Evaluation of Catch Efficiency Transfer Functions for Unshielded and Single-Alter-Shielded Solid Precipitation Measurements

AMANDINE PIERRE AND SYLVAIN JUTRAS

Faculté de Foresterie, de Géographie et de Géomatique, Université Laval, Quebec City, Quebec, Canada

CRAIG SMITH

Climate Research Division, Environment and Climate Change Canada, Saskatoon, Saskatchewan, Canada

JOHN KOCHENDORFER

Atmospheric Turbulence and Diffusion Division, NOAA/Air Resources Laboratory, Oak Ridge, Tennessee

VINCENT FORTIN

Environmental Numerical Weather Prediction Research, Meteorological Research Division, Environment and Climate Change Canada, Dorval, Quebec, Canada

FRANÇOIS ANCTIL

Faculté des Sciences et de Génie, Département de Génie Civil et de Génie des Eaux, Université Laval, Quebec City, Quebec, Canada

(Manuscript received 9 July 2018, in final form 5 March 2019)

ABSTRACT

Solid precipitation undercatch can reach 20%–70% depending on meteorological conditions, the precipitation gauge, and the wind shield used. Five catch efficiency transfer functions were selected from the literature to adjust undercatch from unshielded and single-Alter-shielded precipitation gauges for different accumulation periods. The parameters from these equations were calibrated using data from 11 sites with a WMO-approved reference measurement. This paper presents an evaluation of these transfer functions using data from the Neige site, which is located in the eastern Canadian boreal climate zone and was not used to derive any of the transfer functions available for evaluation. Solid precipitation measured at the Neige site was underestimated by 34% and 21% when compared with a manual reference precipitation measurement for unshielded and single-Alter-shielded gauges, respectively. Catch efficiency transfer functions were used to adjust these solid precipitation measurements, but all equations overestimated amounts of solid precipitation by 2%–26%. Five different statistics evaluated the accuracy of the adjustments and the variance of the results. Regardless of the adjustment applied, the catch efficiency for the unshielded gauge increased after the adjustment. However, this was not the case for the single-Alter-shielded gauges, for which the improvement of the results after applying the adjustments was not seen in all of the statistics tests. The results also showed that using calibrated parameters on datasets with similar site-specific characteristics, such as the mean wind speed during precipitation and the regional climate, could guide the choice of adjustment methods. These results highlight the complexity of solid precipitation adjustments.

1. Introduction

Uncertainty related to hydrometeorological observations is pervasive in the Canadian subarctic or boreal climate zones, where solid precipitation may account for up to 55% of the total precipitation (Barbier et al. 2009; Oreiller 2013). Difficulties arise from a heterogeneous spatial distribution of solid precipitation within the same geographic area, which results from local variability due to the wind and turbulence (Rasmussen et al. 2012; Sevruk et al. 1989) or from other storm characteristics

Corresponding author: Amandine Pierre, amandine.pierre.1@ulaval.ca

DOI: 10.1175/JTECH-D-18-0112.1

typical of cyclonic, convective, and orographic events (McKay 1968). The catch efficiency of snow gauges is influenced by the size and shape of hydrometeors (Thériault et al. 2012), as well as the shape of the gauge and its orifice (Sevruk et al. 1989), the surface materials constituting the apparatus (Devine and Mekis 2008), and the wind shield selection (Rasmussen et al. 2012; Metcalfe and Goodison 1993). The impact of undercatch is especially large in boreal regions, where the winter season may last up to six months (Nalder and Wein 1998) and snowfall is typically high (Jones and Pomeroy 2001), often exceeding depths of 120 cm (Plamondon et al. 1984).

Since the early 2000s, many national climate monitoring organizations have transitioned from manual to automatic precipitation gauges such as the Geonor T-200B (Geonor, Inc.) or OTT Pluvio^2 (OTT Hydromet GmbH), deployed either unshielded or within a single Alter shield (Smith 2009). It has been reported that these devices, like manual precipitation gauges, can underestimate solid precipitation by more than 50% during windy conditions (Goodison et al. 1998; Kochendorfer et al. 2017; Rasmussen et al. 2012; Smith 2009) when compared with a reference gauge such as the double fence intercomparison reference (DFIR) (Metcalfe and Goodison 1993; Yang 2014; Nitu and Wong 2010). The performance of a gauge configuration can be assessed by calculating its catch efficiency (CE), which is defined as the ratio of its accumulated precipitation divided by the accumulated precipitation of the reference over a specified period. It allows for the development of transfer functions that can be used to compensate for undercatch (Goodison et al. 1998; Kochendorfer et al. 2017; Smith 2009).

In 2010, the World Meteorological Organization (WMO) initiated an observation network to document undercatch. The objectives of the WMO Solid Precipitation Intercomparison Experiment (SPICE) were to conduct intercomparisons between automatic precipitation gauges, and ultimately make recommendations for their use and best practices (Nitu et al. 2012; Goodison et al. 1998). Environment and Climate Change Canada (ECCC) and its collaborators also deployed their own network of intercomparison sites called Canadian Solid Precipitation Intercomparison Experiment (C-SPICE). The Neige site is part of C-SPICE and was established in 2014 in the Montmorency Forest, Quebec, Canada, which has a boreal climate. The Neige site hosts about 30 manual and automatic snow-measuring devices, including a DFIR, an unshielded OTT Pluvio^2, two single-Alter-shielded Geonor T-200B, and two single-Alter-shielded OTT Pluvio^2.

Five different catch efficiency transfer functions were evaluated in this study. They were proposed by Goodison et al. (1998), Smith (2009), and Kochendorfer et al. (2017), the latter group assessing the WMO-SPICE unshielded and single-Alter-shielded gauge measurements at daily, 12-h, and 0.5-h time steps. These transfer functions are identified in this document as "equations" G1 24H, S2 12H, K3 12H, K3 0.5H, and K4 0.5H, respectively (the first letter refers to the name of the author, the first number refers to the chronological order in the publication, and the second part refers to the accumulation length in hours). Observations from the Neige site were not used to develop any of the abovementioned transfer functions, so Neige observations can be used to independently verify the available transfer functions. The objective of this work is to evaluate the various transfer functions applied to twice-daily (0800 and 1800 local standard time) and hourly unshielded and single-Alter-shielded automatic measurements based on the precipitation gauges available at the Neige site, for snow events only. The results of this evaluation are thus dependent on the climatic characteristics of the study site. Furthermore, for some of the equations, parameters were recalibrated to the Neige site dataset to help to evaluate their impact on the resultant adjustments.

2. Material and methods

a. *The Neige site*

The Montmorency Forest is a 412-km^2 public forest that has been dedicated to Université Laval's teaching and research activities since 1965. Located 80 km north of Québec City, this balsam fir–dominated ecosystem is representative of eastern Canada (Jones and Pomeroy 2001). According to the Köppen–Geiger system, its climate is classified as continental subarctic (Kottek et al. 2006; Rubel et al. 2017). It hosts several meteorological and hydrological stations, some of which have been reporting continuously since 1965. This is the case of meteorological station 7042388, which was promoted to an ECCC Reference Climate Station (RCS) 7042395 in 2003. The 1981–2010 climate normals were 1583 mm of total precipitation and 620 mm (41%) of solid precipitation.

The Neige site (47°19′20.15″N, 71°9′4.11″W) includes the ECCC meteorological station, two provincial meteorological stations (7042388 and 7042389), and a C-SPICE intercomparison site. It hosts about 30 manual and automatic meteorological sensors that monitor precipitation, wind velocity, air temperature, radiation, snow water equivalent, and hydrometeor phase (Fig. 1).

Two single-Alter-shielded Geonor T-200B gauges, two single-Alter-shielded OTT Pluvio^2 gauges, and one unshielded OTT Pluvio^2 gauge were operated at the

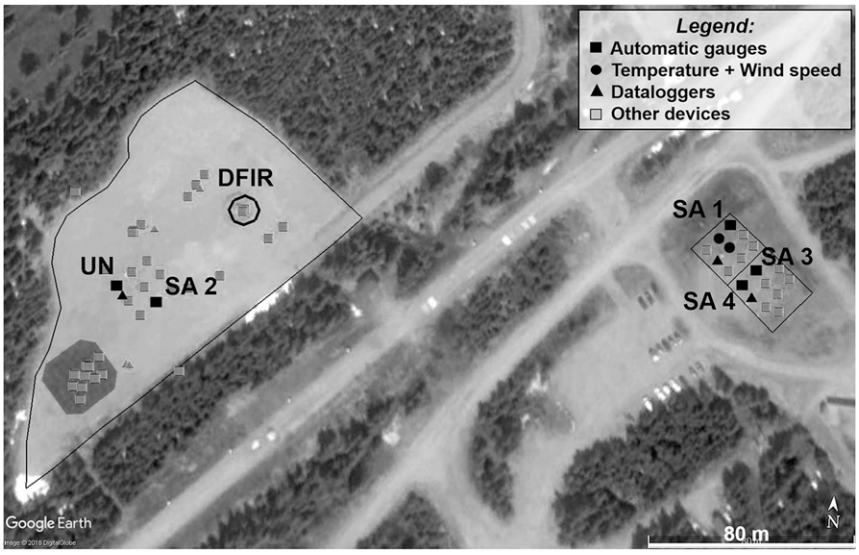

FIG. 1. Map of the Neige site and the locations of the automatic and manual devices used this study (black) and other devices (gray).

site for up to four winters (Fig. 1; Table 1). All of these gauges were unheated. Geonor T-200B are automatic weighing gauges with three vibrating wire transducers and a 200-cm^2 inlet catch area. Output from all three transducers was averaged together to estimate the bucket weight (converted to depth) with a precision of 0.1 mm (Smith 2009). OTT Pluvio^2 automatic gauges (Nitu and Wong 2010) also measure the mass of precipitation falling into a bucket. They have a catch area of 200 cm^2, 0.01-mm resolution, and a maximum 6-s output frequency. Reference observations were taken using an H&H 90 manual weighing precipitation gauge (182.4-cm^2 orifice), sheltered by a DFIR (Fig. 2). A DFIR is a combination of a Tretyakov shield surrounded by two octagonal fences of 4 and 12 m in diameter, made of wood lath spaced to create a porosity of 50% (Metcalfe and Goodison 1993; Yang 2014).

Horizontal wind speed and direction were recorded by a propeller anemometer (05103 Wind Monitor; R. M. Young Co.) installed at a height of 2.5 m. Data were averaged each 1 h from the sixty 1-min cumulative values. Their accuracies were $\pm 0.3\,\mathrm{m\,s^{-1}}$ and ±3°, respectively. Relative humidity and temperature were measured hourly with a Campbell Scientific, Inc., HMP45C installed inside a Stevenson shelter to limit radiation errors from direct sunlight (Henshall and Snelgar 1989). Their accuracies were ±1% and ±0.1°C, respectively.

b. *Transfer functions*

The transfer functions from the literature were developed and calibrated over different accumulation periods than those observed at the Neige site. These transfer functions have been developed using 0.5- (semihourly), 12- (bidaily), and 24-h (daily) measured accumulations, which differ from the 10–14-h (semidaily or twice daily) manual observations and 1-h automatic records made at the Neige site.

1) EQUATION G1 24H, FROM GOODISON ET AL. (1998)

The Goodison et al. (1998) results were based on U.S. standard 8-in. (1 in. = 2.54 cm) nonrecording gauges,

TABLE 1. Configuration of the precipitation gauges installed on the Neige site.

Name	Gauge type	Shield type	Measurement	No. of twice-daily obs			
				2013/14	2014/15	2015/16	2016/17
DFIR	H&H 90	DFIR	Semidaily manual	137	286	277	276
SA1	Geonor T-200B	Single Alter	Hourly automatic	138	275	246	246
SA2	Geonor T-200B	Single Alter	Hourly automatic	0	0	11	246
SA3	OTT Pluvio^2	Single Alter	Hourly automatic	0	283	239	276
SA4	OTT Pluvio^2	Single Alter	Hourly automatic	0	283	239	269
UN	OTT Pluvio^2	None	Hourly automatic	0	0	86	250

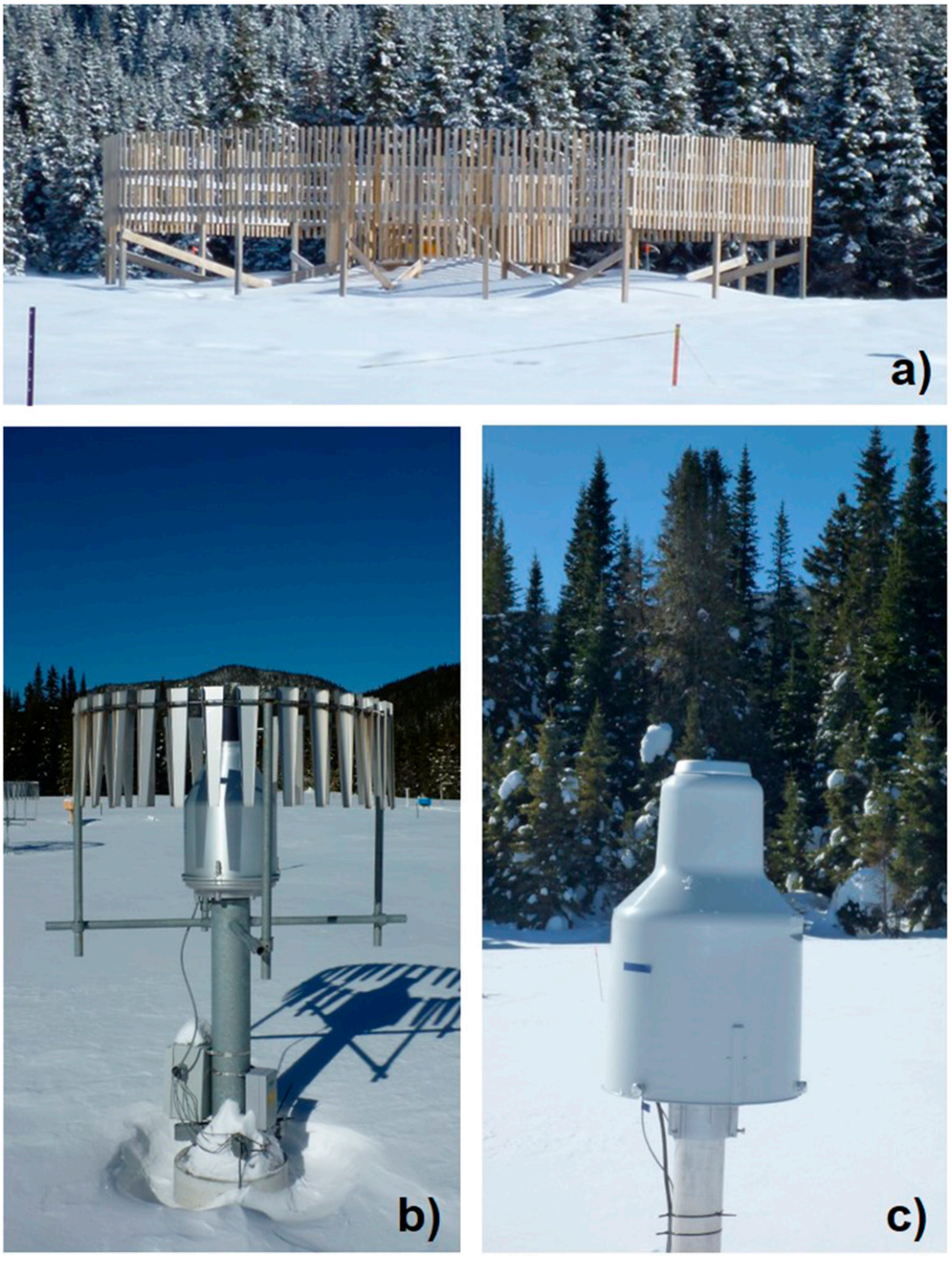

FIG. 2. Photographs of (a) DFIR and Tretyakov shield fencing around an H&H 90 manual weighing gauge, (b) a single-Alter-shielded Geonor T-200B, and (c) an unshielded OTT Pluvio^2 (taken by author A. Pierre on 17 Feb 2017).

also known as NWS 8-in. standard gauges, deployed either without a wind shield or with a single-Alter shield at Valdai, Russia; Reynolds Creek, Idaho; and Danville, Vermont. Daily precipitation measurements were recorded from 1986 to 1993 (Table 2). Reference observations at each site were taken using the same manual gauge positioned inside a DFIR shield. A sigmoidal shaped transfer function (G1 24H) was used:

$$\mathrm{CE} = \exp[a - b(U)^c]/100, \quad (1)$$

TABLE 2. Geographic, climatic, and meteorological characteristics of the study sites. Here, U_{gh} is the wind speed at gauge height.

Site	Country	Elev (m)	Lat	Köppen–Geiger classification	Mean U_{gh} ($m\,s^{-1}$)	Mean T_{air} (°C)	Annual total P (mm)	Annual solid P (mm)
CARE[a]	Canada	251	44°14′N	Boreal	3.2	−3.3	636	430
Haukeliseter[a]	Norway	991	59°49′N	Boreal	6.7	−1.7	2432	594
Sodankylä[a]	Finland	179	67°22′N	Boreal	1.6	−2.1	—	527
Caribou Creek[a]	Canada	519	53°57′N	Boreal	2.6	−6.3	365	252
Weissfluhjoch[a]	Switzerland	2537	46°50′N	Polar	3.8	−7.2	1146	586
Formigal[a]	Spain	1800	42°46′N	Temperate	2.3	−0.7	—	403
Marshall[a]	United States	1742	39°57′N	Arid	2.8	−2.0	560	229
Bratt's Lake[a]	Canada	585	50°12′N	Boreal	4.4	−1.5	390	206
Valdaï[b]	Russia	194	57°59′N	Boreal	3.8	−4.1[c]	1256	357
Reynolds Creek[b]	United States	1193	43°12′N	Arid	2.5	−2.0	371	87
Danville[b]	United States	552	44°29′N	Boreal	1.5	−6.9	1993	1051
Neige	Canada	640	47°19′N	Boreal	2.0	−4.2	1583	620

[a] Data for temperature and wind speed during solid precipitation events were provided by Kochendorfer et al. (2017), and precipitation data were provided by J. Kochendorfer, (unpublished data) of the NOAA Air Resources Laboratory and C. Smith (unpublished data) of Environment and Climate Change Canada.
[b] Data for temperature and wind speed during solid precipitation events and precipitation data were provided by Goodison et al. (1998).
[c] Only maximum temperature measurements were available.

where parameters a, b, and c are the height coefficient of the curve, the inflection point value, and the slope from the sigmoidal function, respectively; U is the daily mean wind speed at gauge height ($m\,s^{-1}$).

2) EQUATION S2 12H, FROM SMITH (2009)

Smith (2009) employed an unheated Geonor T-200B with a single Alter shield deployed at the Bratt's Lake site from 2003 to 2006 (Table 2), which collected data at 15-min time intervals. Manual reference observations were taken daily or twice daily using a DFIR.

Smith (2009) opted for the following exponential transfer function (S2 12H):

$$\mathrm{CE} = \exp[-a(U)], \tag{2}$$

where parameter a is the value of the tangent to the y axis of the decreasing exponential function and U is the mean wind speed at gauge height ($m\,s^{-1}$) over the period of accumulation.

3) EQUATIONS K3 0.5H, K3 12H, AND K4 0.5H, FROM KOCHENDORFER ET AL. (2017)

Kochendorfer et al. (2017) used both heated Geonor T-200B and orifice-heated OTT Pluvio^2 gauges, deployed without wind shields and with single Alter shields at the Centre for Atmospheric Research Experiments (CARE), Ontario, Canada; Haukeliseter, Norway; Sodankylä, Finland; Caribou Creek, Saskatchewan, Canada; Weissfluhjoch, Switzerland; Formigal, Spain; Marshall, Colorado; and Bratt's Lake, Saskatchewan, Canada, from 2013 to 2015 (Table 2). Measurements were recorded every 0.5 h. Reference observations at each site were taken using Geonor T-200B or OTT Pluvio^2 gauges positioned inside an adapted DFIR. This automated reference is called a Double Fence Automatic Reference (DFAR). It is equipped with the same double fences as the DFIR, but a single Alter shield replaces the Tretyakov shield around the collector of an automated gauge rather than a manual gauge.

Kochendorfer et al. (2017) fitted the data collected during SPICE to two transfer function models. The first transfer function was sigmoidal with respect to air temperature and exponential with respect to wind speed at gauge height (K3 0.5H and K3 12H):

$$\mathrm{CE} = \exp(-a(U)\{1 - \mathrm{atan}[b(T_{air})] + c\}), \tag{3}$$

where parameters a, b, and c are the height coefficient of the curve, the inflection point value, and the slope from the sigmoidal function, respectively; U is the wind speed at gauge height ($m\,s^{-1}$); and T_{air} is the air temperature (°C) during the 0.5-h event.

The second transfer function excluded the air temperature (K4 0.5H):

$$\mathrm{CE} = a \times \exp[-b(U)] + c, \tag{4}$$

where parameters a, b, and c are respectively the tangent to the ordinate axis, the slope, and the tangent to the abscissa axis of the exponential function; U is the wind speed at gauge height ($m\,s^{-1}$) during the 0.5-h event. While equation K3 0.5H integrates a temperature parameter, it does not require the user to differentiate precipitation phase. Equation K4 0.5H requires the user to determine phase and select the appropriate parameters accordingly. Kochendorfer et al. (2017) used temperature thresholds to determine phase.

TABLE 3. Original and recalibrated parameter values a, b, and c of the five equations used in this study for unshielded (UN) and single-Alter-shielded (SA) instruments. Source of the original parameters: Kochendorfer et al. (2017), Smith (2009), and Goodison et al. (1998).

			a		b		c	
Equation	Shield	Equation form	Original	Recalibrated	Original	Recalibrated	Original	Recalibrated
G1 24H	UN	$CE = \exp[a - b(U)^c]/100$	4.61	4.26	0.16	0.03	1.28	1.70
G1 24H	SA	$CE = \exp[a - b(U)^c]/100$	4.61	4.41	0.04	0.01	1.75	1.72
S2 12H	SA	$CE = \exp[-a(U)]$	−0.20	0.09	—	—	—	—
K3 0.5H	UN	$CE = \exp(-a(U)\{1 - \text{atan}[b(T_{air})] + c\})$	0.08	—	0.73	—	0.41	—
K3 12H	UN	$CE = \exp(-a(U)\{1 - \text{atan}[b(T_{air})] + c\})$	0.11	18.8	0.34	0.00	0.26	-0.99
K3 0.5H	SA	$CE = \exp(-a(U)\{1 - \text{atan}[b(T_{air})] + c\})$	0.03	—	1.37	—	0.78	—
K3 12H	SA	$CE = \exp(-a(U)\{1 - \text{atan}[b(T_{air})] + c\})$	0.05	−0.13	0.97	−0.04	0.84	−1.42
K4 0.5H	UN	$CE = a \times \exp[-b(U)] + c$	0.86	—	0.37	—	0.23	—
K4 0.5H	SA	$CE = a \times \exp[-b(U)] + c$	0.73	—	0.23	—	0.34	—

4) CATCH EFFICIENCY TRANSFER FUNCTIONS PARAMETER VALUES

The appropriate unshielded and single-Alter-shielded parameter values from each of the abovementioned studies are given in Table 3.

c. *Geographic, climatic, and meteorological characteristics of the studied sites*

Catch efficiency transfer functions from past WMO intercomparisons were either based on eight sites operated during two winter seasons (Kochendorfer et al. 2017), one site during three winter seasons (Smith 2009), or three sites during from three to six winter seasons (Goodison et al. 1998) (Table 2; Fig. 3).

The Köppen–Geiger climate classification (Kottek et al. 2006; Rubel et al. 2017) describes the regional climatic conditions for each of the study sites (Fig. 3). The dominant climate type studied by Goodison et al. (1998) and Smith (2009) was boreal, with 86% and 100% of their data found in this class, respectively. The database used by Kochendorfer et al. (2017) included a wider range of climates: 56% boreal, 18% temperate, 14% polar, and 12% arid. The mean gauge-height wind speed ranged from 1.5 to 6.7 m s^{-1}, the mean winter air temperature ranged from −0.7° to −7.2°C, the total annual precipitation ranged from 365 to 1993 mm, and the total solid precipitation ranged from 87 to 1051 mm. Three virtual sites were created to regroup and illustrate the generic climatic characteristics of the three datasets. Each virtual site reflected the climate repartition of each dataset from the literature. A climate type was estimated for every dataset by weight averaging the climatic characteristics from their respective sites. A principal component analysis (PCA) was performed (Fig. 4) using the climatic characteristics. PCA is a geometrical and statistical tool used to determine the organization of a dataset based on variables chosen by the user (Jolliffe 2011; Cornillon et al. 2012). It employs an orthogonal transformation to convert a set of observations (i.e., data) of possibly correlated variables into a set of linearly uncorrelated variables called principal components. Principal components constitute the axis of the individual and variable factor maps. The first two components are the two main uncorrelated variables resulting from the transformation of the correlated variables. In our study, the PCA variables were climate class (Kottek et al. 2006), mean air temperature [mean T_{air} (°C)], annual amount of solid precipitation [P (mm)], and mean wind speed [mean U_{gh} (m s^{-1})] from all of the sites and

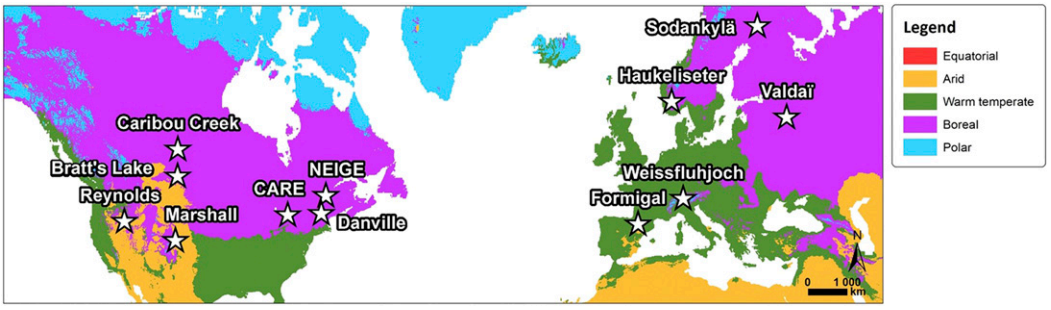

FIG. 3. Map of the SPICE sites used in this study, overlaying the Köppen–Geiger climate classifications (Kottek et al. 2006; Rubel et al. 2017).

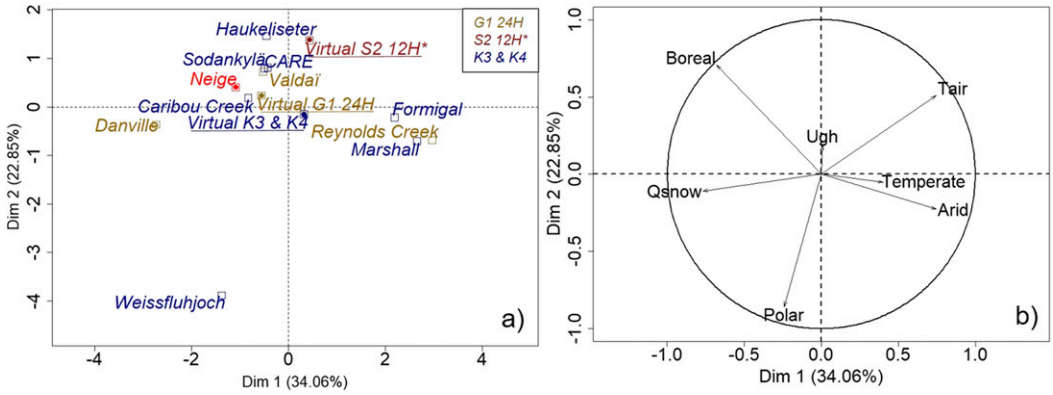

FIG. 4. (a) Individual factor map (PCA) from all of the study sites and the three virtual sites of G1 24H, S2 12H, and K3 and K4. Virtual S2 12H: only Bratt's Lake site for the climate (Kottek et al. 2006), mean T_{air} (°C), annual solid P (mm), and mean U_{gh} (m s^{-1}) variables. (b) Variable factor map (PCA) from all of the variables used to build the PCA.

virtual sites. The geometric proximities of each site with regard to the chosen variable were illustrated by an individual factor map that exposed the structural relationships between the variables and the components (Fig. 4).

d. Precipitation measurements

Manual gauges were observed twice daily [0800 and 1800 LT (UTC − 5 h) during eastern standard time and 0900 and 1900 LT (UTC − 4 h) during eastern daylight (saving) time] by a dedicated technician, from 1 November to mid-April. The H&H 90 containers and the snow within them were weighed on a 10^{-3} g-precision scale (Ohaus Explorer, version 2), resulting in a precision of 0.001 mm. To avoid capping, the technician removed the snow covering all manual and automatic gauges as necessary. The OTT Pluvio^2 gauges were partially prefilled with an antifreeze liquid composed of equal parts of ethanol and 1,2-propanediol and 5%–7% of the total of the mix of kerosene to prevent evaporation. Total bucket weight was recorded hourly and converted into millimeters of liquid water equivalent. The Geonor T-200B were also partially prefilled with the same antifreeze liquid and measured at hourly intervals. Since many authors have merged observations provided by gauge replicates (Goodison et al. 1998; Kochendorfer et al. 2017; Smith 2009), coefficients of correlation were calculated between the four single-Alter gauges at the Neige site (SA1–SA4; see Table 4).

e. Selection of twice-daily solid precipitation events

Reverdin et al. (2016) developed an algorithm to identify and select precipitation events and avoid signal noise due to dysfunctions of the automatic gauges. This was necessary because automatic precipitation gauges can falsely record both positive and negative measurements as a function of the sensitivity of the device (from 0.1 or 0.01 mm of precipitation). Therefore, events that are too small to be used with confidence must be excluded from some analyses (Reverdin et al. 2016; Wolff et al. 2015). The Reverdin et al. (2016) algorithm was applied by Kochendorfer et al. (2017). It was also adapted as follows in this study.

There were three selection criteria applied to each twice-daily precipitation measurement. First, precipitation-phase observations made by the technicians were used to select only solid precipitation events. Second, the measured precipitation must have been greater than 1 mm as measured by the DFIR. Third, a minimum precipitation threshold for the automatic measurement as used by Kochendorfer et al. (2017) at 0.5-h intervals was adapted here to the hourly time step: 0.22 mm for all single-Alter-shielded instruments and 0.12 mm for unshielded instruments. This third step relates to the inherent accuracy of hourly automatic measurements made during the twice-daily manual observations. If any one of the automated gauges measured more precipitation than their respective threshold during an hour in each twice-daily event, then

TABLE 4. Results of intercomparison between twice-daily precipitation events measured by single-Alter-shielded gauges on the Neige site.

Instrument 1	Instrument 2	Slope	RMSE (mm)	r	No. of obs
SA2 (Geonor)	SA1 (Geonor)	1.05	2.0	0.89	55
SA3 (OTT)	SA1 (Geonor)	0.98	2.1	0.89	142
SA4 (OTT)	SA1 (Geonor)	0.98	2.0	0.90	138
SA3 (OTT)	SA2 (Geonor)	1.04	1.1	0.97	66
SA4 (OTT)	SA2 (Geonor)	1.01	1.0	0.98	61
SA3 (OTT)	SA4 (OTT)	1.02	0.4	1.00	154

TABLE 5. Error statistics for transfer functions from the literature and recalibrated transfer functions. The plus sign, minus sign, and equal sign indicate that the criterion is improved, worse, or the same after the use of the transfer function relative to the unadjusted data. Here, |P bias| is the absolute P bias.

		Accuracy			Variance		Bias
Equation	Shield	RMSE	MAE	PE ± 10% DFIR	RMSE − MAE	r	|P bias|
G1 24H recalibrated	UN	+	+	+	+	+	+
K3 12H recalibrated	UN	+	+	+	+	+	+
K4 0.5H	UN	+	+	+	+	+	+
K3 0.5H	UN	+	+	+	+	+	+
G1 24H	UN	+	+	+	+	+	+
K3 12H	UN	+	+	+	=	+	+
G1 24H recalibrated	SA	+	+	+	+	=	+
K3 12H recalibrated	SA	+	+	+	−	−	+
S2 12H recalibrated	SA	+	+	+	+	−	+
K4 0.5H	SA	−	+	+	−	−	+
K3 0.5H	SA	−	+	+	−	−	+
G1 24H	SA	−	=	+	−	−	+
K3 12H	SA	−	−	+	−	−	+
S2 12H	SA	−	−	−	−	−	−

each twice-daily event was included in the analysis. Therefore, even when the DFIR measured an event greater than 1 mm, if the automatic gauges measured less than their respective thresholds, the event was excluded from the analysis.

Other meteorological measurements, such as air temperature, humidity, and wind speed, were averaged to correspond with the hourly and twice-daily precipitation time steps.

f. Evaluation procedure

The reference data provided by the DFIR at the Neige site were recorded twice daily. The twice-daily time step was therefore used to compare all unadjusted and adjusted observations. Adjustments derived using 0.5-h data (Table 3) were applied to hourly automatic precipitation observations, using hourly meteorological values (U_{gh} and T_{air}) in the equations. The adjusted precipitation quantities were accumulated at the twice-daily time step. In parallel, all hourly automatic gauge raw observations were accumulated, and meteorological values were averaged according to the twice-daily time step. Then the adjustment using twice-daily, 12-h, and 24-h time-step parameters (Table 3) was applied to each twice-daily accumulated precipitation observation using the twice-daily averaged meteorological values (U_{gh} and T_{air}). The DFIR measurement was adjusted only for evaluation of the results for equation S2 12H to be consistent with Yang et al. (1993).

Five different statistics were chosen to qualify (Table 5) and quantify (Fig. 7, described in more detail below) the impact of the adjustments. The root-mean-square error (RMSE) and mean absolute error (MAE) (Chai and Draxler 2014; Willmott and Matsuura 2005) are indicators that describe the average magnitude of the error in a set of forecasts (e.g., accuracy). RMSE is nonlinear and based on the square of the error, which gives greater weight to larger errors than the MAE. RMSE is always equal or higher than the MAE, and the difference between the RMSE and the MAE (RMSE − MAE) can be used to describe the variance of the errors. RMSE and MAE can range from 0 to $+\infty$, and smaller values indicate smaller errors.

The percentage of bias (P bias) was calculated as the proportional difference between unadjusted or adjusted precipitation and the DFIR measurement. A value close to 0% would show that, on average, observed or calculated precipitation was the same as the amount measured by the DFIR reference. Positive P-bias values indicate an overestimation of the reference quantity, while negative values indicate an underestimation. The Pearson correlation coefficient r is a linear coefficient that illustrates the association between two variables (Hauke and Kossowski 2011; Rodgers and Nicewander 1988). The closer the value is to 1, the better the variables are correlated. The percentage of events that were within ±10% of the DFIR quantity (called PE ± 10% DFIR) were also calculated. This indicator expressed, in another way, the accuracy of the observed or calculated solid precipitation quantity.

g. Recalibration of the parameters

Using the Neige dataset, recalibrations of the parameters of some equations were performed following the respective methods described in each study (Goodison et al. 1998; Kochendorfer et al. 2017; Smith 2009). The time step of the Neige reference dataset was twice daily, so K3 0.5H and K4 0.5H (Kochendorfer et al. 2017) could

not be recalibrated. For the same reason, the unshielded transfer function from Smith (2009) was not calibrated for the Neige site.

3. Results

a. Event selection

Overall, 980 twice-daily observations were recorded at the Neige site over four winter seasons (2014–17). Forty-seven percent of these events were snow events, 43% were nonprecipitation, and 8% were rain events. From these events, the algorithm described in section 2e selected 200 snow events for the derivation and testing of transfer functions. The selected events represent 80% of the total precipitation that occurred during the study.

b. PCA results

Climate PCA results suggested that approximately 60% of the variance could be explained by the first two components (Fig. 4). The variable factor map indicates that among climate variables, air temperature and snow volume contributed more to the main components than mean wind speed (Fig. 4b).

On the individual factor map, the Caribou Creek site was the closest to the Neige site regarding the chosen variables. The location of the virtual sites on the individual factor map showed that the virtual site most similar to the Neige site was the G1 24H virtual site, followed by the K3 and K4 virtual sites. The S2 12H virtual site (Bratt's Lake) was the most distinct from Neige (Fig. 4a).

c. Gauge data correlations

Kochendorfer et al. (2017) merged Geonor T-200B and OTT Pluvio^2 data, and demonstrated that undercatch depends more on the shield configuration and site conditions than on the type of gauge. Ryu et al. (2012) showed that the correlation between the data of two different instruments (Geonor T-200B and OTT Pluvio^2) equipped with a DFIR depended on the threshold of the minimum precipitation considered for comparison, the type of precipitation, the temperature, the relative humidity, and the heating of the gauge rim. In our study, the correlation between the observed twice-daily selected precipitation events provided by the four single-Alter-shielded gauges was evaluated: the slope, the RMSE, and r were calculated for all single-Alter-shielded instrument precipitation datasets (Table 4).

Regardless of the statistical comparison, a preliminary analysis indicated no significant measurement differences between all the gauges in the present study. The regression coefficients, or slopes, of each linear regression by pair were close to 1, showing values varying between 0.98 and 1.05. However, the highest RMSE (>2 mm) combined with the lowest r (<0.90) corresponded to the SA1 (Geonor) instrument. All other pairs revealed lower RMSEs along with d r values closer to 1. Kochendorfer et al. (2017) showed slopes ranging from 0.997 to 1.01 and RMSE ranging from 1.70 to 2.04 mm. The correlation between all the instruments was considered acceptable, based on the variability presented by Ryu et al. (2012). Therefore, for the remainder of this study, all observations provided by the four single-Alter-shielded instruments were evaluated as a single source of data, following Kochendorfer et al. (2017).

d. Evaluation of the catch efficiency transfer functions at the Neige site

The CE of the 200 selected twice-daily precipitation events was calculated for unshielded (Fig. 5) and single-Alter-shielded (Fig. 6) precipitation gauges. In addition, the CE transfer functions provided in the literature were plotted with the Neige site CE and wind speed measurements (Figs. 5 and 6). Temperature classes were separated according to thresholds suggested by Dubé (2003), which were based on the different crystalline types of snow.

Equations S2 12H, K3 12H, K3 0.5H, and K4 0.5H decrease exponentially, unlike the sigmoid equation G1 24H shape, which shows a slight plateau for low wind speeds before the catch efficiency decreases (Figs. 5 and 6). Therefore, two basic forms of CE curves were evaluated in this study. It appears that there was no temperature dependency in the CE observed at the Neige dataset, as shown by the different shades of the data and the K3 curves.

After all equations were applied to the observations, statistical analyses were used to compare unadjusted and adjusted datasets with the reference precipitation (Fig. 7; Table 5).

When applied to the unshielded measurements, equation G1 24H, provided by Goodison et al. (1998), led to an increase in accuracy, a decrease in both RMSE and MAE (−30% and −40%) without an increase in the variance of the results (RSME − MAE doubled). When applied to the single-Alter-shielded instruments, the adjustment did not improve measurement accuracy; MAE remained the same, but the variance of the results increased, with the RMSE − MAE being 3 times the RMSE − MAE of the unadjusted measurements. The P bias became positive and was reduced by a factor of 6 and 10 for unshielded and single-Alter-shielded measurements, respectively. The r was closer to 1 (from 0.91 to 0.95) for the unshielded instrument data. In contrast to these unshielded results, for the single-Alter measurements r diverged further from 1 (from 0.96 to 0.87)

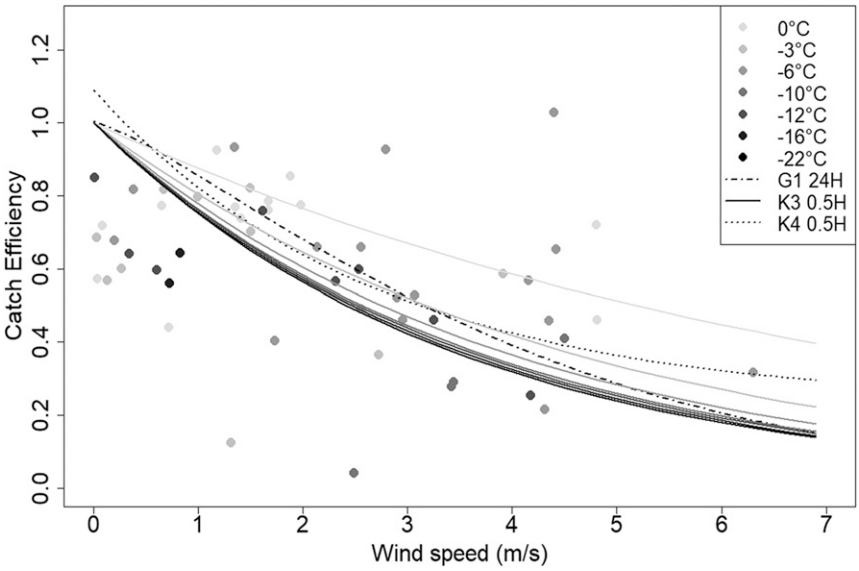

FIG. 5. Catch efficiency transfer functions and CE from selected twice-daily solid precipitation events for the unshielded gauge at various wind speeds. Temperature is represented by different shades of gray. The dot–dashed line is equation G1 24H, the solid lines are equation K3 0.5H, using different gray shades to represent different temperatures, and the dotted line is equation K4 0.5H.

with adjustment. This is consistent with the RMSE and MAE results, which indicated an increase in the variance of the results after applying the G1 24H adjustment to the single-Alter-shielded measurements. The PE ± 10% DFIR, however, was multiplied by 3 and 1.5 for unshielded and single-Alter-shielded gauge data, respectively, after adjustment. This means that more events were close to the DFIR reference, which is contrary to what is shown by some of the other error statistics. In this respect, the adjustments improved the measurements.

Equation S2 12H, provided by Smith (2009), led to a decrease in accuracy, an increase in both RMSE and MAE (around +100% and +60%), and an increase in the variance of the results (RSME-MAE doubled) when applied to single-Alter-shielded measurements. The P bias became positive and higher than the initial P bias (from −21% to +26%). The *r* decreased (from 0.96 to 0.85). The PE ± 10% DFIR remained unchanged after application of equation S2 12H on the data and was the smallest among all of the adjusted PE ± 10% DFIR results.

Equation K3 12H, provided by Kochendorfer et al. (2017), applied to twice-daily unshielded measurements, exhibited an increase in accuracy, a decrease in both RMSE and MAE (by 20% and 30%, respectively), and an increase in the variance of the results (RMSE − MAE increased by 25%). Applied to the single-Alter-shielded measurements, the accuracy increased (RMSE and MAE decreased by 50% and 10%, respectively), and the variance of the results increased (RMSE − MAE doubled). The P bias became positive, and was reduced by a factor of 2.5 and 1.5 for unshielded and single-Alter-shielded twice-daily measurements, respectively. The *r* was closer to 1 (increased from 0.91 to 0.96) for the unshielded twice-daily measurements, whereas it diverged away from 1 (from 0.96 to 0.9) for the adjusted single-Alter-shielded twice-daily gauge data, reflecting a stronger correlation between the DFIR and the adjusted unshielded measurements. The RMSE and MAE variance increased after applying equation K3 12H to the single-Alter-shielded twice-daily instruments data. After adjustment, the PE ± 10% DFIR was multiplied by 3 and 1.5 for unshielded and single-Alter-shielded measurements, respectively, indicating that more events were close to the DFIR reference and the accuracy of the prediction was increased.

Equation K3 0.5H, provided by Kochendorfer et al. (2017), applied to unshielded and single-Alter hourly measurements, resulted in similar results as the application of equation G1 24H, but with better scores on all the statistical criteria. Regarding unshielded measurements, the accuracy increased (RMSE and MAE decreased by 55% and 75%, respectively), with a small decrease in the variance of the results (RSME-MAE decreased by 30%). However, the adjustment did not significantly improve the accuracy of the single-Alter-shielded measurements (RMSE increased by 10%; MAE decreased by 20%). In addition, the RMSE − MAE of the single-Alter-shielded results increased by a factor of

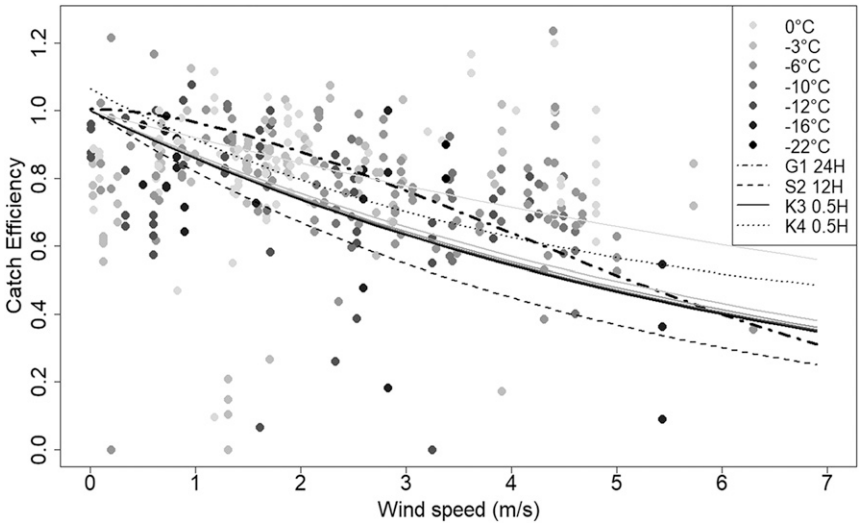

FIG. 6. Similar to Fig. 5, but for single-Alter-shielded instruments at various wind speeds. Here, the dot–dashed line is equation G1 24H, the dashed line is equation S2 12H, the solid lines are equation K3 0.5H, using different gray shades to represent different temperatures, and the dotted line is equation K4 0.5H.

1.5 after adjustment. The P bias was reduced by a factor of 7 and 5 for unshielded and single-Alter-shielded instruments data, respectively. The *r* was closer to 1 (improved from 0.91 to 0.97 with adjustment) for the unshielded instrument data, indicating a stronger association between the variables. On the contrary, this coefficient diverged further from 1 (from 0.96 to 0.92) for the single-Alter-shielded instrument data. The PE ± 10% DFIR was multiplied by 2.5 and 2 for unshielded and single-Alter-shielded gauge data, respectively, after adjustment. More events were close to the DFIR reference, indicating an increase in the prediction accuracy based on this error statistic.

Equation K4 0.5H, provided by Kochendorfer et al. (2017), applied to unshielded hourly measurements, resulted in a decrease in both RMSE and MAE (by 50% and 40%, respectively). It also showed a decrease in the variance of the results (RMSE − MAE decreased by one-third). However, applying equation K4 0.5H actually resulted in a decrease of the accuracy of the single-Alter measurements (RMSE: +15%; MAE: −5%). The variance of the results also increased with this adjustment (RMSE − MAE doubled). The P bias became positive and was reduced by a factor of 8 and 4 for the hourly unshielded and single-Alter-shielded measurements, respectively. For the hourly unshielded data, *r* for the adjusted measurements was closer to 1 (adjusted from 0.91 to 0.98), showing an increased correlation between the variables. However, for the single-Alter-shielded hourly measurements, the correlation decreased after adjustment (from 0.96 to 0.93). The adjustment induced an increase of the PE ± 10% DFIR for both the unshielded and hourly single-Alter-shielded data (by a factor of 2.5 and 1.5, respectively); more adjusted events were close to their corresponding DFIR measurements, indicating an increase in the accuracy of the prediction.

e. Recalibration of catch efficiency transfer functions using the Neige site dataset

Although the objective of this work was not to propose a new set of parameters for the evaluated transfer functions, evaluations of the adjustment made using recalibrated parameters contribute to the interpretation of the primary results (Fig. 7; Table 5).

1) RESULTS OF THE RECALIBRATION OF THE PARAMETERS

After a Neige site-specific recalibration of the parameters for equations G1 24H and K3 12H for unshielded instruments and for equations G1 24H, S2 12H, and K3 12H for single-Alter instruments, the recalibrated parameters of equation G1 24H were closer to the value of the initial parameters than the S2 12H and K3 12H equations (Table 3).

2) RESULTS OF ADJUSTMENT USING THE RECALIBRATED PARAMETERS

Regarding the unshielded gauge data, the RMSE was reduced by 60% (same for MAE) after applying the recalibrated G1 24H and K3 12H equations, and by 15%–30% (by 30%–40% for MAE) for the single-Alter instrument data adjustments (Fig. 7). A smaller absolute P bias was observed for the adjustment using recalibrated G1 24H and

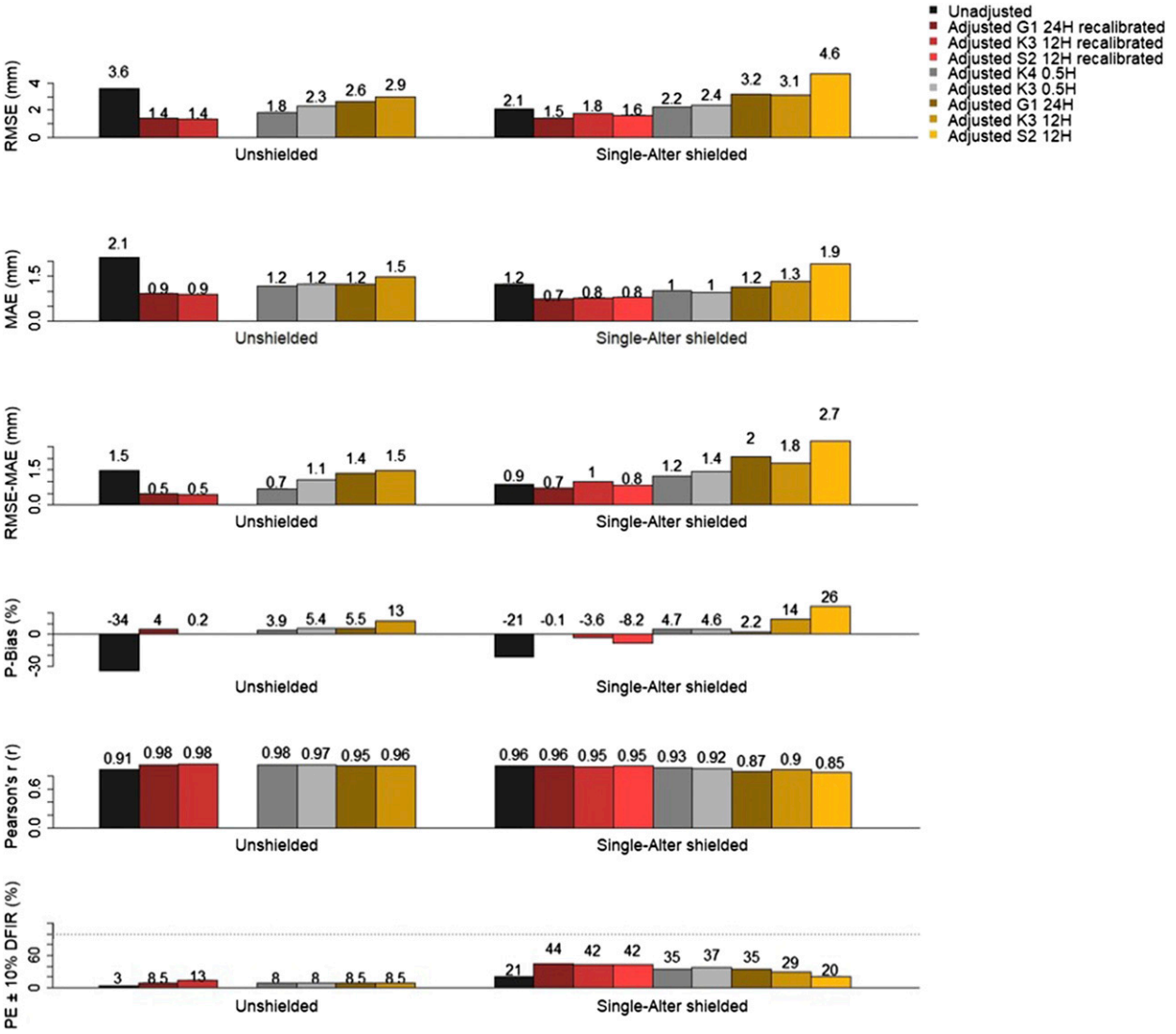

FIG. 7. Comparison of statistics describing unadjusted and adjusted twice-daily solid precipitation events using raw parameters or recalibrated parameters for unshielded, single-Alter-shielded, and DFIR measurements. Shown are RMSE, MAE, RMSE − MAE, P bias, *r*, and PE ± 10% DFIR (the percent of events that were within ±10% of the DFIR reference).

K3 12H equations. For the site-specific adjustments *r* values for unshielded and single-Alter-shielded measurements were 0.98 and 0.96, respectively. For unshielded and single-Alter-shielded gauges, the recalibrated equation K3 12H adjustments increased the PE ± 10% by a factor of 4 and 2, respectively. The highest PE ± 10% DFIR after adjustment was 13% for the unshielded gauge (using recalibrated equation S2 12H) and 44% for the single-Alter-shielded gauges (using recalibrated equation G1 24H) (Fig. 7).

4. Discussion

The Neige site produced about 200 twice-daily snow events for evaluation. In comparison, Goodison et al. (1998) used 55 and 108 daily events to calibrate G1 24H for unshielded and single-Alter-shielded instruments, respectively. Smith (2009) used 21 bidaily or daily data to calibrate equation S2 12H for single-Alter-shielded instruments and Kochendorfer et al. (2017) less than 160 bidaily events to calibrate both unshielded and single-Alter-shielded instruments with equation K3 and K4. This shows the evaluation dataset was large enough if not larger than the calibration datasets, which made the results more trustworthy.

a. Distribution of the raw CE data

Results of the ratio between two relatively small numbers may, of course, lead to noise (Figs. 5 and 6). For

solid precipitation, the precipitation rate is typically much lower than it is for rain, so in order to create representative transfer functions, most of the precipitation measurements available for the derivation of such functions are less than 1.0 mm h^{-1}. At such low rates, random spatial differences in precipitation and other random measurement errors can be significant. Because of this, there is some scatter even between identical or very similar precipitation measurements (Kochendorfer et al. 2017, their Fig. 3; Kochendorfer et al. 2018, their Fig. 7b). Catch efficiency curves have always been plagued by large uncertainties (Wolff et al. 2015). Unfortunately, no practical solution for this problem has been developed.

Another explanation for the scatter in the raw CE data could be related to wind speed fluctuations. At the Neige site, the wind speed was measured at 2.5 m above the ground, rather than 2.5 m above the height of the snowpack. Hultstrand and Fassnacht (2018) demonstrated that the variance and uncertainty related to the wind speed measurements could increase when the anemometer is within 1.5 m of the snowpack (Hultstrand and Fassnacht 2018; their Figs. 2 and 4), such as for the Neige site.

b. *Evaluation of WMO CE transfer functions at the Neige site*

Unadjusted measurements showed that the unshielded instruments had an average CE of about 64% (±19%), which is close to the 66% and 65% that Kochendorfer et al. (2017) and Goodison et al. (1998) measured on their sites. Similarly, with single-Alter-shielded instruments, the CE was around 80% (±17%) on the Neige site, comparable to the 76% and 83% observed in the Kochendorfer et al. (2017) and Goodison (1998) datasets, respectively. Smith (2009) measured an average CE of 36% for single-Alter-shielded instruments at the Bratt's Lake site, but the average wind speed during these snow events was higher than 5 m s^{-1}, which almost never occurred at the Neige site. When compared with unshielded gauges, the single-Alter-shielded observations resulted in a higher CE and a lower standard deviation. Altogether, these observations support the use of a single Alter shield for automated solid precipitation measurements.

A qualitative evaluation (Table 5) of the different transfer functions from the literature was performed for the Neige site. The application of equation G1 24H on unshielded instrument data increased the accuracy of the results, even if it overestimated the total quantity of solid precipitation (Fig. 7; Table 5). For the single-Alter-shielded data, the application of G1 24H equation produced the smallest P bias of all of the single-Alter adjustments applied. Even with the increased variance of the adjusted measurements, the estimated total quantity of solid precipitation was the closest to the DFIR reference for single-Alter-shielded adjustments. Regarding the statistical criteria from adjusted single-Alter-shielded measurements, our results suggest that the accuracy of the results after applying equation G1 24H remained about the same; the variance was increased and the quantity of solid precipitation was overestimated. Because of this, the equation G1 24H for both the unshielded and single-Alter twice-daily data is the recommended adjustment to apply at the twice-daily time step for the Neige site.

The application of equation S2 12H on single-Alter twice-daily data produced the highest RMSE and MAE we observed among all the single-Alter transfer functions. It also resulted in the highest P bias and the smallest *r*. This adjustment produced the least favorable results (Fig. 7; Table 5) and was the least appropriate for the Neige site.

The application of equation K3 12H to unshielded measurements improved the accuracy of the results, but not without increasing their variance (Fig. 7; Table 5). It led to the highest P bias observed for all the adjustments evaluated for the twice-daily unshielded gauge data. Equation K3 12H still overestimated the total quantity of solid precipitation, similar to G1 24H. Therefore, it is the second-least-recommended transfer function for the unshielded gauges at the Neige site. For the K3 12h single-Alter adjustment, the accuracy decreased, the variance of the results increased, and the amount of solid precipitation was overestimated. This is similar to equation S2 12H, but with generally improved results (Fig. 7; Table 5).

The application of equation K4 0.5H to unshielded measurements, although it overestimated the total quantity of solid precipitation, led to the greatest increase in accuracy compared to all the other adjustments (RMSE and MAE criterion) and a decrease in the variance. This adjustment produced the smallest P bias observed among all the unshielded adjustments. The *r* was also closest to 1 of all of the hourly unshielded adjustments. Overall, equation K4 0.5H is the most appropriate transfer function for hourly data at the Neige site (Fig. 7; Table 5).

c. *Hourly and twice-daily adjustment*

Previously, Kochendorfer et al. (2017) recommended the use of their equations with hourly data rather than twice-daily data. At the Neige site, this recommendation was validated for both unshielded instruments and single-Alter-shielded instruments, with an increase in the accuracy (RMSE: −20% on average; MAE: −20%; PE ± 10%

DFIR: +10%), an unchanged variance of the results (RMSE − MAE averaged unchanged; *r*: +1% on average), and a decrease in the bias (P bias: −9%) (Fig. 7; Table 5). Both our results and those of Kochendorfer et al. (2017) suggest that the K3 equations were more accurate when applied to hourly datasets than to twice-daily datasets.

The increased correlation between the average wind speed and the wind speed occurring during precipitation at shorter measurement intervals contributes to the benefits of the hourly adjustments.

d. Temperature effect

For the results illustrated in Figs. 5 and 6, the largest difference between the K3 0.5H and K4 0.5H transfer functions occurred with the unshielded instrument. As for the results of the statistical analyses, the adjustment using equations K3 0.5H and K4 0.5H led to comparable results. The K3 equations were functions of wind speed and temperature, whereas the K4 0.5H equation relied exclusively on wind speed. For the K4 0.5H equation, the temperature must first be used to determine the precipitation phase and the appropriate coefficients. To calibrate equations G1 24H, S2 12H and K4 0.5H, the authors separated rain, mixed and snowfall events before constructing and calibrating their equations, so these past solid precipitation equations did not depend on temperature. In our study, manual observations allowed the elimination of rain and mixed precipitation events with greater confidence. Moreover, we observed that the temperature, as considered in the K3 equations, did not impact the catch efficiency of solid precipitation on the site, but can still be used to help determine the precipitation type as in equations G1 24H, S2 12H, and K4 0.5H.

Then, assuming that the air temperature and cloud-top temperature are correlated, this parameter could affect the size and the shape of the snowflakes combined to the moisture and pressure parameters (Bourgoin 2000; Dubé 2003). Thériault et al. (2012) showed that different types and sizes of snowflakes interact differently with the wind, which affects the measured catch efficiency. The temperature effect on CE seems to be more a combination of multiple parameters, which could explain why it did not appear in an obvious way for the Neige dataset (Figs. 5 and 6).

e. Recalibration of the parameters

1) Recalibration and similarities between the datasets used for initial calibration

The similarity between the initial and recalibrated G1 24H parameters indicates that the measurements from the Neige site were similar to the Goodison et al. (1998) measurements, which supports the PCA results (Fig. 4).

2) Evaluation of the efficiency of the recalibration of the parameters

After applying the adjustment using the recalibrated parameters, the accuracy of the results increased relative to the original parameters found in the literature (RMSE: −50%; MAE: −40%; PE ± 10% DFIR: +45%). The variance of the results decreased (RMSE − MAE: −40%; *r*: +6%), and the P bias was reduced (P bias: −9%) (Fig. 7). Across all adjustment and measurement types, the recalibrated parameters generally outperformed the transfer functions available from past studies. The adjustments using the recalibrated K3 12H and G1 24H equations on unshielded and single-Alter-shielded twice-daily data were more suitable for the Neige site.

f. High– and low–wind speed effect

Some sites used to calibrate equations S2 12H, K3, and K4 had higher wind speeds during precipitation events than other sites (Table 2). High wind speeds can result in very small CE values (Smith 2009) and, therefore, explain why the catch efficiency adjustment curves of equations S2 12H, K3 and K4 were below that of G1 24H (Fig. 5). This helps to explain why the precipitation at the Neige site was frequently overestimated using these adjustments, resulting in an averaged absolute P bias that was 2 times as high as after applying the G1 24H equation.

g. Climate and wind speed impact

The Neige site was located in an eastern boreal climate zone, with the second-highest annual and winter solid precipitation quantity of all of the intercomparison sites listed in Table 2.

The climate characteristics of the Neige site were closer to the sites used to calibrate the G1 24H equation (Goodison et al. 1998) than the other sites, especially the amount of precipitation (Table 2; Fig. 4). When averaged for both instrument types, the application of equation G1 24H to the twice-daily measurements resulted in better accuracy, variance, and bias than the other adjustments (Fig. 7; Table 5). This underscores the importance of site characteristics, such as climate and mean wind speed, rather than measurement technology, as the Goodison et al. (1998) study included only manual gauge measurements, whereas the present study compared automatic and manual measurements.

The dataset used to calibrate the K3 and K4 0.5H equation coefficients encompassed the most variability of climatic characteristics and wind profiles as a result of its provenance from eight sites throughout the Northern

Hemisphere (Table 2). The virtual site of the K3 equations on the PCA individual map was the second most similar to the Neige site (Fig. 4). However, the application of the K3 equations led to inferior statistical results than equation G1 24H. In theory, the increased variability of the climate and wind speed of the precipitation events used to derive the K3 equations should produce a more generic adjustment, but it was still less appropriate for the Neige site than equation G1 24H.

The site used for the calibration of equation S2 12H (Smith 2009) was located in the Canadian prairie region. It had the lowest annual precipitation quantity of all the sites included in this study, and its average wind speeds were 2 times as high as the Neige site (Table 2). The position of the site in the PCA individual map was the farthest from the Neige site (Fig. 4). Moreover, regarding all statistical criteria, the S2 12H adjustment results were worse than applying no adjustment (Fig. 7; Table 5), which confirms that the S2 12H adjustment was not appropriate for the Neige site.

The structure of the transfer functions was not evaluated in detail, but this study shows that the use of transfer function parameters calibrated on datasets from sites with similar climate and wind speeds is the most sensible adjustment choice. Kochendorfer et al. (2017) also found that the application of the adjustment to sites with different climates and/or wind speed distributions produced the least satisfactory results. Both climate and wind speed seem to impact the catch efficiency, emphasizing their importance in choosing the appropriate adjustment for an independent site.

h. *Unshielded and single-Alter-shielded precipitation adjustment*

As S2 12H was not derived for unshielded measurements, only the G1 24H, K3, and K4 adjustment results, using parameters from the literature, were considered for comparison. For the unshielded measurements, the adjustment induced an average increase in the accuracy of the prediction (RMSE: −30%; MAE: −40%; PE ± 10% DFIR: doubled), an increase in the variance of the results (RMSE − MAE: +15%; *r*: +6%), and an improvement in the bias (absolute P bias: −80%). On the contrary, for the single-Alter-shielded instruments' data, application of the adjustment decreased the accuracy of the prediction (RMSE: +30%; MAE: unchanged; PE ± 10% DFIR: +60%), and increased the variance of the results (RMSE − MAE: doubled; *r*: −5%). Considering all statistical criteria, regardless of the adjustment applied, using parameters from the literature, the adjustment of unshielded measurements was at least as accurate as the adjustment of the single-Alter-shielded measurements (Fig. 7; Table 5). This is notable considering that the unshielded measurements were initially less accurate than the single-Alter-shielded measurements. However, the PE ± 10% DFIR of single-Alter-shielded adjusted data was on average 4 times that for the unshielded adjusted data. This is in contrast with the findings of Kochendorfer et al. (2018), who showed that less well-shielded gauges were subject to larger errors even after adjustment by an appropriate transfer function. However, the Neige dataset included only one unshielded instrument for one full winter, but it included four single-Alter-shielded instruments, which could induce some additional noise in the single-Alter-shielded dataset and explain the decreased accuracy of the results after the adjustments. Additional replications of unshielded measurements from the Neige site would help to determine how specific these results are to the gauges that were tested (Bartlett et al. 2001).

i. *Total quantity of precipitation and hydrological relevance*

Total solid precipitation was underestimated by 34% and 21% based on the unadjusted unshielded and single-Alter-shielded measurements, respectively, which would be deleterious to the water balance (Fassnacht 2007). The application of adjustments using parameters from the literature induced an overestimation from 2% to 26% of the total precipitation compared to the DFIR, considering both unshielded and single-Alter-shielded gauges (Fig. 7). The Neige site recorded the second-highest annual and winter precipitation quantity (Table 2) and the most snow events of all the listed study sites. With more precipitation, there is a larger potential accumulated absolute error during the winter. During an average winter at the Neige site, the overestimation could easily exceed 100 mm. The main focus of the research conducted at the Neige site is improving the understanding and estimation of total annual precipitation within the framework of the total hydrologic balance. For flood forecasting, an overestimation of water input would be less grievous than an underestimation. Thus, all adjustments could indirectly reduce the risk of missing a spring flood warning (Whitfield 2012). However, estimated precipitation remains biased.

5. Conclusions

Five transfer functions developed by Goodison et al. (1998), Smith (2009), and Kochendorfer et al. (2017) were evaluated. The data from the Neige site were used to test the appropriateness of available transfer functions for this site and other sites with similar climatic regimes.

Kochendorfer et al. (2017) hypothesized that variability in catch efficiency depends more on site climate

characteristics than on the type of instrument, which helps explain the results presented here. Our work indicates that similarity in the meteorological characteristics at the calibration site and the validation site result in similar CE transfer functions.

For all instruments, the appropriate adjustments were more effective and accurate on hourly data than twice-daily data. This is important to note, especially as Hultstrand and Fassnacht (2018) found that the uncertainty related to measuring solid precipitation could adversely impact the water balance.

Based on many of the error statistics, adjusted single-Alter-shielded data were less accurate than the adjusted unshielded data, but this result is likely specific to the Neige dataset, rather than being generally true for shielded and single-Alter-shielded transfer functions elsewhere. The PE ± 10% DFIR accuracy criterion of single-Alter-shielded adjusted data remained on average 4 times larger than that for the unshielded adjusted data. However, more replicate unshielded gauges at the Neige site are needed to support these conclusions (Bartlett et al. 2001). Adjustment using parameters from the literature induced an overestimation of the winter precipitation, regardless of the instruments considered, which could be deleterious from a hydrological point of view.

The improvement of the results using the recalibrated parameters on the Neige dataset supports the hypothesis that meteorology and siting affect the magnitude of solid precipitation undercatch adjustments. Thus, for sites that do not have a reference measurement such as the DFIR, where it would be impossible to fit custom coefficients to the equations, we recommend 1) choosing a calibration site with climatic and wind speed profiles close to those of the study site and 2) making adjustments to hourly measurements rather than to daily or twice-daily measurements.

Precipitation adjustment errors will typically be largest for the snowiest sites. Moreover, measurement uncertainties increase significantly with high wind speeds, and this impacts the uncertainty of the adjustments. The true complexity of solid precipitation undercatch cannot be encompassed in a single deterministic function of wind speed and air temperature. A probabilistic approach could include a confidence interval for each solid precipitation event, which would be more useful for hydrological modeling than a deterministic solution.

Acknowledgments. We thank the Ministère du Développement Durable et de la Lutte contre les Changements Climatiques (MDDELCC) and Environment and Climate Change Canada (ECCC) for the automated measurements. We thank the team of technicians from Forêt Montmorency for the high-quality manual observations. We also thank N. Francezon, A. Benmoussa, and M. Robinson for reviewing an early draft of the paper, and we thank N. Perreault for the map. This work was supported in part by Ouranos Consortium on Regional Climatology and Adaptation to Climate Change, Hydro-Québec, the Natural Sciences and Engineering Research Council of Canada, the MDDELCC, and Environment and Climate Change Canada through Project RDC-477125-14 entitled ''Modélisation hydrologique avec bilan énergétique (ÉVAP).''

REFERENCES

Barbier, S., P. Balandier, and F. Gosselin, 2009: Influence of several tree traits on rainfall partitioning in temperate and boreal forests: A review. *Annu. For. Sci.*, **66**, 602, https://doi.org/10.1051/forest/2009041.

Bartlett, J. E., J. W. Kotrlik, and C. C. Higgins, 2001: Organizational research: Determining appropriate sample size in survey research. *Inf. Technol. Learn. Perform. J.*, **19**, 43–50.

Bourgoin, P., 2000: A method to determine precipitation types. *Wea. Forecasting*, **15**, 583–592, https://doi.org/10.1175/1520-0434(2000)015<0583:AMTDPT>2.0.CO;2.

Chai, T., and R. R. Draxler, 2014: Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.*, **7**, 1247–1250, https://doi.org/10.5194/gmd-7-1247-2014.

Cornillon, P.-A., A. Guyader, F. Husson, N. Jégou, J. Josse, M. Kloareg, É. Matzner-Løber, and L. Rouvière, 2012: *Statistiques avec R* (*Statistics with R*). 3rd ed. Presses Universitaires de Rennes, 296 pp.

Devine, K. A., and É. Mekis, 2008: Field accuracy of Canadian rain measurements. *Atmos.–Ocean*, **46**, 213–227, https://doi.org/10.3137/ao.460202.

Dubé, I., 2003: De mm à cm. . .: Étude des rapports neige/eau liquide au Québec (From mm to cm. . .: Study of snow/liquid water reports in Quebec). Service Météorologique du Canada Tech. Rep., 131 pp.

Fassnacht, S. R., 2007: Data time step to estimate snowpack accumulation at select United States meteorological stations. *Hydrol. Processes*, **21**, 1608–1615, https://doi.org/10.1002/hyp.6723.

Goodison, B. E., P. Y. T. Louie, and D. Yang, 1998: WMO solid precipitation measurement intercomparison. WMO Instruments and Observing Methods Final Rep. 67, 318 pp., http://www.wmo.int/pages/prog/www/reports/WMOtd872.pdf.

Hauke, J., and T. Kossowski, 2011: Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geogr.*, **30**, 87–93, https://doi.org/10.2478/v10117-011-0021-1.

Henshall, W. R., and W. P. Snelgar, 1989: A small unaspirated screen for air temperature measurement. *N. Z. J. Crop Hortic. Sci.*, **17**, 103–107, https://doi.org/10.1080/01140671.1989.10428016.

Hultstrand, D. M., and S. R. Fassnacht, 2018: The sensitivity of snowpack sublimation estimates to instrument and measurement uncertainty perturbed in a Monte Carlo framework. *Front. Earth Sci.*, **12**, 728–738, https://doi.org/10.1007/s11707-018-0721-0.

Jolliffe, I., 2011: Principal component analysis. *International Encyclopedia of Statistical Science*, M. Lovric, Ed., Springer, 1094–1096, https://doi.org/10.1007/978-3-642-04898-2_455.

Jones, H. G., and J. W. Pomeroy, 2001: Early spring snowmelt in a small boreal forest watershed: Influence of concrete frost on the hydrology and chemical composition of streamwaters during rain-on-snow events. *58th Eastern Snow Conf.*, Ottawa, ON, Canada, Eastern Snow Conference, 209–218, http://www.usask.ca/hydrology/papers/Jones_Pomeroy_2001.pdf.

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.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel, 2006: World map of the Köppen–Geiger climate classification updated. *Meteor. Z.*, **15**, 259–263, https://doi.org/10.1127/0941-2948/2006/0130.

McKay, G. A., 1968: Problems of measuring and evaluating snow cover. *Proc. Workshop Seminar of Snow Hydrology*, Ottawa, ON, Canada, Canadian National Committee, 49–65.

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

Nalder, I. A., and R. W. Wein, 1998: Spatial interpolation of climatic normals: Test of a new method in the Canadian boreal forest. *Agric. For. Meteor.*, **92**, 211–225, https://doi.org/10.1016/S0168-1923(98)00102-6.

Nitu, R., and K. Wong, 2010: CIMO survey on national summaries of methods and instruments related to solid precipitation measurement at automatic weather stations—Preliminary. WMO Rep. 102, 57 pp., https://library.wmo.int/pmb_ged/wmo-td_1544.pdf.

——, and Coauthors, 2012: WMO intercomparison of instruments and methods for the measurement of solid precipitation and snow on the ground: Organization of the experiment. *WMO Tech. Conf. Instruments and Methods of Observation (TECO 2012)*, Brussels, Belgium, World Meteorological Organization, 1-1, https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-109_TECO-2012/Session1/O1_01_Nitu_SPICE.pdf.

Oreiller, M., 2013: Modélisation, par une approche physique, de l'équivalent en eau de la neige (Modeling, by a physical approach, of the water equivalent of snow). M.S. thesis, Dept. of Science des Eaux, Université du Québec, 70 pp.

Plamondon, A. P., M. Prévost, and R. C. Naud, 1984: Accumulation et fonte de la neige en milieux boisé et déboisé (Accumulation and melting of snow in woodland and deforested areas). *Geogr. Phys. Quat.*, **38**, 27–35.

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., M. Earle, A. Gaydos, and M. A. Wolff, 2016: Description of the quality control and event selection procedures used within the WMO-SPICE project. *Proc. WMO Technical Conf. on Meteorological and Environmental Instruments and Methods of Observation*, Madrid, Spain, WMO, P3.10.

Rodgers, J. L., and W. A. Nicewander, 1988: Thirteen ways to look at the correlation coefficient. *Amer. Stat.*, **42**, 59–66, https://doi.org/10.1080/00031305.1988.10475524.

Rubel, F., K. Brugger, K. Haslinger, and I. Auer, 2017: The climate of the European Alps: Shift of very high resolution Köppen–Geiger climate zones 1800–2100. *Meteor. Z.*, **26**, 115–125, https://doi.org/10.1127/metz/2016/0816.

Ryu, S., G. Lee, R. Nitu, C. Smith, E. Lim, and H. Kin, 2012: Automatic double fence reference (DFAR) for measuring solid precipitation: Gauge based characterization. *Second Session of the Int. Organization Committee for the WMO Solid Precipitation Intercomparison Experiment*, Boulder, CO, WMO.

Sevruk, B., J. A. Hertig, and R. Spiess, 1989: Wind field deformation above precipitation gauge orifices. *IAHS Publ.*, **179**, 65–70.

Smith, C. D., 2009: The relationship between snowfall catch efficiency and wind speed for the Geonor T-200B precipitation gauge utilizing various wind shield configurations. *77th Western Snow Conf.*, Canmore, AB, Canada, Western Snow Conference, 115–121.

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.

Whitfield, P. H., 2012: Floods in future climates: A review. *J. Flood Risk Manage.*, **5**, 336–365, https://doi.org/10.1111/j.1753-318X.2012.01150.x.

Willmott, C. J., and K. Matsuura, 2005: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Res.*, **30**, 79–82, https://doi.org/10.3354/cr030079.

Wolff, M. A., K. Isaksen, A. Petersen-Øverleir, K. Ødemark, T. Reitan, and R. Brækkan, 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., 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.

——, J. R. Metcalfe, B. E. Goodison, and É. Mekis, 1993: ''True snowfall'': An evaluation of the double fence intercomparison reference gauge. *61th Western Snow Conf.*, Quebec City, QC, Canada, Western Snow Conference, 105–111.