HAL
open science

How Well Are We Measuring Snow Post-SPICE?

John Kochendorfer, Michael Earle, Roy Rasmussen, Craig Smith, Daqing Yang, Samuel Morin, Eva Mekis, Samuel Buisan, Yves-Alain Roulet, Scott Landolt, et al.

To cite this version:

John Kochendorfer, Michael Earle, Roy Rasmussen, Craig Smith, Daqing Yang, et al.. How Well Are We Measuring Snow Post-SPICE?. Bulletin of the American Meteorological Society, 2022, 103 (2), pp.E370-E388. 10.1175/BAMS-D-20-0228.1 . hal-03971218

HAL Id: hal-03971218

https://hal.science/hal-03971218v1

Submitted on 6 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BAMS Article

How Well Are We Measuring Snow Post-SPICE?

John Kochendorfer, Michael Earle, Roy Rasmussen, Craig Smith, Daqing Yang, Samuel Morin, Eva Mekis, Samuel Buisan, Yves-Alain Roulet, Scott Landolt, Mareile Wolff, Jeffery Hoover, Julie M. Thériault, Gyuwon Lee, Bruce Baker, Rodica Nitu, Luca Lanza, Matteo Colli, and Tilden Meyers

ABSTRACT: Accurate snowfall measurements are necessary for meteorology, hydrology, and climate research. Typical uses include creating and calibrating gridded precipitation products, the verification of model simulations, driving hydrologic models, input into aircraft deicing processes, and estimating streamflow runoff in the spring. These applications are significantly impacted by errors in solid precipitation measurements. The recent WMO Solid Precipitation Intercomparison Experiment (SPICE) attempted to characterize and reduce some of the measurement uncertainties through an international effort involving 15 countries utilizing over 20 types and models of precipitation gauges from various manufacturers. Key results from WMO-SPICE are presented herein. Recent work and future research opportunities that build on the results of WMO-SPICE are also highlighted.

KEYWORDS: Precipitation; Snowfall; Freezing precipitation; Hydrology; Instrumentation/sensors; Measurements

https://doi.org/10.1175/BAMS-D-20-0228.1
Corresponding author: John Kochendorfer, john.kochendorfer@noaa.gov
In final form 4 August 2021
©2022 American Meteorological Society
For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

AFFILIATIONS: Kochendorfer and Baker—Atmospheric Turbulence and Diffusion Division, NOAA/Air Resources Laboratory, Oak Ridge, Tennessee; **Earle**—Meteorological Service of Canada, Environment and Climate Change Canada, Dartmouth, Nova Scotia, Canada; **Rasmussen and Landolt**—National Center for Atmospheric Research, Boulder, Colorado; **Smith**—Climate Research Division, Environment and Climate Change Canada, Saskatoon, Saskatchewan, Canada; **Yang**—Watershed Hydrology and Ecology Research Division, Environment and Climate Change Canada, Victoria, British Columbia, Canada; **Morin**—Univ. Grenoble Alpes, Université de Toulouse, Météo-France, CNRS, CNRM, Centre d'Études de la Neige, Grenoble, France; **Mekis**—Climate Research Division, Environment and Climate Change Canada, Toronto, Ontario, Canada; **Buisan**—Aragon Regional Office, Agencia Estatal de Meteorología, Zaragoza, Spain; **Roulet**—MeteoSwiss, Payerne, Switzerland; **Wolff**—Norwegian Meteorological Institute, Oslo, and Norwegian University of Life Sciences, Ås, Norway; **Hoover, Nitu, and Meyers**—Meteorological Service of Canada, Environment and Climate Change Canada, Toronto, Ontario, Canada; **Thériault**—ESCER Center, Department of Earth and Atmospheric Sciences, University of Quebec at Montreal, Montreal, Quebec, Canada; **Lee**—Kyungpook National University, Daegu, South Korea; **Lanza**—Department of Civil, Chemical and Environmental Engineering, University of Genoa, and WMO/CIMO Lead Centre "B. Castelli" on Precipitation Intensity, Genoa, Italy; **Colli**—Artys Srl, Genoa, Italy

Snowfall is one of the most difficult meteorological variables to measure using automated sensors. These measurements are critically important, as the presence, quantity, and distribution of snow affects regional and global hydrology and climate. Snowfall impacts the albedo of the Earth's surface, ecosystem function, permafrost characteristics, and the mass balance of glaciers, sea ice, and ice sheets. Snowfall also contributes to weather-related hazards, including avalanches and floods, and creates dangerous conditions for both air and land transport. As Earth's average atmospheric temperature increases, alpine and northern regions are experiencing remarkable decreases in snowfall and increases in rainfall (Li et al. 2020; Trenberth 2011; Trenberth et al. 2003; IPCC 2019; Mote et al. 2018), and significant reductions in snow cover duration and extent (Derksen and Brown 2012; Derksen et al. 2019). All of these changes are projected to continue in a warming climate. To quantify these trends and better adapt to their effects, the accurate measurement of snow is critical to predict its variations at local, regional, and global scales.

In spite of their importance, snow measurements are still subject to significant uncertainties and biases, particularly in cold and windy conditions (Goodison et al. 1998; Gugerli et al. 2021; Yang et al. 2005; Rasmussen et al. 2012; Buisán et al. 2020; Daly et al. 2017; Langousis et al. 2018; Milewska et al. 2019; Nitu et al. 2018; Pan et al. 2016, 2019; Yao et al. 2018, Zhang et al. 2020). Despite recent advancements in sensor technology, measurement techniques, and communications, snow cover measurements, such as snow depth and snow water equivalent (SWE), are still primarily recorded manually, and require specialized equipment and well-trained personnel. Snow cover on glaciers, ice sheets, sea ice, and land is notoriously heterogeneous, making it difficult to obtain representative measurements over a given area (Picard et al. 2016). Measurement of the liquid water equivalent of precipitation falling as snow, or other forms of solid precipitation, typically requires heated precipitation gauges to prevent full or partial blockage (capping) of the gauge inlet by snow and ice. In addition, precipitation gauges can significantly underestimate the true amount of solid precipitation, primarily due to wind effects. For these reasons, the improvement of snow cover and solid precipitation measurements is an important subject of climatological and hydrological research in cold regions.

Measurements of solid precipitation and snow cover are especially important in high-altitude regions with complex terrain, as much of the world's water used for irrigation and

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

human consumption in semiarid regions comes from runoff generated by alpine snow cover and glaciers (Schaffer et al. 2019; Orphanopoulos et al. 2013; Arnell 1999). A recent analysis by Lundquist et al. (2019) suggests that gridded gauge-based snowfall estimates in complex terrain are less accurate than snowfall estimates produced using high-resolution models. Likewise, a comparison of Snowpack Telemetry (SNOTEL) and snow course measurements demonstrated that small differences in point measurements of SWE can be large when extrapolated to the basin scale (Dressler et al. 2006). This makes it imperative to improve the spatial coverage and quality of solid precipitation and snow cover measurements in these regions.

Past intercomparison experiments, such as the WMO Solid Precipitation Measurement Intercomparison (Goodison et al. 1998), were used to develop standards for manual solid precipitation measurements. These results were applied to improve precipitation records spanning many countries and regions (Metcalfe and Goodison 1993; Yang 1999; Yang et al. 1998, 1999, 2005; Yang and Ohata 2001; Zhang et al. 2004; Ye et al. 2004; Adam and Lettenmaier 2003). This work significantly improved our understanding of cold region climate and hydrology, including regional climate change (Ding et al. 2007), watershed water balance (Ye et al. 2012), large-scale land surface modeling of the arctic hydrological system (Tian et al. 2007), and precipitation distribution across national borders (Scaff et al. 2015).

Automated precipitation gauges and snow cover sensors have since proliferated in both research and operational measurement networks. In addition, non-catchment precipitation measurements, which involve light scattering, microwave backscatter, or mass and heat transfer, have also come into use. The accuracy of non-catchment solid precipitation measurements, however, needs to be evaluated carefully, because such gauges cannot be calibrated using traditional techniques. Due to the increased availability of all of these measurements, many new uses for automated snow cover and solid precipitation observations have emerged (e.g., climate change monitoring, nowcasting, water supply forecasting and management, aircraft deicing, complex terrain, avalanche warnings). In addition, modern data management and data assimilation techniques are advancing the requirements for improved uncertainty metrics and metadata for snowfall observations.

In recognition of these uncertainty and metadata requirements and the need to standardize automated snow cover and solid precipitation measurements, the WMO Commission for Instruments and Methods of Observation (CIMO) initiated the Solid Precipitation Intercomparison Experiment (WMO-SPICE, hereafter SPICE) in the spring of 2010. SPICE included the assessment of snow cover sensors (including snow depth and SWE sensors), present weather detectors, laser disdrometers, and precipitation gauges designed to measure solid precipitation in real time, in different configurations. These sensors were provided by manufacturers and project participants, and were tested at 20 intercomparison sites across the world (Fig. 1). SPICE also defined and characterized an automated field reference for the measurement of solid precipitation, referred to as the Double Fence Automated Reference (DFAR, shown in Fig. 2b).

The final SPICE report was published in 2018 (Nitu et al. 2018). The report contains assessments of instrument performance and best practices for solid precipitation and snow cover measurements, and was intended to aid in the transition from manual to automated measurements. In addition, recommendations were made for the measurement of solid precipitation and snow cover, including different field configurations, possible improvements, and their applications in various climatic and snowfall regimes.

SPICE also helped create a community of practice among leading solid precipitation and snow cover measurement experts, as well as an extensive, quality-controlled dataset that is now freely available (https://ral.ucar.edu/projects/SPICE/). Many new research opportunities arose from SPICE. For example, a special issue dedicated to SPICE and related work was jointly organized between the journals of *Atmospheric Measurement Techniques*, *Earth System Science Data*, *Hydrology and Earth System Sciences*, and *The Cryosphere*

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

(https://hess.copernicus.org/articles/special_issue400_78.html). This issue includes work performed during and immediately after SPICE. Related work has continued post-SPICE, and is ongoing. This article, succeeding Rasmussen et al. (2012), will summarize the key results of the SPICE project. It will also document the limitations of the SPICE results, progress made since SPICE ended, and future research directions.

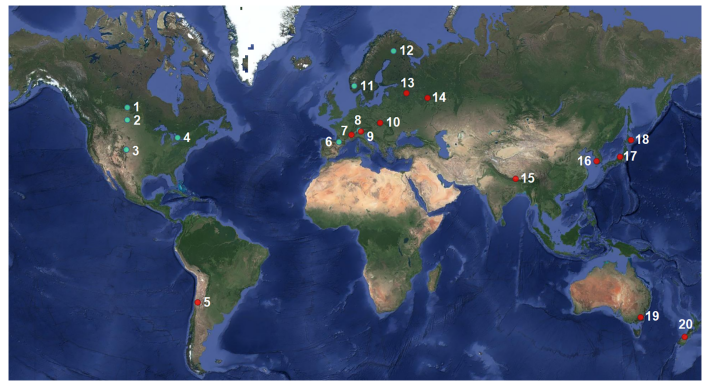

Fig. 1. Map of SPICE sites: Caribou Creek, Canada (1); Bratt's Lake, Canada (2); Marshall, CO, United States (3); CARE, Canada (4); Tapado AWS, Chile (5); Formigal, Spain (6); Col de Porte, France (7); Weissfluhjoch, Switzerland (8); Forni Glacier, Italy (9); Hala Gasienicowa, Poland (10); Haukeliseter, Norway (11); Sodankylä, Finland (12); Valdai, Russia (13); Voljskaya, Russia (14); Pyramid, Nepal (15); Gochang, Korea (16); Joetsu, Japan (17); Rikubetu, Japan (18); Guthega Dam, Australia (19); and Mueller Hut, New Zealand (20). Sites shown with blue dots have a Double Fence Automated Reference (DFAR) solid precipitation system.

Overview of SPICE

SPICE included field measurements from 20 test sites in 16 countries during the Northern Hemisphere winter seasons of 2013/14 and 2014/15, and the Southern Hemisphere winter seasons of 2014 and 2015. Twenty-seven different types of sensors, with over 270 individual sensors in total, were evaluated over the intercomparison period. These assessments provided a broad perspective on the capabilities of automated instruments for the measurement of solid precipitation and snow cover in different climates. Standardized data derivation approaches were developed and used for sensor evaluations. Datasets including only confirmed periods of precipitation were derived for use in the determination of transfer functions to adjust weighing gauge measurements for wind-induced undercatch relative to the reference gauge configuration.

Assessments of gauge performance were conducted for automated weighing gauges, heated tipping-bucket gauges, non-catchment instruments, and snow depth and snow water equivalent sensors. The influence of instrument configuration (e.g., wind shielding, mounting infrastructure, heating) and environmental conditions was also investigated for all sensors included in the intercomparison. Although the physical processes of solid precipitation are linked to snow cover, the assessments of snowfall (solid precipitation) and snow cover (snow depth and snow water equivalent) instrumentation were performed independently in SPICE using

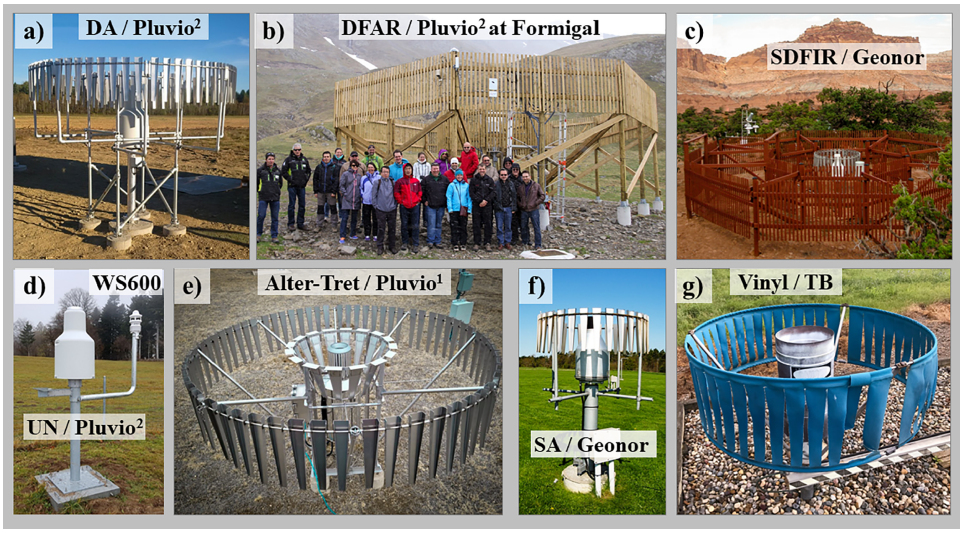

Fig. 2. Photos of operational precipitation gauge configurations: (a) a Pluvio² precipitation gauge in a double Alter shield, (b) a DFAR, (c) a Geonor T200B gauge within a SDFIR (Small DFIR) shield, (d) an unshielded Pluvio² collocated with an automated weather station, (e) a Pluvio¹ within a Tretyakov and Alter shield, (f) a single-Alter shielded Geonor gauge, and (g) an ASOS tipping-bucket gauge with a single vinyl shield.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

separate reference measurements. Intercomparison of solid precipitation and snow cover measurements was not a primary objective of SPICE, although one measurement is often used as a proxy for the other (Sevruk 1983; Broxton et al. 2016; Wen et al. 2016; Goodison et al. 1998), and legacy SPICE datasets, including the reference gauge precipitation, SWE, snow depth, and ancillary measurements at various time intervals, could be useful for further technique development and assessment.

Field reference systems. The DFAR was used to obtain the primary reference measurements for precipitation in SPICE. This configuration consisted of a precipitation detector and an automated weighing gauge installed within a single-Alter shield and a concentric, wooden, double fence (Fig. 2b). This type of fence was used in the manual DFIR configuration, which was established as the primary field reference for the previous WMO intercomparison of solid precipitation measurements (Goodison et al. 1998). The precipitation detector—which was typically an optical disdrometer—was used to enable the identification of precipitating periods with a higher degree of confidence, especially during very light snowfall events.

Manual DFIR and bush-shielded gauges were recommended and used as the primary references for past solid precipitation measurement intercomparisons (Goodison et al. 1998; Yang et al. 1993; Yang 2014). To help establish an automated reference system, the SPICE project compared automated naturally shielded (bush) gauges to the DFAR at the Caribou Creek site in Canada. This work helped to validate the use of the DFAR as an automated field reference for the measurement of solid precipitation (Nitu et al. 2018).

The primary reference measurements for snow cover (snow depth and SWE) were daily manual snow stake observations (for snow depth) and snow course bulk density sampling (for SWE). At several sites, the snow stake observations were semiautomated using photography. At sites with multiple snow depth sensor types [e.g., Centre for Atmospheric Research Experiments (CARE)], the mean of all concurrent sensor observations was utilized as an automated high-frequency reference measurement.

Weighing gauges

Evaluation of primary weighing gauges tested in SPICE. Automated weighing gauges are catchment-type instruments, collecting and storing incident precipitation in a bucket within the gauge housing. The total liquid water equivalent of accumulated precipitation is determined from the weight of the bucket contents. To mitigate the influence of wind effects on gauge collection, which causes significant uncertainties in measurements of solid precipitation, weighing gauges are often installed inside wind shields comprising one or two concentric rings of metal slats. A modified design by Alter (1937) is commonly used, and when used in a single or double ring, is referred to as a single-Alter or double-Alter shield, respectively. Heavy wet snow in low wind conditions and freezing rain can also result in gauge measurement errors. To minimize these errors, heating can be applied at the gauge inlet to reduce the potential for snow capping or freezing rain to influence gauge collection. Additionally, weighing gauge function can be compromised if the bucket contents freeze, necessitating the use of antifreeze and also a layer of oil to prevent evaporation of the antifreeze and bucket contents.

Seven different weighing gauge models were tested in SPICE in different wind shield configurations. Gauge orifices were heated, where possible. Among the eight SPICE sites with a DFAR, there was significant site-to-site variability of solid precipitation undercatch for both unshielded and single-Alter shielded weighing precipitation gauges (Fig. 3). At windier and more exposed sites [e.g., Bratt's Lake (XBK) and Haukeliseter (HKL)], undercatch was more significant than at sheltered and less windy sites [e.g., Sodankylä (SOD)]. For weighing gauges, performance was driven largely by wind conditions, and was found to depend more on the shield configuration than the specific gauge model at a given site (Kochendorfer et al. 2018;

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

Nitu et al. 2018). A typical accumulation time series from the Marshall, Colorado, testbed demonstrates the importance of wind shielding (Fig. 4), with the DFIR-shielded precipitation gauge accumulating about twice as much precipitation as an unshielded gauge over several winter months. An example of the relative effectiveness of other available windshields is also shown in Fig. 4.

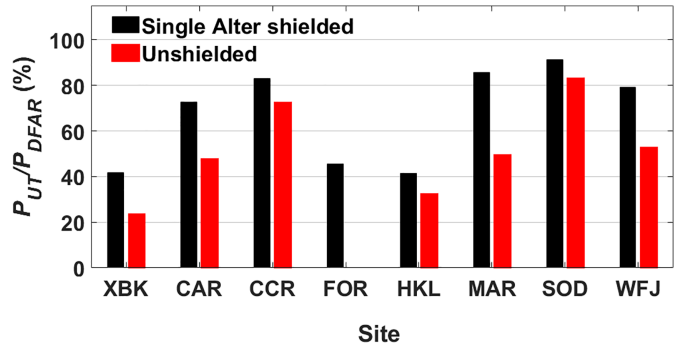

Fig. 3. Solid precipitation accumulated by the gauge under test (P_{UT}), divided by the accumulation of solid precipitation recorded by the DFAR (P_{DFAR}). All solid precipitation recorded during the 2015/16 and 2016/17 seasons is included, for all the SPICE sites that had a DFAR. These sites include Bratt's Lake (XBK, Canada), CARE (CAR, Canada), Caribou Creek (CCR, Canada), Formigal (FOR, Spain), Haukeliseter (HKL, Norway), Marshall, (MAR, United States), Sodankylä (SOD, Finland), and Weissfluhjoch (WFJ, Switzerland). These measurements are described in more detail in Smith et al. (2020).

Derivation of transfer functions for weighing precipitation gauges. Before SPICE measurements were available, and in preparation for the analysis of the SPICE weighing gauge measurements, new single-site transfer functions were derived that included both air temperature and wind speed (Wolff et al. 2015). By combining measurements from eight different SPICE sites, multisite transfer functions and their associated site-specific uncertainties and biases were similarly derived for unshielded and single-Alter shielded weighing gauge measurements (Kochendorfer et al. 2017). These transfer functions were tested using measurements from all of the different types of weighing gauges included in SPICE, and it was determined that the same transfer function could be used for different types of weighing precipitation gauges when the same type of shielding was used (Kochendorfer et al. 2018). However, significant uncertainty and site- or climate-specific biases remain after such transfer functions have been applied to the measurements. The magnitude of these unresolved errors decreases significantly when more effective shielding is used; transfer functions can be used to minimize biases in solid precipitation measurements, but the resultant measurements are still less accurate than well-shielded measurements (Kochendorfer et al. 2018; Smith et al. 2020).

Evaluation and limitations of multisite transfer functions. Following the SPICE intercomparison period, the eight test sites operating DFAR configurations continued to collect data for further intercomparisons and assessments. Smith et al. (2020) used measurements from two post-SPICE winter seasons to evaluate transfer function performance at each site. The performance of the transfer function varied significantly among the sites, supporting earlier evaluations by Kochendorfer et al. (2017). Smith et al. (2020) also evaluated the transfer function for solid precipitation only, and found that biases and errors were exacerbated when rain and mixed precipitation periods were removed from the dataset.

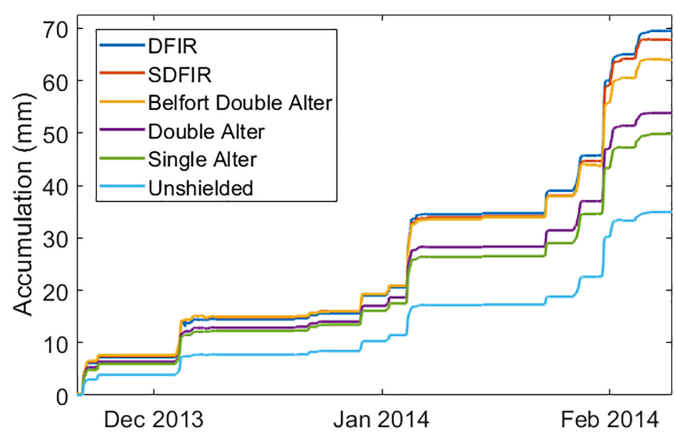

Fig. 4. Winter precipitation accumulations recorded at the Marshall test bed measured using Geonor T-200B gauges with different windshield configurations. The DFIR is the same large wooden shield with a single Alter used for the SPICE DFAR. The SDFIR (Small DFIR) is a 2/3 sized DFIR that also has a single-Alter shield within it. The Belfort double-Alter is a low porosity double-Alter shield manufactured by Belfort Instruments. The double and single Alter shields are produced by Geonor and other manufacturers.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

Pierre et al. (2019) conducted an evaluation of available transfer functions developed in different climate regimes (including the SPICE transfer function) for the adjustment of gauge measurements of solid precipitation at a single site (Forêt Montmorency, Quebec, Canada), and compared the results with the collocated manual DFIR. Site-specific transfer functions developed specifically for Forêt Montmorency, which has a southern boreal climate associated with large amounts of snow, showed lower bias and error values relative to other existing functions, underscoring the importance of climate and site conditions in transfer function performance.

Pierre et al. (2019) and Smith et al. (2020) showed substantial transfer function performance variation by site, mainly due to different wind characteristics. Smith et al. (2020) showed that the multisite SPICE transfer functions produced a significant underadjustment at cold and windy sites, while tending to overadjust at less windy sites. The results also suggested that transfer functions are useful, but should be applied with caution; local meteorology and climate are important for understanding undercatch and assigning appropriate transfer functions to all observation sites.

In addition to transfer function assessment, poorly shielded precipitation gauges tend to miss low-accumulation events in windy conditions. Smith et al. (2020) showed that collocated single-Alter shielded and unshielded weighing gauges at Bratt's Lake (Saskatchewan, Canada) reported a substantial number of 0.0 mm events when the DFAR reported low event accumulations. These missed events were characterized by a mean 10 m height wind speed greater than 7.5 m s^{-1}. For the unshielded and single-Alter shielded gauges, such missed events accounted for 62% and 38% of the solid precipitation reported by the DFAR, respectively (Fig. 5). These results highlight the importance of adequate shielding in high-wind conditions; regardless of how well a transfer function is calibrated to a specific site, it cannot be applied to an event that is not detected by the gauge.

Operational weighing gauge considerations.

Life cycle and change management. For operational applications, different weighing gauge and shield combinations are often used depending on climate and operational requirements. Network-wide changes can be necessitated by life cycle (discontinued sensors) and financial considerations. For example, since SPICE began, new precipitation gauge models [e.g., the Geonor T-200BM 1000 mm and T-200BMD 1500 mm, the Lufft OTT Pluvio²L, and the Lambrecht rain(e)H3] have become available. The use of different gauges and shielding can lead to significant challenges when trying to create homogenous precipitation records within and among regions or countries.

As an example, the Automated Surface Observing Systems (ASOS) and Automated Weather Observing Systems (AWOS) are the primary operational weather reporting networks in the United States. These systems have adopted various gauges and shields over the years, but have not yet established a standard across or within the networks. The ASOS network initially used a heated tipping-bucket with a vinyl shield that had slits cut in it (Fig. 2g). The National Weather Service adopted an OTT Pluvio gauge with a Tretyakov-style shield (referred to as the AWPAG, or Automated Weather Precipitation Accumulation Gauge) to install at their ASOS sites in the mid-2000s, while the Federal Aviation

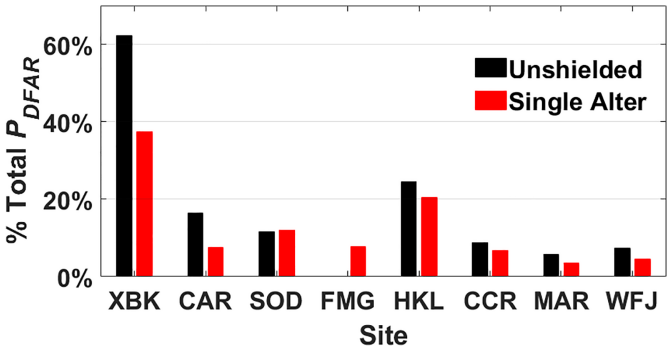

Fig. 5. Percent of total solid precipitation measured by the DFAR (P_{DFAR}) when the test gauge measured zero. The measurements are from Smith et al. (2020), and were recorded during the 2015/16 and 2016/17 winter seasons. Sites are identified in the Fig. 3 caption.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

Administration (FAA) opted to keep the heated tipping-bucket at their ASOS sites. From 2010 to 2011, the National Weather Service upgraded their ASOS OTT Pluvio gauges to include an Alter shield surrounding the Tretyakov shield (Fig. 2e). The AWOS network is maintained by both the FAA and private vendors, and includes an assortment of precipitation gauges, many of which are heated tipping buckets with no shields. This is part of the reason that the U.S. Climate Reference Network (USCRN) was created (Diamond et al. 2013), providing well-shielded reference-quality precipitation measurements (Fig. 2c) throughout the United States.

Canada provides another example of such operational issues. The network-wide automation of the manual observation program started there in the early 2000s, with the introduction of the Geonor T-200B (600 mm) and the higher capacity OTT Pluvio (1000 mm) gauges (Mekis et al. 2018). Beginning in 2008, primarily to decrease maintenance costs, the even larger capacity Pluvio^2 (1500 mm, then eventually the Pluvio^{2}L) gauge was often used at new installations. To modernize, in the Automatic Weather Observing System network, the single-Alter shielded Geonor T-200B gauges are currently being replaced with double-Alter (DA) Pluvio^{2}L gauges (Fig. 2a), with the new double-Alter shielded gauges exhibiting increased solid precipitation catch efficiencies in windy conditions. Many of these transitions have introduced significant spatial and temporal inhomogeneities, with the different configurations requiring homogenization via wind adjustment.

Operational transfer function considerations. The use of wind shields around precipitation gauges may also have some drawbacks. The installation of a shield can create an inhomogeneity in the time series, which can only be accounted for using parallel measurements recorded over several years. The additional structure around the gauge may also be prone to failure, requiring periodic inspection, maintenance, and replacement. For these reasons, Switzerland is investigating an alternative to systematic shielding for its operational network. For wind-exposed sites in mountainous regions, where the wind-induced error is expected to be large, an all-in-one sensor is installed on the same pole as the precipitation gauge, at orifice height, to measure wind speed and temperature (Fig. 2d). This additional information will then be used to apply transfer functions to the measurement ("Evaluation of primary weighing gauges tested in SPICE" and "Derivation of transfer functions for weighing precipitation gauges" sections), and create a new data parameter (corrected and uncorrected measurements will both be available). The impact of this new approach on existing products, like the combination of radar and ground-based measurements for nowcasting, will be evaluated.

Very high winds and blowing snow are also a source of significant operational measurement uncertainties in some areas. For wind speeds greater than 9 m s^{-1}, transfer function adjustments that increase snowfall amounts by up to 5 times can result in very large precipitation measurement errors. This is in part because prolonged windy conditions can generate blowing snow, which may be erroneously measured as new precipitation by a gauge (Nitu et al. 2018; Wagner et al. 2021; Yang and Ohata 2001). Potential solutions to this problem include a maximum wind speed for the transfer function to exclude blowing snow (Yang et al. 1998; Yang and Ohata 2001), or the imposition of a limit on the magnitude of snowfall amounts above which the adjustments should not be applied (MacDonald and Pomeroy 2007). Further research is needed to determine how to adjust precipitation during very windy conditions (Pan et al. 2016), and specifically to examine possible overadjustments and blowing snow impacts on such observations. Blowing snow is particularly problematic because it can also affect reference precipitation measurements, and because the presence or absence of blowing snow varies based on local conditions, snow morphology, and the condition of the snowpack.

Weighing gauge measurement recommendations. For solid precipitation measurement, at least one DFAR configuration should be installed and maintained by monitoring agencies to

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

characterize the measurement biases and uncertainty of national network gauges. Testbeds established at key locations, which include a DFAR and collocated sensors, can be used to preserve the continuity of long-term snow observations and to assess the effect of systematic transitions. The use of instrument testbeds and intercomparison protocols should be formalized to ensure proper intercomparisons with international references. This would reduce the risk of discontinuities associated with instrument transitions, and also facilitate more internationally and globally homogenous solid precipitation records.

Wind shielding is recommended for catchment-type precipitation gauges to reduce the effects of wind-induced undercatch. Shield selection can be based on the gauge exposure and climate. For example, double-ring shields are generally more effective than single-ring shields, but the former may be superfluous for gauges installed at sheltered sites with low wind speeds. Wind shields should be installed above the height of the maximum snow depth, to prevent the shield and gauge from being buried (e.g., Fig. 6e). In addition, shields should be installed separately from the structure used to support the gauge, as wind-induced vibration of shield components can influence gauge measurements.

For weighing gauges, antifreeze and oil should be used to prevent the freezing and evaporation of bucket contents, respectively. Heating is recommended to improve gauge response times for solid precipitation and to mitigate the influence of snow capping; however, heating may impact conditions above the gauge orifice and affect measurements. The heating configuration should be tailored to the specific application and environment, while also taking into consideration power availability.

Gauge capping is a problem that can occur in sheltered areas and in locations with heavy and wet snow. Webcams can be used to detect partial or complete capping events, in addition to other instruments such as precipitation detectors or snow depth sensors. Some gauges are designed to minimize snow accumulation on the gauge. This is done by decreasing the area of surfaces on which snow can accumulate, with an ideal gauge being cylindrical with no shoulders. In addition, gauge surfaces that accumulate snow can be heated when sufficient power is available. In locations prone to capping due to heavy snow and low wind, it may be better not to use a windshield, as snow can accumulate on shield components (e.g., Figs. 6h,i), and some exposure to the wind can actually help prevent snow from building up on the gauge.

Appropriate transfer functions should be used to mitigate the influence of wind-induced undercatch on solid precipitation measurements. Ancillary measurements of wind speed (at gauge height or at 10 m) and temperature must be recorded for the application of transfer functions, as well as all metadata related to the gauge and shield type and configuration of each measurement.

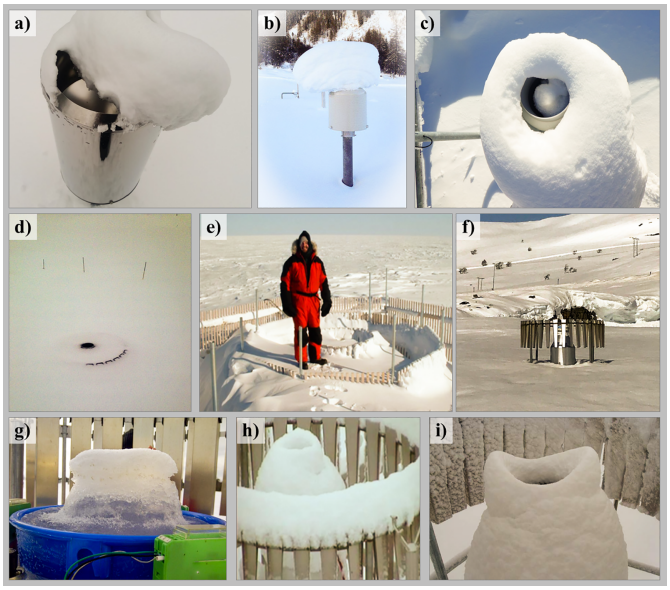

Fig. 6. Photos of common solid precipitation measurement issues, such as (a),(b),(c),(h),(i) fully or partially capped gauges; (d),(f) buried gauges; (e) shields that have accumulated drifting snow; (g) gauges with frozen bucket contents; and (h),(i) shields that have accumulated snow.

Heated tipping-bucket gauges

Tipping-bucket precipitation gauges are widely used in national observation networks (Nitu and Wong 2010), and are available with heating for all-season operation, including solid precipitation measurements. Tipping-bucket gauges are typically less expensive, smaller,

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

and require less expertise to install and operate than weighing gauges. Like weighing gauges, tipping-bucket gauge performance is strongly linked to wind speed (Buisán et al. 2017; Kochendorfer et al. 2020; Nitu et al. 2018). Heated tipping-bucket gauges are also significantly affected by evaporation and response delays, because solid precipitation must be heated and melted prior to its measurement (Buisán et al. 2017; Kochendorfer et al. 2020; Nitu et al. 2018). More research on how to minimize these drawbacks by optimizing heating is still needed. Tipping-bucket gauges are also less accurate than weighing gauges, and typically have a coarser measurement resolution (Nitu et al. 2018). For these reasons, it can be difficult to accurately identify the beginning and end times of solid precipitation events using tipping-bucket measurements (Nitu et al. 2018). Because of this, tipping-bucket gauges are generally more suitable for daily or seasonal precipitation measurements than for 30-min or hourly measurements.

Buisán et al. (2017) and Kochendorfer et al. (2020) derived transfer functions to adjust solid precipitation measurements from unshielded tipping-bucket gauges at SPICE sites, and calculated the resultant uncertainties. To account for tipping-bucket measurement delays, new techniques were developed to derive these transfer functions, some of which may also be useful for the derivation of weighing gauge transfer functions. To include all available measurements in the derivation of tipping-bucket transfer functions, instead of only periods when the tipping-bucket gauge and the DFAR simultaneously caught precipitation, Kochendorfer et al. (2020) optimized transfer function parameters to minimize errors in the seasonal precipitation accumulation. Some of the limitations of the traditional catch efficiency approach were thus overcome, allowing for more representative and customizable transfer functions.

Non-catchment sensors

New technologies that measure precipitation without capturing and collecting hydrometeors (so-called "non-catchment" type instruments, e.g., Fig. 7) are increasingly being used operationally. Three different types of non-catchment sensors were tested in SPICE: optical disdrometers, which determine hydrometeor characteristics based on their attenuation of a laser beam; present weather sensors, which use light scattering to estimate hydrometeor properties; and an evaporative plate, which determines precipitation amounts based on the principles of mass and heat transfer. Instruments using microwave backscatter (i.e., X-band or K-band radars) to determine hydrometeor properties are also available, but were not evaluated in SPICE.

Optical disdrometers and present weather sensors are designed primarily for detecting precipitation type, but many of them also produce a derived precipitation amount or precipitation rate. This makes such sensors attractive to many users. But the SPICE results, which focused on the use of non-catchment instruments for reporting accumulated precipitation amounts, demonstrated that such sensors can significantly over- and underestimate precipitation amounts relative to the DFAR (Reverdin et al. 2016; Nitu et al. 2018). At one of the SPICE sites, Begueria et al. (2018) likewise found that such optical precipitation measurements overestimated solid precipitation amounts relative to the DFAR. The varied performance of such sensors is attributed primarily to assumptions used in data processing methods (e.g., assumed hydrometeor shape, density of

Fig. 7. Disdrometer intercomparison at the Formigal SPICE site in Spain, with the DFAR in the background.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

solid precipitation, etc.), and is more pronounced over shorter periods (hourly to daily) than over longer periods (months to seasons), for which the assumptions are likely more representative. For laser disdrometers, the superposition of hydrometeors along the beam is another potential source of uncertainty.

Within SPICE, non-catchment sensors were generally less susceptible to wind effects relative to weighing gauges and heated tipping-bucket gauges. Wind direction, however, was found to impact performance depending on the specific instrument and its configuration (e.g., sensor heads can block precipitation coming from certain directions). To address this, some instruments have been designed with aerodynamic profiles to mitigate such anisotropic effects.

The Hotplate Precipitation Gauge, which relies on an evaporative plate to determine precipitation amount, was unique among the non-catchment sensors tested in terms of both its operational principles and its measurement results. For example, the solid precipitation measurements recorded using the Hotplate had a lower bias and RMSE compared to unshielded weighing gauge measurements (Nitu et al. 2018). Other recent research on the Hotplate Precipitation Gauge likewise indicated excellent agreement with the SPICE DFAR precipitation measurements (Thériault et al. 2021), and catch efficiencies that increased with wind speed in modeled simulations (Cauteruccio et al. 2021).

Snow cover

Snow cover measurements. In SPICE, automated snow depth and SWE measurements were evaluated relative to manual reference measurements. Examples of such sensors are shown in Fig. 8. Automated snow depth sensors, based on ultrasonic or laser measurement principles, performed well relative to manual measurements. The results demonstrated the importance of spatial variability, driven in particular by the redistribution of snow by wind. At the CARE site in Canada, manual snow depth measurements were used to quantify this spatial variability (Fig. 9), emphasizing the importance of sensor siting. The benefits of heating the snow depth sensors, angled mounting structures, and the use of artificial surface targets under the snow depth sensors were also analyzed within SPICE.

The performance of automated SWE sensors varied depending on the measurement principle and environmental conditions. Like the snow depth measurements, the spatial variability of the snow cover at the scale of a few meters distance significantly affected the SWE measurements, regardless of the measurement principle. The final report from SPICE listed a number of emerging approaches for measuring snow depth or SWE, based on varying measurement principles. Some of these new methods assess the spatial variability of snow cover characteristics, as opposed to more traditional single-point measurements (see Nitu et al. 2018, section 4.1.4.5); however, most of these emerging techniques are not used operationally.

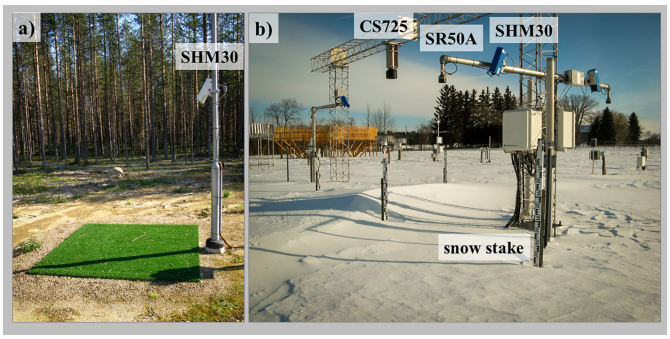

Fig. 8. Photos of snow depth sensor installations under assessment during SPICE: (a) an optical sensor (Lufft/Jenoptik SHM30) at Sodankylä prior to snow accumulation on a surface target constructed from artificial turf; (b) optical (Lufft/Jenoptik SHM30) and ultrasonic (Campbell Scientific SR50A) snow depth sensors concurrently measuring the same target area at CARE alongside manually observed graduated snow stakes used as the reference, and a Campbell Scientific CS725 passive gamma snow water equivalent sensor.

Operational snow cover considerations. While the research community has remained active in developing novel approaches for measuring snow cover or assessing the uncertainties of measurements (e.g., for manual SWE measurements by López-Moreno et al. 2020), a large gap remains between the state of technological development of current prototypes and the

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

way snow cover is measured in operational networks. The assimilation of SPICE results relevant to snow cover measurements (snow depth or SWE) in operational networks has, up until now, remained negligible or anecdotal. A recent synthesis of snow cover measurements in Europe (Haberkorn 2019) describes how snow depth, presence of snow on the ground, depth of new snowfall, and SWE are measured in 38 countries in Europe. With over 13,251 snow depth measurement stations across Europe, the vast majority (85%) employ manual measurements, especially at low elevation. Automated snow depth measurements are more prominent at high elevation, with 355 automated snow depth stations compared with 200 manual stations above 2,000 m elevation. The partitioning between automated and manual measurements is, however, very heterogeneous across Europe, not only due to topography (fraction of mountainous terrain), but also due to differences in network development and management strategy.

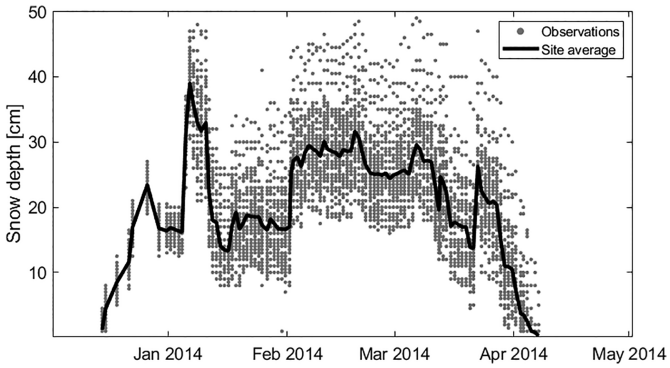

Fig. 9. Manual snow depth measurements recorded at the CARE site over the course of the 2013/14 winter season. Individual measurements (gray circles) were recorded at 62 different snow stakes, typically daily, to quantify spatial variability. The average snow depth (black line) was calculated for each observation period.

The situation for SWE measurements is even more heterogeneous, with 4,044 SWE measurement stations, 97% of which are manual. The lack of harmonization of snow cover measurements, in terms of both measurement principles and quality control methods, hinders the use of the data across national borders (Beniston et al. 2018; Matiu et al. 2021). Progress in snow cover monitoring is therefore predicated on further advances in the exchange and homogenization of snow cover data. This gap has been identified in the latest IPCC Special Report on the Ocean and Cryosphere in Changing Climate (IPCC 2019).

Although operational snow cover measurement networks are advancing slowly, a strong emphasis has been placed on the development of remote sensing products (e.g., Gascoin et al. 2019). This emphasis has now expanded to a pan-European domain as a dedicated Copernicus Land Monitoring service (https://land.copernicus.eu/pan-european/biophysical-parameters/high-resolution-snow-and-ice-monitoring), and includes the combination of remotely sensed data with in situ observations and model results through data assimilation (e.g., Largeron et al. 2020; Cluzet et al. 2021). In situ observations of snow cover are a key component of these innovative snow monitoring and prediction systems, amplifying the need for further harmonization and consolidation of operational snow cover networks.

Snow cover measurement recommendations. For the measurement of snow depth, heating of the sensors and the use of angled (and heated) mounting arms in regions of high snow accumulation are recommended to mitigate potential measurement impacts caused by the accumulation of snow or ice on or around the sensor. Heating of sensor mounting infrastructure is recommended to prevent built-up snow or ice from falling onto the sampling area and impacting the snow cover. Artificial surface targets can be used to provide uniform, level, stable, and low maintenance surfaces for snow depth measurements (Fig. 8a); however, the relative distance between the sensor and target can still be impacted by frost heave and settling. These targets are especially advantageous for ultrasonic sensors, which have larger sample areas than optical sensors, and are more susceptible to nonuniform features of natural targets (e.g., protruding vegetation). Artificial targets should be similar to the surrounding soil in terms of height, color, reflectivity, and thermal properties, to minimize unrepresentative measurements due to surface heterogeneity when snow is accumulating or melting.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

Collection efficiency modeling

Advances in modeling the flow and trajectory of hydrometeors around gauges. Significant advances in the modeling of flow fields in the vicinity of precipitation gauge and trajectories of various snowflake types have been made since Rasmussen et al. (2012). Thériault et al. (2012) simulated the flow around a single-Alter shield and showed the impact of snowflakes types on the scatter in the collection efficiency. Colli et al. (2015) computed the flow field around unshielded and single-Alter shielded Geonor gauges and demonstrated the sensitivity of collection efficiency results to the specific hydrometeor drag model used. Further studies by Colli et al. (2016a,b) showed that snowflake trajectories past and into the gauge were significantly impacted by the airflow distortion, both spatially and temporally.

Baghapour et al. (2017), Baghapour and Sullivan (2017), and Colli et al. (2016a,b) used numerical Reynolds averaged Navier Stokes (RANS) and large-eddy simulation (LES) models to study the turbulence and eddy dynamics around a Geonor precipitation gauge in unshielded, single-Alter shielded, and double-Alter shielded configurations. The simulation results illustrated that shielding decreased the flow momentum above the gauge orifice at the expense of increased turbulence generation, while tilting the Alter shield slats in high winds reduced the shielding benefits. Further study by Baghapour and Sullivan (2017) and Colli et al. (2015, 2016a,b) demonstrated how the effective catchment area of precipitation gauges changes with wind speed and snowflake characteristics. For small hydrometeors in high winds, the effective catchment area can be reduced significantly, but can be partially improved by shielding. The modeled results also suggest that the flow field around the DFAR impacts the angle of attack and produces flow convergence, which increases the number of snowflakes falling in the gauge (Thériault et al. 2015).

Collection efficiency and hydrometeor fall velocity. Thériault et al. (2012) showed that slow falling hydrometeors (e.g., dry snow) have a lower catch efficiency than faster-falling hydrometeors (e.g., wet snow), which explains some of the uncertainty in the collection efficiency at a given wind speed. This was demonstrated using both CFD simulations and detailed field measurements of snowflake habit. Building upon these findings, Hoover et al. (2021) characterized how the collection efficiency of unshielded Geonor gauges changes with wind speed and hydrometeor fall velocity using CFD simulations and Lagrangian particle tracking. Based on the numerical results, a universal transfer function was developed.

Recent studies by Leroux et al. (2021) and Hoover et al. (2021) have demonstrated that hydrometeor fall velocity measurements can be used to develop new transfer functions with wind speed and fall velocity dependence. Such hydrometeor fall velocity measurements can be obtained from optical disdrometers and Doppler radar instruments. These new transfer functions reduce the uncertainty in adjusted precipitation measurements (lower RMSE) relative to existing transfer functions (Leroux et al. 2021; Hoover et al. 2021). These studies demonstrate the importance of snowflake habit on gauge collection efficiency and the utility of fall velocity measurements for developing and applying adjustments.

Transfer function applications

Transfer functions for use in validating weather models. Numerical weather models are traditionally verified with precipitation gauge observations without any form of adjustment for wind-induced undercatch, though more recent studies acknowledge the undercatch of solid precipitation as one of the major reasons for the discrepancies between the measured and modeled solid precipitation (Gowan et al. 2018; Wang et al. 2019). Køltzow et al. (2019, 2020) and Buisán et al. (2020) quantified the possible impact with their respective case studies, using transfer functions derived from SPICE sites.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

Køltzow et al. (2019) used raw and adjusted observational data from the Year of Polar Prediction project to verify solid precipitation forecasts from a number of numerical models. The models showed an unambiguous underestimation of the mean precipitation and a clear change in model behavior. In a follow-up study, Køltzow et al. (2020) investigated the verification of solid precipitation forecasts in Norway. They compared the model results to both automated and manual precipitation measurements, applying universal and site-specific transfer functions to the observations. The results demonstrated the dependency of catch efficiency on temperature and wind speed. All three studies (Buisan et al. 2020; Køltzow et al. 2019, 2020) concluded that the application of adjustment functions reduces precipitation measurement errors, thus improving model verification and decreasing systematic forecast biases. However, caution when applying adjustments was advised by all of these studies due to the uncertainties in identifying the optimum adjustment for a given observational site, and the errors inherent in applying any transfer function (Buisan et al. 2020; Køltzow et al. 2019, 2020; Pierre et al. 2019; Smith et al. 2020).

Precipitation biases and regional hydrology. Precipitation biases can affect hydrologic studies at all scales, because precipitation measurements are often used to drive hydrologic models (Coustau et al. 2015). Robinson and Clark (2020) recently used a water balance approach based on the Gravity Recovery and Climate Experiment (GRACE) total water storage to infer the amount of cold-season precipitation in four large Arctic rivers. They evaluated four gridded meteorological datasets as inputs to a land surface model, and reported that the cold-season precipitation in these datasets needed to increase by up to 55%. Undercatch-adjusted precipitation gauge measurements were also compared with the GRACE-derived correction. The undercatch correction increased the amount of cold-season precipitation by 23%, indicating that some, but not all, of the underestimation was removed.

Lundquist et al. (2019) and Lussana et al. (2018) also demonstrated discrepancies between model simulations and hydrological observations. These differences were likely caused by wind-induced undercatch, as they were associated with solid and mixed precipitation under windy conditions and in complex terrain. Using gridded wind data, Lussana et al. (2019) applied the adjustment function from Wolff et al. (2015) to improve a gridded precipitation dataset from the Norwegian Meteorological Institute. In the process, new constants for the adjustment functions were derived by optimizing the functions to reproduce extreme precipitation events based on geographical parameters characterizing site exposure. The qualitative analysis by Lussana et al. (2019) indicated an improved agreement between the gridded dataset and independent hydrological observations, as well as the requirement for further evaluation.

Other WMO, metrology, and snow cover work

WMO measurement guidance and reports. Recommendations from SPICE and best practices for the automated measurement of snow cover have been incorporated into the WMO Guide to Instruments and Methods (WMO 2018b). The existing guidelines on precipitation measurement (WMO 2018a) will also be updated. In addition, the Expert Team on Surface and Subsurface Measurements of the WMO Infrastructure Commission is modernizing the guidance for the measurement of precipitation, with a shift in focus from manual to automated measurements. This work will include non-catchment measurement techniques, such as disdrometers. The significant SPICE contributions to the field of automated solid precipitation measurement will also be incorporated into this guidance.

Contributions from the metrology community. The metrology community (MeteoMet) is also developing traceability and calibration methods for non-catchment precipitation sensors

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

(www.meteomet.org/incipit/). This project will include new modeling tools developed to better describe non-catchment measurements and their errors, a raindrop generator for calibration and experiments, and field and laboratory validation of new methodologies.

WMO-managed metadata. The proliferation of new sensors for the measurement of solid precipitation and snow cover has greatly increased the importance and challenge of documenting existing measurement configurations. Metadata describing sensor types and their field configuration (presence and type of shield, surface target, etc.) needs to be made available in standardized and accessible formats. Most operational stations, globally, are represented in the OSCAR/Surface database maintained by WMO, based on the Metadata standard. Currently, however, only a small subset of operationally available stations measuring snow depth and total precipitation (solid precipitation included) are fully represented and characterized in OSCAR/Surface. This limits our ability to standardize precipitation measurements and derive and apply transfer functions.

Conclusions

SPICE was an international collaboration developed to assess instrument performance and provide recommendations on best practices for solid precipitation and snow cover measurements. It was intended to aid in the transition from manual to automated measurements and to establish an automated standard for solid precipitation measurement. The inclusion of observing sites from 20 locations around the world allowed for data collection in a variety of climate regimes and under varying types of snowfall conditions, providing a robust dataset of observations from over 200 different sensors.

Analysis of the data from SPICE resulted in many outcomes and accomplishments. One of the significant accomplishments was the establishment of the DFAR as an automated field reference configuration for solid precipitation measurements. This subsequently allowed for the development of transfer functions to adjust for gauge undercatch due to wind, which was the primary performance issue for weighing gauges. Further analysis indicated that transfer functions were largely independent of the type of weighing gauge used, and depended primarily on the shield configuration. This mirrored the trends for catch efficiency results among different weighing gauges in the same shield configuration at SPICE sites.

The application of the transfer functions derived from multiple SPICE site datasets to post-SPICE datasets from individual sites was found to produce a significant underadjustment in cold and windy locations, while tending to overadjust at less windy sites. The results also suggest that these transfer functions, while useful, should be applied carefully, as local meteorological, topographical, and climatological conditions are important for understanding the undercatch at a given site. It was also noted that single-Alter shielded and unshielded precipitation gauges tend to miss low-accumulation events in windy conditions, supporting the requirement for improved and well-maintained gauge shielding.

The SPICE results also highlighted some of the issues associated with heated tipping-bucket gauges. Similar to weighing gauges, the performance of tipping buckets is strongly impacted by wind speed. Tipping-bucket gauges are less accurate than weighing gauges, leading to difficulties in identifying the beginning and end of solid precipitation events. Heated tipping-bucket gauges are additionally affected by evaporation and response delays, and more research is needed to develop methods to minimize these drawbacks. Transfer functions were also developed for heated tipping-bucket gauges, with novel methods developed to mitigate the influence of both wind speed and response delays on seasonal datasets.

Disdrometers and present weather sensors were found to both overestimate and underestimate precipitation amounts relative to the DFAR. This was attributed primarily to assumptions used in the data processing methods, and was more pronounced over shorter periods (hourly

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

to daily) than longer periods (months to seasons). For laser disdrometers in particular, the juxtaposition of hydrometeors along the beam was found to be another potential source of uncertainty. For the evaporative plate instrument, solid precipitation measurements had a lower bias and RMSE as compared to unshielded weighing gauge measurements.

Automated snow depth and SWE measurements were evaluated relative to manual reference measurements. The performance of these sensors varied depending on the measurement principle and environmental conditions, but they generally behaved according to the manufacturer's specification. An overarching theme of this research was the importance of spatial variability, and the necessity of using more than a single point measurement to establish a representative value for all types of snow cover measurements.

SPICE also provided opportunities for collaborative studies to model the flow and trajectory of hydrometeors around gauges and wind shields. Results showed that snowflake trajectories around the gauge were significantly impacted by airflow distortion (both spatially and temporally) caused by the presence of the gauge. Furthermore, simulations of airflow past a Geonor gauge with an Alter shield (either single or double) illustrated that shielding decreased the flow above the gauge orifice and increased turbulence. Additionally, it was shown that the deflection of the Alter shield slats in high winds reduced the shielding benefits. Hydrometeor fall speeds were also shown to have an impact on collection efficiency, with slow falling hydrometeors having a lower catch efficiency than faster-falling hydrometeors.

Finally, a series of recommendations and best practices were developed to help inform sensor users and observation programs regarding the installation, operation, and maintenance experiences of the site teams. These include 1) installing at least one DFAR configuration to characterize the measurement biases and uncertainty of gauges within a given agency or network; 2) applying appropriate transfer functions to mitigate the influence of wind-induced undercatch on solid precipitation measurements; 3) installation of wind shields for catchment-type precipitation gauges to reduce the effects of wind on gauge undercatch; 4) utilization of antifreeze and oil in weighing gauges to prevent the freezing and evaporation of bucket contents; 5) adding orifice heating to improve gauge response times for solid precipitation and to mitigate the influence of snow capping; 6) heating of snow depth sensors and the use of angled (and heated) mounting arms in regions of high snow accumulation; 7) adding artificial surface targets to provide uniform, level, stable, and low maintenance surfaces for ultrasonic snow depth measurements; and 8) improving and sharing metadata describing measurement type and configuration to aid in national and international data homogenization.

Acknowledgments. The authors thank Hagop Mouradian from Environment and Climate Change Canada for contributing the mapped site locations (Fig. 1). We also thank the World Meteorological Organization for supporting SPICE. This research was also sponsored by the National Oceanic and Atmospheric Administration's (NOAA) USCRN program. The authors also acknowledge the support of Howard Diamond, program manager for the USRCRN. The views expressed are those of the authors and do not necessarily represent the official policy or position of the U.S. government. The authors also thank the Formigal Ski Resort for its support in the installation and maintenance of the Spanish SPICE site.

Data availability statement. All SPICE data used in this study are available from the NCAR Research Applications Laboratory database at https://ral.ucar.edu/projects/SPICE/. The post-SPICE data from Smith et al. (2020), which were used in Figs. 3 and 5, are available at https://doi.org/10.1594/PANGAEA.907379.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

References

Adam, J., and D. P. Lettenmaier, 2003: Adjustment of global gridded precipitation for systematic bias. *J. Geophys. Res.*, **108**, 4257–4272, https://doi.org/10.1029/2002JD002499.

Alter, J. C., 1937: Shielded storage precipitation gages. *Mon. Wea. Rev.*, **65**, 262–265, https://doi.org/10.1175/1520-0493(1937)65<262:SSPG>2.0.CO;2.

Arnell, N. W., 1999: Climate change and global water resources. *Global Environ. Change*, **9** (Suppl. 1), S31–S49, https://doi.org/10.1016/S0959-3780(99)00017-5.

Baghapour, B., and P. E. Sullivan, 2017: A CFD study of the influence of turbulence on undercatch of precipitation gauges. *Atmos. Res.*, **197**, 265–276, https://doi.org/10.1016/j.atmosres.2017.07.008.

——, C. Wei, and P. E. Sullivan, 2017: Numerical simulation of wind-induced turbulence over precipitation gauges. *Atmos. Res.*, **189**, 82–98, https://doi.org/10.1016/j.atmosres.2017.01.016.

Begueria, S., S. T. Buisán, J. L. Collado, and J. Alastrue, 2018: Impact of wind and temperature on snowfall measurements by three Thies LPM and three OTT Parsivel2 compared with DFAR (Double Fence Automated Reference) measurements at WMO.SPICE Formigal-Sarrios site. *WMO Tech. Conf. on Meteorological and Environmental Instruments and Methods of Observation, 2018*, Amsterdam, Netherlands, WMO, P3_18, https://community.wmo.int/publications-and-iom-reports/teco-2018-presentations.

Beniston, M., and Coauthors, 2018: The European mountain cryosphere: A review of its current state, trends, and future challenges. *Cryosphere*, **12**, 759–794, https://doi.org/10.5194/tc-12-759-2018.

Broxton, P. D., N. Dawson, and X. Zeng, 2016: Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth. *Earth Space Sci.*, **3**, 246–256, https://doi.org/10.1002/2016EA000174.

Buisán, S. T., and Coauthors, 2017: Assessment of snowfall accumulation underestimation by tipping bucket gauges in the Spanish operational network. *Atmos. Meas. Tech.*, **10**, 1079–1091, https://doi.org/10.5194/amt-10-1079-2017.

——, and Coauthors, 2020: The potential for uncertainty in numerical weather prediction model verification when using solid precipitation observations. *Atmos. Sci. Lett.*, **21**, e976, https://doi.org/10.1002/asl.976.

Cauteruccio, A., E. Brambilla, M. Stagnaro, L. G. Lanza, and D. Rocchi, 2021: Experimental evidence of the wind-induced bias of precipitation gauges using particle image velocimetry and particle tracking in the wind tunnel. *J. Hydrol.*, **600**, 126690, https://doi.org/10.1016/j.jhydrol.2021.126690.

Cluzet, B., and Coauthors, 2021: CrocO_v1.0: A particle filter to assimilate snowpack observations in a spatialised framework. *Geosci. Model Dev.*, **14**, 1595–1614, https://doi.org/10.5194/gmd-14-1595-2021.

Colli, M., and Coauthors, 2015: An improved trajectory model to evaluate the collection performance of snow gauges. *J. Appl. Meteor. Climatol.*, **54**, 1826–1836, https://doi.org/10.1175/JAMC-D-15-0035.1.

——, L. G. Lanza, R. Rasmussen, and J. M. Thériault, 2016a: The collection efficiency of shielded and unshielded precipitation gauges: Part I: CDF airflow modelling. *J. Hydrometeor.*, **17**, 231–243, https://doi.org/10.1175/JHM-D-15-0010.1.

——, ——, ——, and ——, 2016b: The collection efficiency of shielded and unshielded precipitation gauges: Part II: Modeling particle trajectories. *J. Hydrometeor.*, **17**, 245–255, https://doi.org/10.1175/JHM-D-15-0011.1.

Coustau, M., F. Rousset-Regimbeau, G. Thirel, F. Habets, B. Janet, E. Martin, C. de Saint-Aubin, and J.-M. Soubeyroux, 2015: Impact of improved meteorological forcing, profile of soil hydraulic conductivity and data assimilation on an operational hydrological ensemble forecast system over France. *J. Hydrol.*, **525**, 781–792, https://doi.org/10.1016/j.jhydrol.2015.04.022.

Daly, C., M. E. Slater, J. A. Roberti, S. H. Laseter, and L. W. Swift, 2017: High-resolution precipitation mapping in a mountainous watershed: Ground truth for evaluating uncertainty in a national precipitation dataset. *Int. J. Climatol.*, **37**, 124–137, https://doi.org/10.1002/joc.4986.

Derksen, C., and R. Brown, 2012: Spring snow cover extent reductions in the 2008–2012 period exceeding climate model predictions. *Geophys. Res. Lett.*, **39**, L19504, https://doi.org/10.1029/2012GL053387.

——, and Coauthors, 2019: Changes in snow, ice, and permafrost across Canada. *Canada's Changing Climate Report*, E. Bush and D. S. Lemmen, Eds., Government of Canada, 194–260.

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

Ding, Y., D. Yang, B. Ye, and N. Wang, 2007: Effects of bias correction on precipitation trend over China. *J. Geophys. Res.*, **112**, D13116, https://doi.org/10.1029/2006JD007938.

Dressler, K. A., S. R. Fassnacht, and R. C. Bales, 2006: A comparison of snow telemetry and snow course measurements in the Colorado River basin. *J. Hydrometeor.*, **7**, 705–712, https://doi.org/10.1175/JHM506.1.

Gascoin, S., M. Grizonnet, M. Bouchet, G. Salgues, and O. Hagolle, 2019: Theia Snow collection: High-resolution operational snow cover maps from Sentinel-2 and Landsat-8 data. *Earth Syst. Sci. Data*, **11**, 493–514, https://doi.org/10.5194/essd-11-493-2019.

Goodison, B., P. Louie, and D. Yang, 1998: WMO solid precipitation measurement intercomparison. WMO/TD-872, IOM Rep. 67, 212 pp., https://library.wmo.int/doc_num.php?explnum_id=9694.

Gowan, T. M., W. J. Steenburgh, and C. S. Schwartz, 2018: Validation of mountain precipitation forecasts from the convection-permitting NCAR ensemble and operational forecast systems over the western United States. *Wea. Forecasting*, **22**, 739–765, https://doi.org/10.1175/WAF-D-17-0144.1.

Gugerli, R., M. Guidicelli, M. Gabella, M. Huss, and N. Salzmann, 2021: Multisensor analysis of monthly gridded snow precipitation on alpine glaciers. *Adv. Sci. Res.*, **18**, 7–20, https://doi.org/10.5194/asr-18-7-2021.

Haberkorn, A., 2019: European Snow Booklet – An inventory of snow measurements in Europe. EnviDat, https://doi.org/10.16904/envidat.59.

Hoover, J., M. E. Earle, P. I. Joe, and P. E. Sullivan, 2021: Unshielded precipitation gauge collection efficiency with wind speed and hydrometeor fall velocity. *Hydol. Earth Syst. Sci.*, **25**, 5473–5491, https://doi.org/10.5194/hess-25-5473-2021.

IPCC, 2019: High mountain areas. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, H.-O. Pörtner et al., Eds., 131–202, www.ipcc.ch/srocc/.

Kochendorfer, J., and Coauthors, 2017: Analysis of single-Alter-shielded 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.

——, and Coauthors, 2018: Testing and development of transfer functions for weighing precipitation gauges in WMO-SPICE. *Hydrol. Earth Syst. Sci.*, **22**, 1437–1452, https://doi.org/10.5194/hess-22-1437-2018.

——, and Coauthors, 2020: Undercatch adjustments for tipping bucket gauge measurements of solid precipitation. *J. Hydrometeor.*, **21**, 1193–1205, https://doi.org/10.1175/JHM-D-19-0256.1.

Køltzow, M., B. Casati, E. Bazile, T. Haiden, and T. Valkonen, 2019: An NWP model intercomparison of surface weather parameters in the European Arctic during the year of polar prediction special observing period Northern Hemisphere 1. *Wea. Forecasting*, **34**, 959–983, https://doi.org/10.1175/WAF-D-19-0003.1.

——, ——, T. Haiden, and T. Valkonen, 2020: Verification of solid precipitation forecasts from numerical weather prediction models in Norway. *Wea. Forecasting*, **35**, 2279–2292, https://doi.org/10.1175/WAF-D-20-0060.1.

Langousis, A., R. Deidda, A. A. Carsteanu, C. Onof, P. Burlando, R. Uijlenhoet, and A. Bárdossy, 2018: Precipitation measurement and modelling: Uncertainty, variability, observations, ensemble simulation and downscaling. *J. Hydrol.*, **556**, 824–826, https://doi.org/10.1016/j.jhydrol.2017.09.016.

Largeron, C., and Coauthors, 2020: Toward snow cover estimation in mountainous areas using modern data assimilation methods: A review. *Front. Earth Sci.*, **8**, 325, https://doi.org/10.3389/feart.2020.00325.

Leroux, N. R., J. M. Thériault, and R. Rasmussen, 2021: Improvement of snow gauge collection efficiency through a knowledge of solid precipitation fall speed. *J. Hydrometeor.*, **22**, 997–1006, https://doi.org/10.1175/JHM-D-20-0147.1.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

Li, Z., and Coauthors, 2020: Declining snowfall fraction in the alpine regions, Central Asia. *Sci. Rep.*, **10**, 3476, https://doi.org/10.1038/s41598-020-60303-z.

López-Moreno, J. I., and Coauthors, 2020: Intercomparison of measurements of bulk snow density and water equivalent of snow cover with snow core samplers: Instrumental bias and variability induced by observers. *Hydrol. Processes*, **34**, 3120–3133, https://doi.org/10.1002/hyp.13785.

Lundquist, J., M. Hughes, E. Gutmann, and S. Kapnick, 2019: Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks. *Bull. Amer. Meteor. Soc.*, **100**, 2473–2490, https://doi.org/10.1175/BAMS-D-19-0001.1.

Lussana, C., and Coauthors, 2018: seNorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day. *Earth Syst. Sci. Data*, **10**, 235–249, https://doi.org/10.5194/essd-10-235-2018.

——, O. E. Tveito, A. Dobler, and K. Tunheim, 2019: seNorge_2018, daily precipitation, and temperature datasets over Norway. *Earth Syst. Sci. Data*, **11**, 1531–1551, https://doi.org/10.5194/essd-11-1531-2019.

MacDonald, J., and P. W. Pomeroy, 2007: Gauge undercatch of two common snowfall gauges in a prairie environment. *Proc. 64th Eastern Snow Conf.*, St. John's, NL, Canada, Eastern Snow Conference, 119–124, www.easternsnow.org/esc-2007.

Matiu, M., and Coauthors, 2021: Observed snow depth trends in the European Alps: 1971 to 2019. *Cryosphere*, **15**, 1343–1382, https://doi.org/10.5194/tc-15-1343-2021.

Mekis, E., and Coauthors, 2018: An overview of the surface-based precipitation observations at environment and climate change Canada. *Atmos.–Ocean*, **56**, 71–95, https://doi.org/10.1080/07055900.2018.1433627.

Metcalfe, R. J., and E. B. Goodison, 1993: Correction of Canadian winter precipitation data. Preprints, *Eighth Symp. on Meteorological Observations and Instrumentation*, Anaheim, CA, Amer. Meteor. Soc., 338–343.

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

Mote, P. W., S. Li, D. P. Lettenmaier, M. Xiao, and R. Engel, 2018: Dramatic declines in snowpack in the western US. *npj Climate Atmos. Sci.* **1**, 2, https://doi.org/10.1038/s41612-018-0012-1.

Nitu, R., and K. Wong, 2010: CIMO survey on national summaries of methods and instruments for solid precipitation measurement at automatic weather stations. World Meteorological Organization Instruments and Observing Methods Rep. 102, WMO/TD-1544, 57 pp., https://library.wmo.int/doc_num.php?explnum_id=9443.

——, and Coauthors, 2018: WMO Solid Precipitation Intercomparison Experiment (SPICE) (2012–2015). IOM Rep. 131, 1445 pp., https://library.wmo.int/doc_num.php?explnum_id=5686.

Orphanopoulos, D., K. Verbist, A. Chavez, and G. Soto, 2013: Water use efficiency in an arid watershed: A case study. *Sci. Cold Arid Reg.*, **5**, 16–26, https://doi.org/10.3724/SP.J.1226.2013.00016.

Pan, X., and Coauthors, 2016: Bias corrections of precipitation measurements across experimental sites in different ecoclimatic regions of western Canada. *Cryosphere*, **10**, 2347–2360, https://doi.org/10.5194/tc-10-2347-2016.

——, D. Yang, K. P. Chun, J. Zhang, and Y. You, 2019: Under-measured daily maximum precipitation from manual gauge observations over the northern regions. *Sci. Total Environ.*, **715**, 136970, https://doi.org/10.1016/j.scitotenv.2020.136970.

Picard, G., L. Arnaud, J.-M. Panel, and S. Morin, 2016: Design of a scanning laser meter for monitoring the spatio-temporal evolution of snow depth and its application in the Alps and in Antarctica. *Cryosphere*, **10**, 1495–1511, https://doi.org/10.5194/tc-10-1495-2016.

Pierre, A., S. Jutras, C. Smith, J. Kochendorfer, V. Fortin, and F. Anctil, 2019: Evaluation of catch efficiency transfer functions for unshielded and single-Alter-shielded solid precipitation measurements. *J. Atmos. Oceanic Technol.*, **36**, 865–881, https://doi.org/10.1175/JTECH-D-18-0112.1.

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

Reverdin, A., S. Buisán, Y. A. Roulet, J. L. Collado, and J. Alastrue, 2016: Intercomparison of snowfall measurements using disdrometers in two mountainous environments: Weissfluhjoch (Switzerland) and Formigal-Sarrios (Spain). *WMO Technical Conf. on Meteorological and Environmental Instruments and Methods of Observation, 2016*, Madrid, Spain, WMO, IOM Rep. 125, P3(11), https://library.wmo.int/doc_num.php?explnum_id=3226.

Robinson, E. L., and D. B. Clark, 2020: Using Gravity Recovery and Climate Experiment data to derive corrections to precipitation data sets and improve modelled snow mass at high latitudes. *Hydrol. Earth Syst. Sci.*, **24**, 1763–1779, https://doi.org/10.5194/hess-24-1763-2020.

Scaff, L., D. Yang, Y. Li, and E. Mekis, 2015: Inconsistency in precipitation measurements across the Alaska–Yukon border. *Cryosphere*, **9**, 2417–2428, https://doi.org/10.5194/tc-9-2417-2015.

Schaffer, N., and Coauthors, 2019: Rock glaciers as a water resource in a changing climate in the semiarid Chilean Andes. *Reg. Environ. Change*, **19**, 1263–1279, https://doi.org/10.1007/s10113-018-01459-3.

Sevruk, B., 1983: Correction of measured precipitation in the Alps using the water equivalent of new snow. *Hydrol. Res.*, **14**, 49–58, https://doi.org/10.2166/nh.1983.0005.

Smith, C. D., and Coauthors, 2020: Evaluation of the WMO Solid Precipitation Intercomparison Experiment (SPICE) transfer functions for adjusting the wind bias in solid precipitation measurements. *Hydrol. Earth Syst. Sci.*, **24**, 4025–4043, https://doi.org/10.5194/hess-24-4025-2020.

Thériault, J. M., R. Rasmussen, K. Ikeda, and S. Landolt, 2012: Dependence of snow gauge collection efficiency on snowflake characteristics. *J. Appl. Meteor. Climatol.*, **51**, 745–762, https://doi.org/10.1175/JAMC-D-11-0116.1.

——, ——, E. Petro, J. Trépanier, M. Colli, and L. G. Lanza, 2015: Impact of wind direction, wind speed, and particle characteristics on the collection efficiency of the Double Fence Intercomparison Reference. *J. Appl. Meteor. Climatol.*, **54**, 1918–1930, https://doi.org/10.1175/JAMC-D-15-0034.1.

——, N. R. Leroux, and R. M. Rasmussen, 2021: Improvement of solid precipitation measurements using a hotplate precipitation gauge. *J. Hydrometeor.*, **22**, 877–885, https://doi.org/10.1175/JHM-D-20-0168.1.

Tian, X., A. Dai, D. Yang, and Z. Xie, 2007: Effects of precipitation-bias corrections on surface hydrology over northern latitudes. *J. Geophys. Res.*, **112**, D14101, https://doi.org/10.1029/2007JD008420.

Trenberth, K. E., 2011: Changes in precipitation with climate change. *Climate Res.*, **47**, 123–138, https://doi.org/10.3354/cr00953.

——, A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.*, **84**, 1205–1218, https://doi.org/10.1175/BAMS-84-9-1205.

Wagner, D. N., and Coauthors, 2021: Snowfall and snow accumulation processes during MOSAiC. *EGU General Assembly 2021*, Online, EGU21-12692, https://doi.org/10.5194/egusphere-egu21-12692.

Wang, G., X. Zhang, and S. Zhang, 2019: Performance of three reanalysis precipitation datasets over the Qinling-Daba Mountains, eastern fringe of Tibetan Plateau, China. *Adv. Meteor.*, **2019**, 7698171, https://doi.org/10.1155/2019/7698171.

Wen, Y., A. Behrangi, B. Lambrigtsen, and P.-E. Kirstetter, 2016: Evaluation and uncertainty estimation of the latest radar and satellite snowfall products using SNOTEL measurements over mountainous regions in western United States. *Remote Sens.*, **8**, 904, https://doi.org/10.3390/rs8110904.

WMO, 2018a: Measurement of Meteorological Variables. Vol. I, Guide to Instruments and Methods of Observation, WMO-8, World Meteorological Organization, 548 pp., https://library.wmo.int/doc_num.php?explnum_id=10616.

——, 2018b: Measurement of Cryospheric Variables. Vol. II, Guide to Instruments and Methods of Observation, WMO-8, World Meteorological Organization, 42 pp., https://library.wmo.int/doc_num.php?explnum_id=9870.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC

Wolff, M. A., and Coauthors, 2015: Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: Results of a Norwegian field study. *Hydrol. Earth Syst. Sci.*, **19**, 951–967, https://doi.org/10.5194/hess-19-951-2015.

Yang, D., 1999: An improved precipitation climatology for the Arctic Ocean. *Geophys. Res. Lett.*, **26**, 1625–1628, https://doi.org/10.1029/1999GL900311.

——, 2014: Double Fence Intercomparison Reference (DFIR) vs. Bush Gauge for "true" snowfall measurement. *J. Hydrol.*, **509**, 94–100, https://doi.org/10.1016/j.jhydrol.2013.08.052.

——, and T. Ohata, 2001: A bias-corrected Siberian regional precipitation climatology. *J. Hydrometeor.*, **2**, 122–139, https://doi.org/10.1175/1525-7541(2001)002<0122:ABCSRP>2.0.CO;2.

——, J. R. Metcalfe, B. E. Goodison, and E. Mekis, 1993: "True snowfall": An evaluation of the double fence intercomparison reference gauge. *Proc. 50th Eastern Snow Conf./61st Western Snow Conf.*, Quebec City, QC, Canada, Western Snow Conference, 105–111, https://westernsnowconference.org/sites/westernsnowconference.org/PDFs/1993Yang.pdf.

——, B. E. Goodison, C. S. Benson, and S. Ishida, 1998: Adjustment of daily precipitation at 10 climate stations in Alaska: Application of World Meteorological Organization intercomparison results. *Water Resour. Res.*, **34**, 241–256, https://doi.org/10.1029/97WR02681.

——, S. Ishida, B. E. Goodison, and T. Gunther, 1999: Bias correction of daily precipitation for Greenland. *J. Geophys. Res.*, **104**, 6171–6181, https://doi.org/10.1029/1998JD200110.

——, D. Kane, Z. Zhang, D. Legates, and B. Goodison, 2005: Bias corrections of long-term (1973–2004) daily precipitation data over the northern regions. *Geophys. Res. Lett.*, **32**, L19501, https://doi.org/10.1029/2005GL024057.

Yao, N., Y. Li, N. Li, D. Yang, and O. O. Ayantobo, 2018: Bias correction of precipitation data and its effects on aridity and drought assessment in China over 1961–2015. *Sci. Total Environ.*, **639**, 1015–1027, https://doi.org/10.1016/j.scitotenv.2018.05.243.

Ye, B., D. Yang, Y. Ding, T. Han, and T. Koike, 2004: A bias-corrected precipitation climatology for China. *J. Hydrometeor.*, **5**, 1147–1160, https://doi.org/10.1175/JHM-366.1.

——, ——, and L. Ma, 2012: Effect of precipitation bias correction on water budget calculation in upper Yellow River China. *Environ. Res. Lett.*, **7**, 025201, https://doi.org/10.1088/1748-9326/7/2/025201.

Zhang, Y., T. Ohata, D. Yang, and G. Davaa, 2004: Bias correction of daily precipitation measurements for Mongolia. *Hydrol. Processes*, **18**, 2991–3005, https://doi.org/10.1002/hyp.5745.

Zhang, Y., Y. Ren, G. Ren, and G. Wang, 2020: Precipitation trends over mainland China from 1961–2016 after removal of measurement biases. *J. Geophys. Res. Atmos.*, **125**, e2019JD031728, https://doi.org/10.1029/2019JD031728.

Unauthenticated | Downloaded 02/06/23 08:00 AM UTC