorrectly measure a broad range of velocities for small drops, decreasing the true low velocity ($<1.0 \text{ m s}^{-1}$) count. Since the Thies LPMs use only DSD and DVD for PT, these anomalies in observation will lead to incorrect diagnoses of PT when particle sizes are small.

d. Precipitation type

Both Thies LPMs have high drizzle counts, higher than rain which does not match the site climatology (Ventouras et al.

2006). This signals a systematic design error which may be due to the sensitivity of the instrument to any small particles in the atmosphere like insects or debris. Both the PWS100 and PWD21 counted fewer drizzle cases than the Thies LPMs. In section 4a it was shown that the PWS100 underestimates small drops; drizzle is defined as drops < 0.5 mm in diameter, so the result is consistent with earlier findings. Similarly, the PWD21 does not output DSDs, so cannot explicitly distinguish drizzle. Ice and wet ice are the most disagreed classes simply because the climatology of the study location has rare occurrences of pristine ice crystals. Hail is also rare in the United Kingdom (Hand and Cappelluti 2011) and is therefore unreasonable to evaluate over a 12-month period.

The results for PT show weak agreement across the instruments, which is consistent with the literature for similar instruments. Bloemink (2005) showed that for a simple PT classifier (none, liquid, mixed, solid), there was little difference between the Thies LPM and Vaisala FD12P (akin to the PWD21 with a visiometer, thermometer, and capacitive plate) when each were compared with a trained observer. They noted a particular weakness for the instruments to detect mixed phase precipitation, also seen here.

There is little difference in the HSS for different time periods, which demonstrates that the amalgamation algorithm presented here is successful; if there were large increases or decreases in the HSS over time, then the algorithm is introducing a bias. The PT scheme and amalgamation algorithm should be applied to future instruments and to research analysis, so that there is consistency between instrument manufacturers in this field. The explicit PT scheme will also require instruments to be explicit, which the current PW code does not require since it is ambiguous.

e. Study limitations

Several limitations exist with the current study which must be noted and considered. First there is no single instrument which can be identified as the truth, and therefore the results are limited to comparisons between two instruments or an average of all instruments (referred to as a "composite working reference" in intercomparison studies). By contrast, Lanza and Vuerich (2009) use a composite working reference of four gauges (two TBRs and two weighing gauges) identified from laboratory comparisons. The effect of wind on precipitation measurements has long been known (Heberden 1769; Jevons 1862; Koschmieder 1934) but here no considerations are made to isolate precipitation events by wind speed thresholds. This study was based on existing instruments on-site and their preinstalled locations which could not be changed due to the need to maintain consistent long-term measurements at CAO. Only the RAL and TBR could receive a validation of the manufacturer calibration on-site. The PWD21 and PWS100 instruments are on the roof of a small building, which is suboptimal due to turbulence generated by the building-however, these instruments are raised above the building as shown in Fig. 1 which reduces the turbulence within the instrument sampling volume. Finally, while 12 months of data covers seasonal changes, the study period cannot be considered a true climatological sample since 30 years of observations would be

required. Note that other intercomparison studies also use a data collection period on the order of 1 year (Lanza and Vuerich 2009; Kochendorfer et al. 2017). The study is also conducted in a single location and therefore the results only apply to locations with similar precipitation climatologies as that of CAO (see section 2b) and up to PRs of 50 mm h^{-1} .

6. Summary

We examined six techniques for measuring precipitation to inform the future usage of their observations, sensor deployments, and to provide a benchmark for new sensor development. Comparison techniques from the literature were modified and improved upon to further explore the consistency and reliability of the observations from sensors with different measurement techniques. Rainfall depth and rate were examined using linear regression. The regressions of each instrument were cross compared to reveal consistent poor agreement and other patterns. A novel technique for comparing drop size and velocity distributions was described and employed. By selecting cases of stable rainfall rate, the drop distributions are kept consistent for the comparisons. A new precipitation type lookup table was created to convert and group the widely used PW code format into more explicit classes which enables a standardized comparison with other instruments to be performed. An amalgamation algorithm was also introduced which merges multiple 1-min precipitation type observations into a single code. The new PT codes and amalgamation algorithm were then used to compare four instruments over three time scales. Though the comparison of the instruments showed poor instrument agreement, the results demonstrate the successful application of the PT scheme and amalgamation algorithm.

For robust measurements, observing sites should employ multiple sensor designs. No single sensor in this study could satisfy all user applications. The Thies LPM makes reasonable observations for all variables, although the PT data is difficult to verify without a human observer and the PD between two Thies LPMs over 12 months had a 26% difference, signaling poor manufacturing calibration consistency. The PWS100 reported unrealistically low velocity measurements, fewer PTs than the Thies LPM and struggled to observe drizzle. The PWD21 performed poorly for every variable; PR and PD showed large negative biases of around 50% compared with the average of all instruments in the study, and PTs were narrowly reported (few classes). The JWD and PWS100 12-month PD were close to the mean PD but the DSD from the JWD underdiagnosed small and large drops, especially in higher PRs. The JWD and RAL had reasonable PR r^2 values with a fast response time. The RAL 12-month PD was 14% higher than the average for all instruments. The TBR has a slight negative PD bias compared with the average for all instruments over 12 months with 7.2% below the average. The low PD resolution of 0.2 mm makes the TBR unsuitable for response times less than 5 min for typical PRs observed at CAO.

A clear outcome of this study is that observations of PT, while useful and in growing demand, are poor because there is only moderate agreement between instruments with mixed precipitation disagreed upon the most. Other studies still use human observers as the most trusted PT which shows that there is a need for improved precipitation type sensors. Overall, the sensor design choice should change based on the user requirements, and this study serves as a reference for such a decision to be made.

Acknowledgments. This work is funded by a NERC Industrial CASE studentship (NE/N008359/1). Publication fees were funded by the University of Leeds Library through the UKRI Open Access grant. Ryan R. Neely III, Judith Jeffery, David Dufton, and Maryna Lukach were funded through the NCAS underpinning science programme through UKRI. Thanks goes to the National Centre for Atmospheric Science and Atmospheric Measurement and Observation Facility (AMOF) for site support at CAO and Joshua Hampton for classifying precipitation mode from NXPol radar observations which was not included in the final manuscript and for also proofreading an early draft.

Data availability statement. The PWD21, RAL, JWD, and TBR data used in this study are available from Wrench (2003a,b,c,d). The PWS data are available from Agnew (2013), and the Thies LPM data are available from Pickering et al. (2019b).

REFERENCES

- Adolf Thies GmbH, 2011: Laser Precipitation Monitor Instruction for use. Adolf Thies GmbH & Co. KG Tech. Rep., 66 pp.
- Agnew, J., 2013: Chilbolton Facility for Atmospheric and Radio Research (CFARR) Campbell Scientific PWS100 present weather sensor data. NCAS British Atmospheric Data Centre, accessed 1 November 2019, https://catalogue.ceda.ac.uk/uuid/ e490cd13d86d832bd2d62f1650d7b265.
- —, 2014: Final report on the operation of a Campbell Scientific CS135 ceilometer at Chilbolton Observatory. Science and Technology Facilities Council, 30 pp., https://s.campbellsci.com/ documents/us/miscellaneous/cs135_Chilbolton_final_report.pdf.
- Al-Sakka, H., A. A. Boumahmoud, B. Fradon, S. J. Frasier, and P. Tabary, 2013: A new fuzzy logic hydrometeor classification scheme applied to the French X-, C-, and S-band polarimetric radars. J. Appl. Meteor. Climatol., 52, 2328–2344, https:// doi.org/10.1175/JAMC-D-12-0236.1.
- Atlas, D., and F. H. Ludlam, 1961: Multi-wavelength radar reflectivity of hailstorms. *Quart. J. Roy. Meteor. Soc.*, 87, 523–534, https://doi.org/10.1002/qj.49708737407.
- Battan, L. J., 1973: Radar Observation of the Atmosphere. University of Chicago Press, 324 pp.
- Bentley, W. A., 1904: Studies of raindrops and raindrop phenomena. *Mon. Wea. Rev.*, **32**, 450–456, https://doi.org/10.1175/ 1520-0493(1904)32<450:SORARP>2.0.CO;2.
- Bloemink, H., 2005: Precipitation type from the Thies disdrometer. WMO Tech. Conf. on Meteorological and Environmental Instruments and Methods of Observation (TECO-2005), Bucharest, Romania, World Meteorological Organization, 1–7, http://www.wmo.int/pages/prog/www/IMOP/publications/ IOM-82-TECO_2005/Papers/3(11)_Netherlands_4_Bloemink.pdf.
- Campbell Scientific, 2012: PWS100 present weather sensor user guide. Tech. Rep., 122 pp.
- Chandrasekar, V., R. Meneghini, and I. Zawadzki, 2003: Global and local precipitation measurements by radar. *Radar and Atmospheric Science: A Collection of Essays in Honor of David*

Atlas, R. M. Wakimoto and R. Srivastava, Eds., Amer. Meteor. Soc., 215–236, https://doi.org/10.1007/978-1-878220-36-3_9.

- Distromet Ltd, 2002: Disdrometer RD-80 Instruction Manual. Tech. Rep., 18 pp.
- —, 2012: Disdrometer RD-80 Operating Instructions. Tech. Rep., 23 pp.
- Eisenberg, D., S. Laustsen, and J. Stege, 2018: Wind turbine blade coating leading edge rain erosion model: Development and validation. *Wind Energy*, **21**, 942–951, https://doi.org/10.1002/ we.2200.
- Erkal, A., D. D'Ayala, and L. Sequeira, 2012: Assessment of winddriven rain impact, related surface erosion and surface strength reduction of historic building materials. *Build. Environ.*, 57, 336–348, https://doi.org/10.1016/j.buildenv.2012.05.004.
- Feingold, G., Z. Levin, and S. Tzivion, 1991: The evolution of raindrop spectra. Part III: Downdraft generation in an axisymmetrical rainshaft model. J. Atmos. Sci., 48, 315–330, https://doi.org/10.1175/1520-0469(1991)048<0315:TEORSP> 2.0.CO;2.
- Frasson, R. P. M., L. K. da Cunha, and W. F. Krajewski, 2011: Assessment of the Thies optical disdrometer performance. *Atmos. Res.*, **101**, 237–255, https://doi.org/10.1016/j.atmosres. 2011.02.014.
- Fujiyoshi, Y., T. Endoh, T. Yamada, K. Tsubokj, Y. Tachibana, and G. Wakahama, 1990: Determination of a *Z*–*R* relationship for snowfall using a radar and high sensitivity snow gauges. *J. Appl. Meteor.*, **29**, 147–152, https://doi.org/10.1175/1520-0450(1990) 029<0147:DOARFS>2.0.CO;2.
- Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting*, 13, 377–395, https://doi.org/10.1175/1520-0434(1998) 013<0377:TWRA>2.0.CO;2.
- Gunn, R., and G. D. Kinzer, 1949: The terminal velocity of fall for water droplets in stagnant air. J. Meteor., 6, 243–248, https:// doi.org/10.1175/1520-0469(1949)006<0243:TTVOFF>2.0.CO;2.
- Hand, W. H., and G. Cappelluti, 2011: A global hail climatology using the UK Met Office convection diagnosis procedure (CDP) and model analyses. *Meteor. Appl.*, 18, 446–458, https:// doi.org/10.1002/met.236.
- Harimaya, T., 1978: Observation of size distribution of graupel and snow flake. J. Fac. Sci., 5, 67–77.
- Harrison, D., S. J. Driscoll, and M. Kitchen, 2000: Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteor. Appl.*, 7, 135–144, https:// doi.org/10.1017/S1350482700001468.
- Heberden, W., 1769: XLVII. Of the different quantities of rain, which appear to fall, at different heights, over the same spot of ground. *Philos. Trans. Roy. Soc. London*, **59**, 359–362, https:// doi.org/10.1098/rstl.1769.0047.
- Heidke, P., 1926: Berechnung Des Erfolges Und Der Güte Der Windstärkevorhersagen Im Sturmwarnungsdienst. *Geogr. Ann.*, 8, 301–349, https://doi.org/10.1080/20014422.1926.11881138.
- Humphrey, M. D., J. D. Istok, J. Y. Lee, J. A. Hevesi, and A. L. Flint, 1997: A new method for automated dynamic calibration of tipping-bucket rain gauges. *J. Atmos. Oceanic Technol.*, 14, 1513–1519, https://doi.org/10.1175/1520-0426(1997)014<1513: ANMFAD>2.0.CO;2.
- Jameson, A. R., and A. B. Kostinski, 2001: What is a raindrop size distribution? *Bull. Amer. Meteor. Soc.*, 82, 1169–1177, https:// doi.org/10.1175/1520-0477(2001)082<1169:WIARSD>2.3.CO;2.
- Jevons, W. S., 1862: On the deficiency of rain in an elevated raingauge, as caused by wind. J. Franklin Inst., 73, 332, https:// doi.org/10.1016/0016-0032(62)90997-3.

- Jorgensen, D. P., and P. T. Willis, 1982: A Z–R relationship for hurricanes. J. Appl. Meteor. Climatol., 21, 356–366, https://doi.org/ 10.1175/1520-0450(1982)021<0356:AZRRFH>2.0.CO;2.
- Joss, J., and A. Waldvogel, 1967: A raindrops spectrograph with automatic analysis. *Pure Appl. Geophys.*, 68, 240–246, https:// doi.org/10.1007/BF00874898.
- —, and —, 1970: Raindrop size distributions and Doppler velocities. *14th Radar Meteorology Conf.*, Tucson, AZ, Amer. Meteor. Soc., 153–155.
- Keegan, M. H., D. H. Nash, and M. M. Stack, 2013: On erosion issues associated with the leading edge of wind turbine blades. J. Phys., 46D, 383001, https://doi.org/10.1088/0022-3727/46/38/383001.
- Kinnell, P. I., 1981: Rainfall intensity-kinetic energy relationships for soil loss prediction. Soil Sci. Soc. Amer. J., 45, 153–155, https://doi.org/10.2136/sssaj1981.03615995004500010033x.
- Kochendorfer, J., and Coauthors, 2017: Analysis of single-Altershielded and unshielded measurements of mixed and solid precipitation from WMO-SPICE. *Hydrol. Earth Syst. Sci.*, 21, 3525–3542, https://doi.org/10.5194/hess-21-3525-2017.
- Koschmieder, H., 1934: Methods and results of definite rain measurements. *Mon. Wea. Rev.*, **62**, 5–7, https://doi.org/10.1175/ 1520-0493(1934)62<5:MARODR>2.0.CO;2.
- Langleben, M. P., 1954: The terminal velocity of snowflakes. Quart. J. Roy. Meteor. Soc., 80, 174–181, https://doi.org/10.1002/ qj.49708034404.
- 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.
- Lanzinger, E., 2006: Rainfall amount and intensity measured by the Thies laser precipitation monitor. TECO-2006, WMO-IOM-94, WMO-TD 135, 9 pp., http://www.wmo.int/pages/prog/ www/IMOP/publications/IOM-94-TECO2006/3(3)_Lanzinger_ Germany.pdf.
- Laws, J. O., and D. A. Parsons, 1943: The relation of raindrop-size to intensity. *Eos, Trans. Amer. Geophys. Union*, 24, 452–460, https://doi.org/10.1029/TR024i002p00452.
- Locatelli, J. D., and P. V. Hobbs, 1974: Fall speeds and masses of solid precipitation particles. J. Geophys. Res., 79, 2185–2197, https://doi.org/10.1029/JC079i015p02185.
- 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.
- Lumb, F. E., 1963: Downward penetration of snow in relation to the intensity of precipitation. *Meteor. Mag.*, **92**, 1–14.
- Maeso, J., V. N. Bringi, S. Cruz-Pol, and V. Chandrasekar, 2005: DSD characterization and computations of expected reflectivity using data from a two-dimensional video disdrometer deployed in a tropical environment. *Int. Geoscience and Remote Sensing Symp.*, Seoul, South Korea, IEEE, 5073–5076, https://doi.org/ 10.1109/IGARSS.2005.1526820.
- Marshall, J. S., and W. M. Palmer, 1948: The distribution of raindrops with size. J. Meteor., 5, 165–166, https://doi.org/10.1175/ 1520-0469(1948)005<0165:TDORWS>2.0.CO;2.
- Met Office, 2016: Southern England: Climate. United Kingdom Meteorological Office, https://www.metoffice.gov.uk/climate/ uk/regional-climates/so.
- Norbury, J. R., and W. J. White, 1971: A rapid-response rain gauge. J. Phys., 4E, 601–602, https://doi.org/10.1088/0022-3735/4/8/013.

- Oppenheim, I., and D. Shinar, 2012: A context-sensitive model of driving behaviour and its implications for in-vehicle safety systems. *Cognition Technol. Work*, **14**, 261–281, https://doi.org/ 10.1007/s10111-011-0178-3.
- Park, H. S., A. V. Ryzhkov, D. S. Zrnić, and K.-E. Kim, 2009: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Wea. Forecasting*, 24, 730–748, https://doi.org/10.1175/2008WAF2222205.1.
- Pickering, B. S., R. R. Neely III, and D. Harrison, 2019a: The Disdrometer Verification Network (DiVeN): A UK network of laser precipitation instruments. *Atmos. Meas. Tech.*, **12**, 5845–5861, https://doi.org/10.5194/amt-12-5845-2019.
- —, —, and —, 2019b: The Disdrometer Verification Network (DiVeN): Particle diameter and fall velocity measurements from a network of Thies Laser Precipitation Monitors around the UK (2017-2019). Centre for Environmental Data Analysis, accessed 1 November 2019, https://doi.org/ 10.5285/602f11d9a2034dae9d0a7356f9aeaf45.
- Ralph, F. M., and Coauthors, 2005: Improving short-term (0-48 h) coolseason quantitative precipitation forecasting: Recommendations from a USWRP Workshop. *Bull. Amer. Meteor. Soc.*, 86, 1619–1632, https://doi.org/10.1175/BAMS-86-11-1619.
- Reeves, H. D., 2016: The uncertainty of precipitation-type observations and its effect on the validation of forecast precipitation type. *Wea. Forecasting*, **31**, 1961–1971, https://doi.org/10.1175/WAF-D-16-0068.1.
- Rosewell, C. J., 1986: Rainfall kinetic energy in eastern Australia. J. Appl. Meteor. Climatol., 25, 1695–1701, https://doi.org/ 10.1175/1520-0450(1986)025<1695:RKEIEA>2.0.CO;2.
- Rubel, F., and K. Brugger, 2009: 3-hourly quantitative precipitation estimation over central and northern Europe from rain gauge and radar data. *Atmos. Res.*, 94, 544–554, https://doi.org/ 10.1016/j.atmosres.2009.05.005.
- Saltikoff, E., and L. Nevvonen, 2011: First experiences of the operational use of a dual-polarisation weather radar in Finland. *Meteor.* Z., 20, 323–333, https://doi.org/10.1127/0941-2948/2011/0197.
- Sevruk, B., 1996: Adjustment of tipping-bucket precipitation gauge measurements. Atmos. Res., 42, 237–246, https://doi.org/10.1016/ 0169-8095(95)00066-6.
- Slot, H. M., E. R. Gelinck, C. Rentrop, and E. Van der Heide, 2015: Leading edge erosion of coated wind turbine blades: Review of coating life models. *Renewable Energy*, 80, 837–848, https:// doi.org/10.1016/j.renene.2015.02.036.
- Smith, P. L., 2016: Sampling issues in estimating radar variables from disdrometer data. J. Atmos. Oceanic Technol., 33, 2305–2313, https://doi.org/10.1175/JTECH-D-16-0040.1.
- Tang, W., and C. I. Davidson, 2004: Erosion of limestone building surfaces caused by wind-driven rain: 2. Numerical modeling. *Atmos. Environ.*, 38, 5601–5609, https://doi.org/10.1016/j.atmosenv.2004.06.014.
- Tapiador, F. J., and Coauthors, 2016: On the optimal measuring area for pointwise rainfall estimation: A dedicated experiment with

fourteen laser disdrometers. J. Hydrometeor., **18**, 753–760, https://doi.org/10.1175/JHM-D-16-0127.1.

- Thurai, M., and V. N. Bringi, 2018: Application of the generalized gamma model to represent the full rain drop size distribution spectra. J. Appl. Meteor. Climatol., 57, 1197–1210, https:// doi.org/10.1175/jamc-d-17-0235.1.
- Tokay, A., W. A. Petersen, P. Gatlin, and M. Wingo, 2013: Comparison of raindrop size distribution measurements by collocated disdrometers. *J. Atmos. Oceanic Technol.*, 30, 1672–1690, https://doi.org/10.1175/JTECH-D-12-00163.1.
- Ulbrich, C. W., 1983: Natural variations in the analytical form of the raindrop size distribution. J. Appl. Meteor. Climatol., 22, 1764–1775, https://doi.org/10.1175/1520-0450(1983)022<1764: NVITAF>2.0.CO;2.
- —, and D. Atlas, 2007: Microphysics of raindrop size spectra: Tropical continental and maritime storms. J. Appl. Meteor. Climatol., 46, 1777–1791, https://doi.org/10.1175/2007JAMC1649.1.
- Vaisala Oyj, 2001: Present weather detector PWD21 user's guide. Vaisala Oyj Tech. Rep., 115 pp.
- Ventouras, S., S. A. Callaghan, and C. L. Wrench, 2006: Longterm statistics of tropospheric attenuation from the Ka/U band ITALSAT satellite experiment in the United Kingdom. *Radio Sci.*, **41**, RS2007, https://doi.org/10.1029/ 2005RS003252.
- Villermaux, E., and B. Bossa, 2009: Single-drop fragmentation determines size distribution of raindrops. *Nat. Phys.*, 5, 697–702, https://doi.org/10.1038/nphys1340.
- Wexler, R., 1955: The Melting Layer. Vol. 3, Meteorological Radar Studies, Harvard University, 29 pp.
- Wiesner, J., 1895: Beiträge zur Kenntniss des tropischen Regens. Ber. Akad. Wiss. Wein, 104, 1397–1434.
- WMO, 1988: Manual on Codes. WMO Publ. 306, 203 pp.
- —, 2017: Guide to meteorological instruments and methods of observation. Tech. Rep. WMO 8, 1166 pp.
- Wrench, C. L., 2003a: Chilbolton Facility for Atmospheric and Radio Research (CFARR) Vaisala PWD21 Present Weather Detector data, Chilbolton Site. NCAS British Atmospheric Data Centre, accessed 1 November 2019, https://catalogue.ceda.ac.uk/uuid/ aac5f8246987ea43a68e3396b530d23e.
- —, 2003b: Chilbolton Facility for Atmospheric and Radio Research (CFARR) drop counting raingauge data, Chilbolton Site. NCAS British Atmospheric Data Centre, accessed 1 November 2019, https:// catalogue.ceda.ac.uk/uuid/aac5f8246987ea43a68e3396b530d23e.
- —, 2003c: Chilbolton Facility for Atmospheric and Radio Research (CFARR) disdrometer data, Chilbolton Site. NCAS British Atmospheric Data Centre, accessed 1 November 2019, https:// catalogue.ceda.ac.uk/uuid/aac5f8246987ea43a68e3396b530d23e.
- —, 2003d: Chilbolton Facility for Atmospheric and Radio Research (CFARR) multiple raingauges data, Chilbolton Site. NCAS British Atmospheric Data Centre, accessed 1 November 2019, https:// catalogue.ceda.ac.uk/uuid/aac5f8246987ea43a68e3396b530d23e.