Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Hydrology and Earth System Sciences
Open Access
Discussions

The Quantification and Correction of Wind-Induced Precipitation Measurement Errors

John Kochendorfer[1], Roy Rasmussen[2], Mareile Wolff[3], Bruce Baker[1], Mark E. Hall[1], Tilden Meyers[1], Scott Landolt[2], Al Jachcik[2], Ketil Isaksen[3], Ragnar Brækkan[3], and Ronald Leeper[4,5]

[1]ARL/Atmospheric Turbulence and Diffusion Division, National Oceanic and Atmospheric Association, Oak Ridge, TN, 37830, US
[2]National Centers for Atmospheric Research, Boulder, 80305, US
[3]Norwegian Meteorological Institute, Oslo, 0313, Norway
[4] N Carolina State Univ., Cooperative Inst of Climate and Satellites, Asheville, 28801 US
[5] National Center for Environmental Information, National Oceanic and Atmospheric Association, Asheville, 28801 US

Correspondence to: John Kochendorfer (john.kochendorfer@noaa.gov)

Abstract

Hydrologic measurements are becoming increasingly important for both the short and long term management of water resources. Of all the terms in the hydrologic budget, precipitation is the typically most important input. However, measurements of precipitation are still subject to large errors and biases. For example, a high-quality but unshielded weighing precipitation gauge can collect less than 50% of the actual amount of solid precipitation when wind speeds exceed 5 ms^{-1}. Using results from two different precipitation testbeds, such errors have been assessed for unshielded weighing gauges and for four of the most common windshields currently in use. Functions used to correct wind-induced undercatch were developed and tested. In addition, corrections for the single Altar weighing gauge were developed using the combined results of two separate sites, one of which was in Norway and other in the US. In general the results indicate that corrections described as a function of air temperature and wind speed effectively remove the undercatch bias that affects such precipitation measurements. In addition, a single 'universal' function developed for the single Altar gauges effectively removed the bias at both sites, with the bias at the US site improved from -12% to 0%, and the bias at the Norwegian site improved from -27% to -3%. These correction functions require only wind speed and air temperature, and were developed for use in national and local precipitation networks, hydrological monitoring, roadway and airport safety work, and climate change research. The techniques used to develop and test these transfer functions at more than one site can also be used for other more comprehensive studies, such as the WMO Solid Precipitation Intercomparison Experiment.

1 Introduction

Precipitation measurements are used by policy makers, hydrologists, farmers, and watershed managers to quantify and allocate the water available for society's needs. Precipitation measurements are necessary for public safety in areas as

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

diverse as avalanche control, flood forecasting, roadway safety, and aircraft de-icing operations. Precipitation measurements are also used to evaluate radar-based estimates of rainfall, to monitor climate change, and help improve climate and weather models. More specifically, monitoring changes in the frequency, intensity, duration and phase of precipitation is critical for current and future climate research (Trenberth et al., 2003;Barnett et al., 2005). Although precipitation has been monitored for many centuries, the need to accurately measure precipitation will become even more important in the future, as changes in precipitation are expected to be complex and variable with location, requiring robust and accurate precipitation measurements. Despite this critical need, precipitation observations are still beset with significant biases and errors (eg. Forland and Hanssen-Bauer, 2000;Adam and Lettenmaier, 2003;Yang et al., 2005;Groisman and Legates, 1994).

Solid precipitation is particularly difficult to measure accurately, and biases between wintertime precipitation measurements made using different technologies, different measurement networks, or across different regions can be larger than 50% (Rasmussen et al., 2012;Yang et al., 1999). Previous studies have identified wind effects as one of the primary causes for snow undercatch (eg. Folland, 1988). This is due to two important factors: 1) the relatively slow fall velocity of snow, and 2) the creation of flow distortions of similar magnitude to the snow fall velocity by the gauge itself as air flows past it. These two factors cause a snowflake trajectory to be significantly deflected by the airflow past the gauge. In particular, the updraft at the leading edge of the gauge can lead to an upward deflection of snowflake trajectories, causing snow to miss the gauge orifice and not be measured. The flow distortion around the gauge increases as the wind increases, while snowflake terminal velocity remains the same, causing more snowflakes to be deflected around the gauge. Thus, the collection efficiency of snow by a gauge decreases with increasing wind speed. The collection of rain in a weighing gauge suffers from this same problem, but to a much lesser extent due to the order of magnitude higher fall velocity of raindrops, allowing them to be subject to only minimal trajectory deflection due to flow distortions around the gauge.

To mitigate this deflection and undercatch of snow by weighing gauges, scientists have designed wind shields to surround the gauge with the goal of slowing the oncoming flow, and lessen the resulting flow distortions around the gauge. Prominent among these is the Alter shield (Alter, 1937). This shield consists of vertical slats of ~40 cm length suspended from a circular rim ~100 cm in diameter surrounding the gauge. The flow speed is indeed decreased (Rasmussen et al., 2012), and the flow distortion around the gauge reduced leading to an increase in snow collection. This shield is widely used around the world to improve the measurement of snow and rain.

To quantify the impact of the Alter shield and other types of shields (Yang et al., 1995;Alter, 1937;Nipher, 1878), the World Meteorological Organization (WMO) sponsored a solid precipitation measurement program for manual weighing gauge in the early 1990's (Goodison et al., 1997). A significant result of this study was the creation of a reference standard gauge for snowfall. This was determined to be a precipitation gauge embedded within a carefully pruned bush without leaves. The height of the bush was the same level as the orifice of the gauge. A secondary reference was established as the Double Fence

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Intercomparison Reference (DFIR, Yang, 2014;Groisman et al., 1991). This allowed field sites to establish a standard for snow measurement using a DFIR without having to grow and prune a bush system. A key characteristic of the reference was that the snowfall rate did not significantly depend on the magnitude of the wind.

The 1990's WMO study used both the bush and DFIR as references to compare to a variety of gauges with different shielding (Goodison et al., 1997). A key result from this study was that the collection efficiency of the gauge was primarily determined by the type of wind shield used (Yang et al., 1999). Rasmussen et al. (2012) compiled a review of snowfall measurements to date that confirmed this result for automatic weighing gauges and other wind shield types.

These studies confirmed that the weighing gauge collection efficiency for snow as compared to the above-mentioned reference gauge systems decreased with increasing wind speed, depending on the type of the shield. An unexpected result was that the collection efficiency for a given wind speed and wind shield/gauge type varied significantly (Yang et al., 1999, Rasmussen et al., 2012). While functions describing the decrease in collection efficiency could be derived, there was often as much variability at a given wind speed as across the range of wind speeds (Yang et al., 1999). These results revealed that wind speed is not the only factor that impacts the trajectory of a snowflake past a gauge. Since the fall speed of a snowflake is often close to the magnitude of the flow distortion, an obvious candidate was the fall speed of the snow. Using numeric modelling, Theriault et al. (2012) showed that the difference between wet and dry snow fall speeds can lead to significant changes in the collection efficiency of snow by a GENOR weighing gauge with and without an Alter shield. Other factors include airflow turbulence, time dependence of the flow past a gauge (Colli et al., 2015), snow size distribution (Theriault et al., 2012), and snow density (Colli et al., 2015).

As a result, the performance of weighing snow gauges with various types of wind shields may be expected to vary by climate regions. The previous WMO Solid Precipitation Intercomparison (Goodison et al., 1997) examined solid precipitation from manual gauges at a limited number of field sites. The more recent WMO Solid Precipitation Intercomparison Experiment (SPICE) expanded the number of climate regimes covered, and used automatic gauges instead of manual (Nitu et al., 2016). This paper examines results from two sites participating in this experiment, and includes several years of measurements that pre-date the WMO experiment for more robust results. These results include the assessment of various weighing gauge/shield combinations across these sites and: 1) the dependence of snow collection efficiency as a function of wind speed as well as its uncertainty at the various measuring sites with characteristic climatological conditions, 2) identifies a functional form for the collection efficiency/wind speed relationship (referred to as a "transfer function"), 3) assesses whether each climatological site requires a different transfer function or if a universal function can be used to reduce wind induced snow undercatch, and 4) quantifies the expected uncertainty of the correction as a function of gauge/shield type and climatological condition.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

2 Methods

2.1 Site Descriptions

The US field site is located just outside of Boulder, Colorado along the eastern slopes of the Colorado Front Range. The site resides on top of the Marshall Mesa at 39.949° north, 105.195° west, and is ~1740 m above sea level. Prairie grasses and low scrub are the primary flora found at the site, and its proximity to the mountains makes it an ideal location for studying upslope snowfall events. The US site typically receives ~200 cm of snowfall throughout the course of the winter months. Snowfall at the site typically occurs from October through April but can occasionally occur as early as September or as late as June. The site is relatively flat and the lack of trees, large buildings and other obstacles allow for uninterrupted wind flow around the gauges. All of the gauges included in this study were mounted with their inlets approximately 1.9 m above the ground. The layout was designed to minimize the effects of the gauges and their wind shields on each other. The site generally stretches out perpendicular to the wind direction that prevails during snow storms, and special care was taken not to put larger shields upwind of unshielded and smaller-shielded gauges. Comparison of replica single- and double-Alter shielded gauges from differing locations at the field site and a lack of any significant wind direction effects on catch efficiency confirmed the fact that wind direction did not play an important role in catch efficiency at the site.

The Norwegian test site (Haukeliseter) is situated at 59.812° N, 7.214° E, and is at 991 m a.s.l. on a plateau in an alpine region in southwestern Norway. Snowfall at the site typically occurs October through May. The site has one Double Fence Intercomparison Reference (DFIR) surrounding an automated weighing gauge within a single Alter shield (SA). The DFIR and several gauges, most of them within single Alter shields, are installed in two lines perpendicular to the main wind directions (east-southeast and west-northwest). An evaluation of precipitation and wind measurements recorded at different locations throughout the site indicated that the site was homogeneous (Wolff et al., 2010). The gauges at the Norwegian test site are all mounted at 4.5 m altitude in order to mitigate the effects of blowing snow and to allow for the increasing snow depth through the season. Annual maximum snow depth reaches ~3 m. The site is described in more detail in previous publications (Wolff et al., 2013;Wolff et al., 2015).

2.2 Precipitation Gauges and Shields

To reduce potential sources of uncertainty, all of the precipitation measurements presented here were recorded using the same model weighing precipitation gauge (3-wire T200B, Geonor Inc., Oslo, Norway) and all of the gauge inlets were heated using the same type of inlet heaters (described in NOAA Technical Note NCDC No. USCRN-04-01). Both the outer and inner tubes of the GEONOR were heated to prevent snow melted at the orifice from re-freezing while it drips into the collection bucket. The inlet heaters were activated only when the inlet temperature and the air temperature were both < 2 °C.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

The DFIR shield has the largest footprint of any of the shields, and consists of three concentric shields. The outer two shields are octagonal in design with the outer shield having a diameter of twelve meters and the middle shield having a diameter of four meters. The DFIR shield has a porosity of 50%, and both the outer and middle shields are perpendicular to the ground. For the third innermost shield, an Alter-style shield of standard size and configuration is used. The DFIR shield is described in more detail in the first WMO Solid Precipitation Intercomparison (Goodison et al., 1997).

The 2/3 scale version of the DFIR, hereafter referred to as the small DFIR (SDFIR), was designed for the US Climate Reference Network program and tested at the US site. The SDFIR laths are 1.2 meters long, and the diameter of the outer shield is eight meters in diameter and the middle shield is 2.6 meters in diameter. Additionally, the middle shield height is 10 cm lower than the outer shield. A standard diameter Alter shield is used as the innermost shield, which is 10 cm lower than the middle shield and located at the same height as the gauge orifice inlet.

Alter shields, hereafter referred to as single Alter shields (SA) to avoid confusion with the double Alter shields, consist of metal laths about 40 centimeters in length (though some versions of the Alter use slightly longer laths that are 46 centimeters in length). The laths on the SA shield are typically attached near the top to a circular ring, 1.2 meters in diameter, and allowed to move freely in the wind. The double Alter shield (DA) is a variation of the single Alter shield and has two concentric shields instead of one (Rasmussen et al., 2001). This shield consists of a standard 1.2 meters diameter single Alter shield surrounded by an additional outer ring of laths measuring two meters in diameter. Like the single Alter, the laths on both rings are approximately 40 cm in length, secured only at the top, allowing them to move freely at the bottom.

The Belfort double Alter shield is a modified version of the standard double Alter shield. The diameter of the inner shield is 1.2 meters and the laths are 46 cm long. The diameter of the outer shield is 2.4 meters and the laths are 61 cm long. Unlike the standard single and double Alter shield laths, these laths don't taper at the bottom and are only allowed to swing inwards or outwards at a maximum 45-degree angle. The Belfort double Alter shield is also only approximately 30% porous, which is significantly less than the ~50% porous double Alter shield.

2.3 Other Measurements

2.3.1 US Site

At the US testbed the air temperature was measured using fan-aspirated (Model 076B Radiation Shield, Met One Instruments, Grants Pass, OR, US) platinum resistance thermometers (Thermometrics, Northridge, CA, US) mounted at a height of 1.5 m. Three wetness sensors (Model DRD11A, Vaisala, Helsinki, Norway) also mounted at a height of 1.5 m were used to independently detect precipitation. Wind speed was measured at 1.5 m using a cup anemometer (Model 014A Wind Speed Sensor, Met One Instruments, Grants Pass, OR, US), at 2.0 and 3.0 m using propeller anemometers (Model 05103 Wind Monitor, RM Young, Traverse City, MI), and at 10 m using both a propeller anemometer (Model 05103 Wind

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Monitor, RM Young) and a two-dimensional sonic anemometer (Model 86004 Ultrasonic Anemometer, RM Young). The two anemometers at 10 m were found to interfere with each other due to wind shadowing when winds were from the north or the south, and a composite 10 m wind speed was therefore produced using the ultrasonic anemometer measurements to replace the propeller anemometer measurements when winds were from the north (wind directions < 30° or > 340°). Likewise, as identified by plotting the ratio of the measured wind speed to the 10 m wind speed as a function of wind direction, the lower wind speeds were found to be subject to interference at some wind directions. Using 30-min mean wind speeds measured in a clear sector (wind direction < 30°) from all measurement locations, the roughness length (z_0 = 0.01 m) and displacement height (d = 0.4 m) were determined based on the log wind profile (Thom, 1975):

$$U_z \approx \ln\left[{}^{(z-d)}/_{z_0}\right] \quad (1)$$

where U_z is the wind speed (U) at a height z (U_z). Using the same relationship, the wind speed at the gauge height of 1.9 m was estimated be equal to U_{10m}×0.71. Only wind speed measurements recorded during precipitation events were used to develop this relationship, minimizing the neglected effects of stability on the wind profile in the typically overcast and near-neutral surface layer conditions associated with precipitation. Due to the effects of near-field obstructions on the near-surface wind speed measurements, the 10 m wind speed measurement was used to estimate the gauge-height wind speed measurement using this method throughout the study. The 10 m wind speed was more generally representative of the wind speed affecting all of the gauges throughout the site, and it produced more accurate precipitation corrections than the gauge-height wind speed. Errors in the method used to estimate the gauge height wind speed from the 10 m wind speed were evaluated using the mean half-hour 2 m wind speed measurements recorded during precipitation events when the recorded wind speed was greater than 1 m/s from unobstructed wind directions (WD < 30 deg). The 2 m wind speeds were compared to the gauge height (1.9 m) wind speed estimated using the log profile, resulting in an RMSE of 0.4 ms^{-1} (10.1%) and a bias of -0.12 ms^{-1} (-3.1%).

2.3.2 Norwegian Site

At the Norwegian site (hereafter NOR), unobstructed 10 m and gauge height (4.5 m) wind speed measurements recorded during precipitation events were used to determine the relationship between the gauge height and 10 m wind speeds: U_{gh} = 0.93×U_{10m}, R^2 = 0.99, RMSE = 0.54 m s^{-1}, and the wind speed at 10 m was used to predict the gauge-height wind speed from all wind directions. Following Wolff et al. (2015), precipitation events were also screened for wind directions associated with shadowing between gauges and shields, and were excluded from the analysis.

2.4 Precipitation type

Transfer functions have commonly been developed separately for snow, mixed precipitation, and rain (eg. Goodison et al., 1997;Yang et al., 2005). Other proposed classification schemes also include differentiation between wet and dry snow. In the past, manual observations of precipitation type were recorded and used to develop such transfer functions, but modern

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

automated measurement networks now rarely include such manual measurements. Airport weather stations often include precipitation type measurements, but hydrological, meteorological, and climate stations do not typically include precipitation type measurements, and defensible methods to correct wind-induced errors without precipitation-type measurements are therefore needed.

At the US testbed, precipitation type was determined using a present weather detector (Vaisala PWD22, Helsinki, Finland). Half hour increments of rain, mixed, and snow were identified using more than 15 min of any of these precipitation types as detected from the present weather detector measurements recorded every minute. For rain and snow, less than 5 min total of any other precipitation type were allowed to occur. These identifiers were used only for the analysis of precipitation type shown in Figures 2, 3 and 5, but not for the development of transfer functions.

At the US site, below -2.5 ˚C more than 95% of the precipitation in every 1 ˚C bin was classified as snow using the present weather detector. Above 2.5 ˚C, more than 95% of the precipitation was classified as rain. These thresholds will of course change depending upon the climate of a given site, and the 95% threshold could also be adjusted to suit the needs of a given study. However, based on these temperature thresholds, the precipitation type that was most sensitive to classification methods was mixed precipitation. The present weather detector classified only 5% of the available half hours of precipitation as mixed, whereas 19% of the precipitation occurred between -2.5 ˚C and 2.5 ˚C. A significant amount of rain and snow may therefore be misclassified as mixed precipitation when temperature thresholds are used to predict precipitation type, assuming that the present weather detector accurately identified precipitation type.

An alternative to using temperature thresholds to differentiate between different precipitation types is to use a continuous function of both air temperature and wind speed to define the catch efficiency (Wolff et al., 2015). Although the functions produced using such methods are more complex than a relationship between wind speed and catch efficiency for a single precipitation type, such an approach is arguably more convenient because only one equation is needed to determine catch efficiency for all conditions. More importantly, a continuous function of temperature may more accurately represent reality, as catch efficiency varies continuously with air temperature, especially near 0 °C. The transition from liquid precipitation to dry snow is continuous, without any well-defined step changes, and any temperature threshold chosen to differentiate between different precipitation types will inevitably be somewhat arbitrary. In addition, a continuous function of air temperature and wind speed eases the comparison of catch efficiency results from different sites with potentially different climates. This in turn aides in the evaluation of the uncertainty inherent in a ‘universal’ transfer function describing more than one site.

Following Wolff et al. (2015), for the transfer functions developed here we use a continuous f(T_{air}, U) as air temperature measurements are more universally available than precipitation type, different precipitation type detectors do not always

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

agree on the precipitation type (Merenti-Valimaki et al., 2001;Wong, 2012), and air temperature thresholds used to separate different precipitation types must be chosen somewhat arbitrarily.

2.5 Transfer Function Development

2.5.1 Data analysis and event selection

For the US results, the US Climate Reference Network precipitation algorithm was used to determine 5-minute accumulations from all the 3-wire Geonor precipitation gauges (Leeper et al., 2015). The algorithm relied upon wetness sensor measurements to detect periods of precipitation, and it calculated the average accumulation of the three wires by inversely weighting the individual wire accumulations using the variance of the individual depths over the last 3 hours. This was done to lessen the contribution of noisier wires to the total gauge depth, and thereby decrease the amount of noise in the precipitation measurements. The algorithm was modified for the purposes of this study by increasing the 5-minute precipitation resolution from 0.1 mm to 0.01 mm.

In addition at the US site, more than 10 minutes total of any type of precipitation as identified by the present weather detector had to occur within each half-hour to be included in the transfer function analyses. Half-hour increments with unrealistic air temperatures or wind speeds were also excluded from the transfer function analyses. For example, half-hour increments with 10 m anemometer measurements affected by ice accumulation were identified by comparing the 10 m propeller anemometer measurements with the 1.5 m cup anemometer measurements.

For the NOR results the methods described in Wolff et al. (2015) were used for the selection of 60-min periods included in the analysis.

2.5.2 Selection of precipitation threshold.

For the US site, 30-minutes was selected as the most suitable time interval for the creation of transfer functions, as 30-minutes is near the centre of the gap in spectra describing atmospheric motions within the surface layer (eg. Stull, 1988), allowing averages of air temperature and wind speed measurements that are representative of the field site as a whole and also relatively stationary. A longer time period would be subject to increased mesoscale, synoptic, and diurnal changes in precipitation type, wind speed, and air temperature. Under quiescent conditions a shorter averaging period would approach turbulent time scales, where the presence or absence of an individual eddy would affect the results, making the results less representative of the entire site. At the NOR site, 60-minute periods of precipitation were used following Wolff et al. (2015), with many of the same arguments supporting the 30-minute period equally valid for 60-minutes.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Because catch errors are best described using catch efficiency, described as the ratio between a gauge under test and a standard gauge ($CE = {P_{UT}}/{P_{DFIR}}$, where *CE* is catch efficiency, P_{UT} is the accumulation of precipitation from a gauge under test, and P_{DFIR} is the accumulated DFIR precipitation used as the standard), a minimum threshold is necessary to constrain errors in the denominator of this ratio. Using all the 30-minute periods of snow measured within the SDFIR and DFIR gauges from the US site, a minimum 30-minute precipitation amount of 0.25 mm in the DFIR gauge was found to provide a good balance between reducing the effects of measurement noise while simultaneously maintaining a large sample size of events. This was examined by iteratively increasing the DFIR precipitation threshold from zero in 0.01 mm steps, and calculating a simple linear transfer function for each threshold. The number of 30-minute events (n) and standard deviation (σ) of the *CE* model were estimated for every 0.01 mm increase in threshold, and a minimum in the standard error ($SE = {\sigma}/{\sqrt{n}}$) was encountered at 0.25 mm (Fig. 3). For this threshold test the SDFIR gauge data was used to develop transfer functions, as it was the most comparable to the DFIR gauge and was described well by a simple linear transfer function for snow events. In reality, some variability occurs in this 'ideal' threshold based on the specific gauge/shield under test, the sample size, precipitation type, and the amount of noise in the measurements. However for the sake of simplicity and consistency the same 0.25 mm reference threshold was used for all of the gauges included in this study.

Using the original 30-minute SDFIR precipitation measurements, which were generally nearly equal to the DFIR measurements, we discovered that a minimum threshold for the gauge under test was also necessary to produce unbiased transfer functions. This was because both the gauge under test and the standard DFIR gauge were affected by random measurement error and random spatial variability in precipitation; the standard gauge/shield combination is capable of measuring more than 0.25 mm of precipitation in a 30-minute period even when the actual site-average rate is below this threshold. The minimum thresholds for the gauges under test were determined from Eq. 2, using the entire multi-year datasets available.

$$THOLD_{UT} = \frac{median(P_{UT})}{median(P_{DFIR})} 0.25\ mm \quad (2)$$

Where $THOLD_{UT}$ is the threshold of the gauge under test, P_{UT} is the 30-minute accumulation from the gauge under test, and P_{DFIR} is the 30-minute DFIR accumulation. The resultant thresholds were 0.18 mm for the unshielded gauge, 0.20 mm for the single Alter gauges, 0.21 mm for the double Alter gauges, 0.22 mm for the Belfort double Alter gauge, and 0.25 mm for the SDFIR gauge.

2.5.3 Choice of transfer function model

Wolff et al. (2015) tested many sigmoidal type transfer functions for the determination of catch efficiency from a single Alter gauge as a function of T_{air} and U. The equation they selected using Bayesian analysis was sigmoidal both in respect to its response to T_{air} and U. It is included here for reference, as it is used throughout this study:

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Hydrology and Earth System Sciences
Discussions

$$CE = \left[1 - \tau_1 - (\tau_2 - \tau_1)\frac{e^{\left(\frac{T_{air}-T_\tau}{s_\tau}\right)}}{1+e^{\left(\frac{T_{air}-T_\tau}{s_\tau}\right)}} e^{-\left(\frac{U}{\theta}\right)^\beta}\right] + \tau_1 + (\tau_2 - \tau_1)\frac{e^{\left(\frac{T_{air}-T_\tau}{s_\tau}\right)}}{1+e^{\left(\frac{T_{air}-T_\tau}{s_\tau}\right)}}, \quad (3)$$

where T_{air} is the air temperature, U is the wind speed, and $\tau_1, \tau_2, T_\tau, s_\tau, \theta, \beta,$, are coefficients fit to the data described in more detail in Wolff et al (2015). An example of the sigmoid function fit to SA *CE* measurements from both NOR and US is shown in Figure 4. We also propose an alternative function developed for the sake of simplicity, which responds with an exponential decrease in *CE* to wind speed, and with a simple sigmoid ($\tan^{-1}$) response to T_{air}:

$$CE = e^{-a(U)(1-[\tan^{-1}(b(T_{air}))+c])}, \quad (4)$$

where *a*, *b*, and *c* are coefficients fit to the data. The form of both of these functions follows the same form presented by others, with catch efficiency rather than correction factor determined, such that the inverse (CE^{-1}) must be used to correct actual precipitation data ($P_{DFIR} = P_{UT}/CE$). Comparisons of corrections based on correction factors (P_{DFIR}/P_{UT}) and catch efficiency (P_{UT}/P_{DFIR}) revealed no significant differences, so *CE* is used in the present study because it enables comparison with past experiments. However it is difficult to interpret *CE* error estimates (eg. Wolff et al., 2015), so special care has been taken in the present study to estimate transfer function uncertainties and biases that are relevant to the measurement of precipitation, rather than the measurement of *CE*.

2.6 Uncertainty Assessment

2.6.1 Measurement Uncertainty

The repeatability or random error of the 30-minute precipitation measurements can be estimated using the three different sets of replicate or near-replicate sets of identical gauge-shield combinations. This type of analysis includes both random gauge measurement uncertainty and uncertainty caused by the spatial variability in precipitation occurring across a field site. At the US site there were two SA-shielded gauges and two DA-shielded gauges recording precipitation measurements throughout the field study. There was also one full-sized DFIR gauge and SDFIR-shielded gauge, which were similar enough in their catch to be considered identical for the purposes of estimating measurement uncertainty. Figure 5 shows the results of the comparison of these three sets of identical or near-identical gauge-shield combinations. The root mean square errors (RMSE) were calculated from differences between the identical gauges, and were < 0.1 mm. These results include only snowfall, with precipitation type determined using the present weather detector by classifying 30-min periods with more than 15 min of snow and less than 5 min of other precipitation types as snow.

2.6.2 Uncertainty of Transfer Functions

Uncertainty in the transfer functions was estimated by applying the transfer functions to the different gauges under test and then comparing the results to the standard DFIR precipitation. To assess the SA NOR and the SA US transfer function uncertainty, the transfer function developed using the combined results from both sites was applied at the individual sites.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

This was done to assess to what degree a 'universal' transfer function was valid for the two sites. For the other US gauges, in order to maintain some independence between the data used to develop the transfer function and the data used to test the transfer function, the uncertainty of the transfer functions was estimated using a 10-fold cross-validation. This model validation technique randomly separated the available measurements of a given shield type into 10 equally sized groups, determined the transfer function using 90% of the data, tested the transfer function on the remaining 10% of the data, and was repeated for all 10 groups. The coefficients describing the transfer function were based on the entire data set, but the uncertainty estimates were based on this 10-fold cross-validation. The uncertainty estimated from 10-fold cross validation and the uncertainty estimated by circularly developing and testing the transfer function on identical data were very similar, but the 10-fold validation was used where possible for more defensible estimates of the transfer function uncertainty. For the SDFIR-shielded gauge under test it proved impossible to constrain the sigmoid function for all 10 iterations of the model fitting, so the cross-validation was not used and the uncertainty and the transfer function were circularly determined using exactly the same data.

3 Results and Discussion

3.1 Shielding Errors and Biases

Errors in uncorrected precipitation measurements were estimated by comparing the DFIR-shielded precipitation measurements to the other shielded and unshielded measurements. Based on the differences between the 60-min DFIR accumulation and the single Alter (SA) accumulation at the NOR site, root mean square errors (RMSE) and biases were calculated (Table 1, NOR SA). The RMSE of the NOR SA measurements was 0.64 mm (51.6%) and the bias was -0.34 mm (27.1%), indicating that errors in the uncorrected data were significant. At the US site the 30-min precipitation measurements were similarly used to calculate RMSEs and biases (Table 1), with the RMSEs and biases generally decreasing in absolute magnitude as the size of the shield increased. For example the absolute magnitudes of the unshielded gauge (US UN) RMSE (0.30 mm or 28.6%) and bias (-0.17 mm or -16.2%) at the US site were much larger than the SDFIR (US SDFIR) RMSE (0.14 mm or 14.7%) and bias (-0.03 mm or -3.6%). As expected, the RMSE and bias for the combined SA dataset (All SA) that included US and NOR measurements fell between the US SA and the NOR SA results. In addition to quantifying the errors occurring with the different shields at the sites, these uncorrected results also serve to provide some prospective for the corrections that were developed and applied to the precipitation measurements. Error and bias estimates reported in both mm and percent further demonstrate that a relatively small error or bias reported in mm can actually be quite significant in terms of percent. This is due to the fact that many of the 30 or 60 minute precipitation measurements, particularly for snow, included less than 0.5 mm of accumulation.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

3.2 Transfer Functions and Uncertainty

From the US site, the resultant precipitation catch efficiencies were described as a function of T_{air} and U using both the sigmoid function (Eq. 3) and exponential function (Eq. 4) for the unshielded (UN), double Alter (DA), Belfort double Alter (BDA), and small DFIR (SDFIR) Geonor T-200B precipitation gauges. In addition, the single alter (SA) results from the NOR and US sites were combined and used to create exponential (Exp) and sigmoid (Sig) transfer functions (labelled as 'All SA' in Table 3 and 4). The SA transfer functions developed using all the data were used to determine the uncertainty in the combined dataset ('All SA' in Table 2 and 3). This was done by individually applying transfer functions developed from SA data to the two separate sites and calculating the uncertainty and biases (shown in NOR SA and US SA in Tables 2 and 3). This approach was chosen to evaluate site-biases and the effects of different climates on the transfer functions. Results are described separately for the gauge-height wind speeds (Table 2) and the 10 m wind speeds (Table 3). For example, for the SA gauges, at the US site the gauge height wind speed Exp transfer function reduced the RMSE from 23.6% (Table 1) to 15.4% (Table 2) and improved the bias from -11.7% (Table 1) to 0.1% (Table 2). Likewise the NOR SA RMSE was decreased from 51.6% (Table 1) to 36.9% (Table 2) and the bias was improved from -27.1% (Table 1) to -3.3% (Table 2). This indicates that the transfer function effectively removed the bias at both sites. The RMSE of the corrected results reflects the significant residual variability in the corrected catch efficiency. It is worth noting that the RMSE of even the corrected SDFIR measurements was greater than 0.1 mm, indicating such uncorrected errors may be due to random measurement error and site variability rather than crystal type and wind speed; even a gauge that is relatively unaffected by wind speed with respect to the DFIR .

Generally, differences between the Exp and Sig functions were quite small, indicating that the Exp function can be used as a simpler alternative to the Sig function developed by Wolff et al. (2015). For the US site, the transfer function RMSE was about 0.15 mm (15%) for all the gauges irrespective of whether the sigmoid function or the exponential function was used. The errors in the uncorrected data (Table 1) were generally much larger than the errors in the corrected data, with the errors in the uncorrected data dependent upon the efficacy of the shield. For example the corrected SDFIR gauge RMSE (0.13 mm or ~14%) and bias (-0.01 mm or -1%) shown in both Table 2 and Table 3 were only slightly better than the uncorrected error (0.14 mm or 15%) and bias (-0.03 mm or -4%) shown in Table 1, whereas application of the transfer function resulted in a much more significant improvement in the unshielded and SA gauge error, (UN and SA, in Tables 1, 2, and 3).

For the combined NOR and US SA transfer functions, the RMSE were much larger for the NOR measurements than the US measurements. This was due primarily to a generally more noisy gauge at the NOR site than at the US site. Random measurement error from these vibrating-wire weighing gauges can be reduced via trial and error by rotating and remounting the vibrating wires within the gauge and also by mounting the shield separately from the gauge, but such noise can vary significantly from gauge to gauge and even from wire to wire within a gauge equipped with redundant vibrating wires. The

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

NOR site is also windier than the US site and the undercatch is generally larger, and because of this one can expect the RMSE to increase. A well-shielded gauge that requires less correction will be less affected by variability in precipitation type and crystal habit, for example. In addition blowing snow may have increased the RMSE of the NOR results, with 2.6% of the events occurring wind speeds greater than 15 ms^{-1}.

The bias found for the NOR SA gauge is perhaps however more notable, as the bias should be relatively unaffected by random measurement noise. The NOR SA measurements corrected using the gauge-height transfer function had a larger bias (approximately -0.05 mm or -4% as the mean of the Exp and Sig function results) than the US SA gauge (0.00 mm or ~0%) possibly because there were more US SA events (N = 843) than NOR SA events (N = 352), and the resultant transfer function was therefore weighted more towards the US SA catch efficiency (Table 2 and 3). More importantly, this NOR SA bias indicates that small site-biases exist that can effect such 'universal' transfer functions, and such a bias has been quantified here using automated gauges for the first time.

For the gauges at the US site, there was no difference in the RMSE or bias between the 10 m wind speed and the gauge-height wind speed transfer functions. However, this is neither surprising nor noteworthy as the gauge-height wind speed was estimated based on the 10 m wind speed. It is notable however that the RMSE for the combined US and NOR SA gauge results was also not significantly affected by the choice of wind speed measurement height (Table 2 and 3). These RMSE data indicate that although the gauge heights at the US site and the NOR site were significantly different, there was no significant loss in accuracy when transfer functions were created for both sites using the 10 m wind speed. There was however a more negative bias in the 10 m *WS* NOR SA results (eg. Sig: 0.10 mm, or -7.8%) than for the gauge height *WS* NOR SA results (eg. Sig: -0.06 mm, or -5.2%), indicating that there may be a small advantage to using the gauge height wind speed in preference to the 10 m wind speed for the development and application of precipitation transfer functions.

To demonstrate both the importance and the limitations of the transfer function corrections, the 30-min uncorrected (Fig 6a) and corrected (Fig 6b) SA snow (T_{air} < -2.5 °C) measurements were compared to the DFIR precipitation. The uncorrected SA snow measurements are subject to significant errors as a result of variability in wind speed and crystal type, with dense, wet, warm snow and low-wind speed snow less affected by shielding than cold, dry, light and windy snow. In addition, the corrected SA snow measurements reveal the effects of gauge noise, the spatial variability of precipitation, and also variability in crystal habit that are inadequately captured by temperature. For example, at a given temperature variability in crystal type has a significant effect on hydrometeor fall velocity, drag coefficient, and the resultant relationship between wind speed and *CE* (Theriault et al., 2012; Colli et al., 2015).

To further demonstrate the effects of errors and variability in the transfer functions, 1 min accumulations during the Figure 1 March, 2013 event were corrected using the appropriate transfer function and the mean 1-min T_{air} and U (Fig. 7). This 'typical' event was in fact somewhat atypical, with *CE* lower than predicted by the transfer functions. This is another

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

example of the results shown in Figure 6b, with some periods significantly overcorrected and others significantly under-corrected by transfer functions. These types of events serve as a good example of why it is always preferable to make the most accurate measurement possible, and only rely upon corrections when absolutely necessary. In comparison with the uncorrected precipitation values shown in Figure 1 however, Figure 7 also demonstrates the necessity of such transfer functions. For example, with a standard DFIR accumulation of 22.5 mm for the entire event, the UN accumulation was improved from 8.5 mm (38% of the DFIR) to 17.8 mm (79% of the DFIR).

4 Discussion

4.1 Wind Speed for Transfer Functions

Although transfer functions based on the 10 m wind speed are included in this study, we recommend using the gauge height wind speed when a gauge height wind speed measurement is available or an approximation of the gauge height wind speed. Because differences in gauge height are common due to the necessity of mounting gauges and shields well above the highest expected snow depth, even when only 10 m wind speed measurements are available they should be adjusted to gauge height for a more defensible application of shielding corrections. For example, at a gauge height of 5 m, the wind speed affecting the catch efficiency of the gauge would typically be ~90% of the 10 m wind speed, while the wind speed affecting a gauge at 2 m would be ~70% of the 10 m wind speed. Using the log wind profile and/or other available wind speed profile measurements as demonstrated here, the commonly available 10 m wind speed can be used to estimate gauge height wind speeds and defensibly correct wind speed precipitation errors at all gauge heights. However in the present study, where an estimated gauge height wind speed and a 10 m wind speed were both used to create transfer functions for gauges at significantly different heights, the uncertainty (RMSE) of the combined NOR and US transfer function showed no significant change when tested on the US, NOR, or combined datasets. The only notable change caused by the use of the 10 m wind speed version of the combined transfer function was in the bias calculated from the NOR SA, which was about -4% (-0.05 mm) for the gauge height transfer function and -8% (-0.11 mm) for the 10 m *WS* transfer function.

Wind speed measurements at the gauge-height and the standard 10 m measurement height both have advantages. The 10 m height wind speed is widely available as the standard wind speed measurement height, and measurements at this height are also less likely to be affected by obstacles such as towers and precipitation gauge shields. The effects of obstacles at gauge height were apparent in both the NOR and US sites, for example. However the catch efficiency of a shielded or unshielded gauge is more closely linked to the wind speed at the gauge height. If, for example, the wind speed at gauge height is affected by the changing height of the snowpack or by vegetation or other obstacles this will affect the relationship between the 10 m and gauge height wind speed, and potentially lead to additional sources of error. It is worth noting however that at the NOR site, where snow depth during this measurement campaign was a maximum of almost 2 m, changing snow depth did not significantly affect the relationship between the 10 m wind speed and the gauge height wind speed.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

4.2 A 'Universal' Transfer Function

Uncertainties and biases associated with the development and application of a single transfer function for multiple sites have been presented here. The fact that the NOR measurements were generally noisier than the US results and the NOR site is much windier than the US site makes it difficult to draw site-specific conclusions based on the RMSE. However the bias found for the NOR SA results was more significant than for the US SA measurements. This indicates that the combined transfer function was biased towards the US SA results, which outnumbered the NOR SA results by over 2:1. This suggests that small site biases do indeed exist, and the application of a 'universal' transfer function based on wind speed and air temperature will be subject to such site-to-site variability. The soon-to-be released results of the WMO Solid Precipitation Intercomparison Experiment will provide a better opportunity for quantifying such site-biases, as measurements from the two sites from this study will be included along with many other sites. Such a 'universal' transfer function is needed because site-specific transfer functions only exist at sites where a DFIR is present, and therefore a more generally applicable transfer function is necessary for real-world precipitation corrections where the actual or DFIR-shielded amount of precipitation is unknown. In this study, the bias was shown to be minimized at two separate sites using one transfer function that was developed from combined results, and we suggest using this same approach in future studies such as the WMO Solid Precipitation Intercomparison, for which many more sites will be included and 'universal' transfer functions can be defined more defensibly.

5. Conclusions

Methods to address the effects of wind on precipitation measurements have been presented, and significant improvements to the measurements have been made available for unshielded gauges and other widely used wind shields. A simpler function to describe catch efficiency as a function of air temperature and wind speed performed as well as the more complex function suggested by Wolff et al. (2015). In addition, the remaining uncertainty in the transfer functions used to correct or standardize precipitation measurements has also been carefully described and quantified. Significant errors persisted in the measurements even after correction for undercatch, with the RSME reduced by less than 50% for every windshield examined. Measurement error, the random spatial variability of precipitation, and variability in the type, size, density, and fall speed of hydrometeors all likely contributed to the errors that remained uncorrected. This is an active ongoing area of research that merits more attention. In addition, this study indicates that low-porosity windshields like the Belfort double Alter show great promise in reducing undercatch with a small-footprint, low-maintenance shield and merit further research.

Significant errors exist in our historical and present-day precipitation measurements. For weighing gauges that are designed to measure snowfall, these errors are affected primarily by shielding, precipitation type, crystal habits, wet vs. dry snow, and wind speed. Such errors affect the measurement of the amount of water in both seasonal and ephemeral snowpack, and therefore affect our ability to quantify the availability of water for communities and ecosystems that rely upon water from

snowfall. The results and techniques presented here can be used to help create precipitation records that are traceable to a common standard, ultimately leading to a better constrained and more accurate understanding of the earth's hydrological balance.

5. References

Adam, J. C., and Lettenmaier, D. P.: Adjustment of global gridded precipitation for systematic bias, Journal of Geophysical Research-Atmospheres, 108, 10.1029/2002jd002499, 2003.

Alter, J. C.: Shielded storage precipitaiton gages, Monthly Weather Review, 65, 262-265, doi:10.1175/1520-0493(1937)65<262:SSPG>2.0.CO;2, 1937.

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303-309, 10.1038/nature04141, 2005.

Colli, M., Rasmussen, R., Thériault, J. M., Lanza, L. G., Baker, C. B., and Kochendorfer, J.: An improved trajectory model to evaluate the collection performance of snow gauges, Journal of Applied Meteorology and Climatology, 54, 1826-1836, 10.1175/JAMC-D-15-0035.1, 2015.

Folland, C. K.: Numerical models of the raingauge exposure problem, field experiments and an improved collector design, Quarterly Journal of the Royal Meteorological Society, 114, 1485-1516, 10.1256/smsqj.48406, 1988.

Forland, E. J., and Hanssen-Bauer, I.: Increased precipitation in the Norwegian Arctic: True or false?, Climatic Change, 46, 485-509, 10.1023/a:1005613304674, 2000.

Goodison, B., Louie, P., and Yang, D.: The WMO solid precipitation measurement intercomparison, World Meteorological Organization-Publications-WMO TD, 65-70, 1997.

Groisman, P. Y., Koknaeva, V. V., Belokrylova, T. A., and Karl, T. R.: Overcoming biases of precipitation measurement - a history of the USSR experience, Bulletin of the American Meteorological Society, 72, 1725-1733, 10.1175/1520-0477(1991)072<1725:obopma>2.0.co;2, 1991.

Groisman, P. Y., and Legates, D. R.: The accuracy of United-Sates precipitaiton data, Bulletin of the American Meteorological Society, 75, 215-227, 10.1175/1520-0477(1994)075<0215:taousp>2.0.co;2, 1994.

Leeper, R. D., Palecki, M. A., and Davis, E.: Methods to Calculate Precipitation from Weighing-Bucket Gauges with Redundant Depth Measurements, Journal of Atmospheric and Oceanic Technology, 32, 1179-1190, 10.1175/jtech-d-14-00185.1, 2015.

Merenti-Valimaki, H. L., Lonnqvist, J., and Laininen, P.: Present weather: comparing human observations and one type of automated sensor, Meteorological Applications, 8, 491-496, 10.1017/s1350482701004108, 2001.

Nipher, F. E.: On the determination of the true rainfall in elevated gauges, American Association for the Advancement of Science, 103-108, doi:10.1175/1520-0493(1937)65<262:SSPG>2.0.CO;2, 1878.

Nitu, R., Rasmussen, R., Smith, C. D., Earle, M., Roulet, Y. A., Reverdin, A., Wolff, M., Baker, C. B., and Kochendorfer, J.: WMO solid precipitation intercomparison: from experiments to results, 96th American Meteorological Society Annual Meeting, New Orleans, LA, 2016.

Rasmussen, R., Dixon, M., Hage, F., Cole, J., Wade, C., Tuttle, L., McGettigan, S., Carty, T., Stevenson, L., Fellner, W., Knight, S., Karplus, E., and Rehak, N.: Weather support to deicing decision making (WSDDM): A winter weather nowcasting system, Bulletin of the American Meteorological Society, 82, 579-595, 10.1175/1520-0477(2001)082<0579:wstddm>2.3.co;2, 2001.

Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A. P., Black, J., Theriault, J. M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., and Gutmann, E.: How Well Are We Measuring Snow: The NOAA/FAA/NCAR Winter Precipitation Test Bed, Bulletin of the American Meteorological Society, 93, 811-829, 10.1175/bams-d-11-00052.1, 2012.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Springer Netherlands, 1988.

Theriault, J. M., Rasmussen, R., Ikeda, K., and Landolt, S.: Dependence of Snow Gauge Collection Efficiency on Snowflake Characteristics, Journal of Applied Meteorology and Climatology, 51, 745-762, 10.1175/jamc-d-11-0116.1, 2012.

Thom, A. S.: Momentum, mass and heat exchange of plant communities, Vegetation and the Atmosphere, Vol. 1, edited by: Monteith, J. L., Academic Press, 1975.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, Bulletin of the American Meteorological Society, 84, 1205-+, 10.1175/bams-84-9-1205, 2003.

Wolff, M., Brækkan, R., Isaksen, K., and Ruud, E.: A new testsite for wind correction of precipitation measurements at a mountain plateau in southern Norway, Proceedings of WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO–2010), Geneva, 2010,

Wolff, M., Isaksen, K., Braekkan, R., Alfnes, E., Petersen-Overleir, A., and Ruud, E.: Measurements of wind-induced loss of solid precipitation: description of a Norwegian field study, Hydrology Research, 44, 35-43, 10.2166/nh.2012.166, 2013.

Wolff, M. A., Isaksen, K., Petersen-Overleir, A., Odemark, K., Reitan, T., and Braekkan, R.: Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study, Hydrology and Earth System Sciences, 19, 951-967, 10.5194/hess-19-951-2015, 2015.

Wong, K.: Performance of several present weather wensors as precipitation gauges, Proceedings of WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO–2012), Brussels, Belgium, 2012, 25,

Yang, D.: Double Fence Intercomparison Reference (DFIR) vs. Bush Gauge for "true" snowfall measurement, Journal of Hydrology, 509, 94-100, http://dx.doi.org/10.1016/j.jhydrol.2013.08.052, 2014.

Yang, D. Q., Goodison, B. E., Metcalfe, J. R., Golubev, V. S., Elomaa, E., Gunther, T., Bates, R., Pangburn, T., Hanson, C. L., Emerson, D., Copaciu, V., and Miklovic, J.: Accuracy of Tretyakov precipitation gauge: Result of WMO intercomparison, Hydrological Processes, 9, 877-895, 10.1002/hyp.3360090805, 1995.

Yang, D. Q., Goodison, B. E., Metcalfe, J. R., Louie, P., Leavesley, G., Emerson, D., Hanson, C. L., Golubev, V. S., Elomaa, E., Gunther, T., Pangburn, T., Kang, E., and Milkovic, J.: Quantification of precipitation measurement discontinuity induced by wind shields on national gauges, Water Resources Research, 35, 491-508, 10.1029/1998wr900042, 1999.

Yang, D. Q., Kane, D., Zhang, Z. P., Legates, D., and Goodison, B.: Bias corrections of long-term (1973-2004) daily precipitation data over the northern regions, Geophysical Research Letters, 32, 10.1029/2005gl024057, 2005.

6. Tables and Figures

Shield	Uncor RMSE	Uncor Bias
US UN	0.30 mm, 28.6%	-0.17 mm, -16.2%
All SA	0.35 mm, 34.3%	-0.16 mm, -16.1%
NOR SA	0.64 mm, 51.6%	-0.34 mm, -27.1%
US SA	0.22 mm, 23.6%	-0.11 mm, -11.7%
US DA	0.21 mm, 21.6%	-0.10 mm, -10.6%
US BDA	0.16 mm, 17.5%	-0.05 mm, -5.6%
US SDFIR	0.14 mm, 14.7%	-0.03 mm, -3.6%

Table 1. Errors and biases in the uncorrected 30-min precipitation from gauges under test, estimated using the DFIR precipitation measurements as the standard.

Shield	Sig Coef (τ_1, τ_2, T_τ,	Sig RMSE	Sig Bias	Exp Coef (a,b,c)	Exp RMSE	Exp Bias	n	Max U

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Hydrology and Earth System Sciences
Discussions

	s_τ, θ, β)							
US UN	0.31, 0.94, -0.08, 0.92 2.58, 1.23	0.18 mm, 16.8%	0.00 mm, 0.4%	0.063, 1.22, 0.66	0.18 mm, 16.8%	0.00 mm, 0.4%	843	6 m/s
All SA	0.20, 0.96, 0.22, 1.11, 4.70, 1.97	0.23 mm, 22.3%	-0.02 mm, -1.5%	0.04, 1.1, 0.54	0.25 mm, 24.0%	-0.02 mm, -1.7%	1501	9 m/s
NOR SA	--	0.46 mm, 36.9%	-0.06 mm, -5.2%	--	0.45 mm, 36.1%	-0.04 mm, -3.3%	352	9 m/s
US SA	--	0.14 mm, 14.5%	0.00 mm, 0.0%	--	0.15 mm, 15.4%	0.00 mm, -0.1%	1156	9 m/s
US DA	0.00, 0.92, -1.19, 1.89, 7.04, 1.36	0.13 mm, 13.8%	0.00 mm, 0.0%	0.028, 0.74, 0.66	0.14 mm, 13.8%,	0.00 mm, 0.1%	1392	6 m/s
US BDA	0.00, 1.00, 1.81, 0.57, 8.73, 2.87	0.12 mm, 13.6%	-0.01 mm, -1.7%	0.115, 0.32, 0.38	0.13 mm, 14.6%	0.00 mm, -0.9%	1204	6 m/s
US SDFIR	0.99, 0.96, 0.52, 0.10, 0.14, 6.16	0.13 mm, 14.2%	-0.01 mm, -1.6%	0.006, 0.00, 0.00	0.13 mm, 14.2%	-0.01 mm, -1.6%	1508	6 m/s

Table 2. Transfer function results for estimated gauge height wind speeds. Coefficients for the sigmoid transfer function (Sig, Eq. 3) and the exponential transfer function (Exp, Eq. 4) and the associated RMSEs and Biases as well as the number of periods available (*n*) and the maximum wind speed (Max *U*) are described for the US unshielded (UN), single Altar (SA), double Altar (DA), Belfort double Alter (BDA), small DIFR (SDFIR), the NOR single Altar (SA) gauge, and the combined US and NOR SA results (All SA). The same coefficients determined by the All SA transfer function were used to separately correct the results from NOR and US.

Shield	Sig Coef ($\tau_1, \tau_2, T_\tau, s_\tau, \theta, \beta$)	Sig RMSE	Sig Bias	Exp Coef (a,b,c)	Exp RMSE	Exp Bias	*n*	Max *U*
US UN	0.31, 0.94, -0.08, 0.92 3.58, 1.23	0.18 mm, 16.8%	0.00 mm, 0.4%	0.045, 1.21, 0.66	0.19 mm, 17.5%	0.00 mm, 0.6%	843	8 m/s
All SA	0.17, 0.96, 0.23, 1.11,	0.24 mm, 22.9%	-0.02 mm, -1.8%	0.03, 1.04,	0.25 mm, 24.4%	-0.02 mm, -2.0%	1501	12 m/s

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Hydrology and Earth System Sciences
Discussions

	6.46, 2.01			0.57				
NOR SA	--	0.47 mm, 37.8	-0.10 mm, -7.8%	--	0.45 mm, 35.8%	-0.12 mm, -8.5%	352	12 m/s
US SA	--	0.14 mm, 14.7%	0.00 mm, 0.5%	--	0.15 mm, 15.5%	0.00 mm, 1.00%	1156	12 m/s
US DA	0.00, 0.92, -1.19, 1.89 9.75, 1.36	0.13 mm, 13.9%,	0.00 mm, 0.0%	0.021, 0.74, 0.66	0.14 mm, 13.8%,	0.00 mm, 0.1%	1390	8 m/s
US BDA	0.31, 1.00, 1.79, 0.58, 10.0, 3.15	0.12 mm, 13.6%	-0.01 mm, -1.7%	0.01, 0.48, 0.51	0.13 mm, 14.6%	0.00 mm, -0.9%	1204	8 m/s
US SDFIR	0.99, 0.96, 0.52, 0.10, 0.14, 10.75	0.13 mm, 14.1%	-0.01 mm, -0.09%	0.004, 0.00, 0.00	0.13 mm, 14.2%	-0.01 mm, -1.6%	1508	8 m/s

Table 3. Transfer function results for 10 m height wind speeds. Coefficients for the sigmoid transfer function (Sig, Eq. 3) and the exponential transfer function (Exp, Eq. 4) and the associated RMSEs and Biases as well as the number of periods available (*n*) and the maximum wind speed (Max U) are described for the US unshielded (UN), single Altar (SA), double Altar (DA), Belfort double Alter (BDA), small DIFR (SDFIR), the NOR single Altar (SA) gauge, and the combined US and NOR SA results (All SA). The same coefficients determined by the All SA transfer function were used to separately correct the results from NOR and US.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Hydrology and Earth System Sciences
Discussions

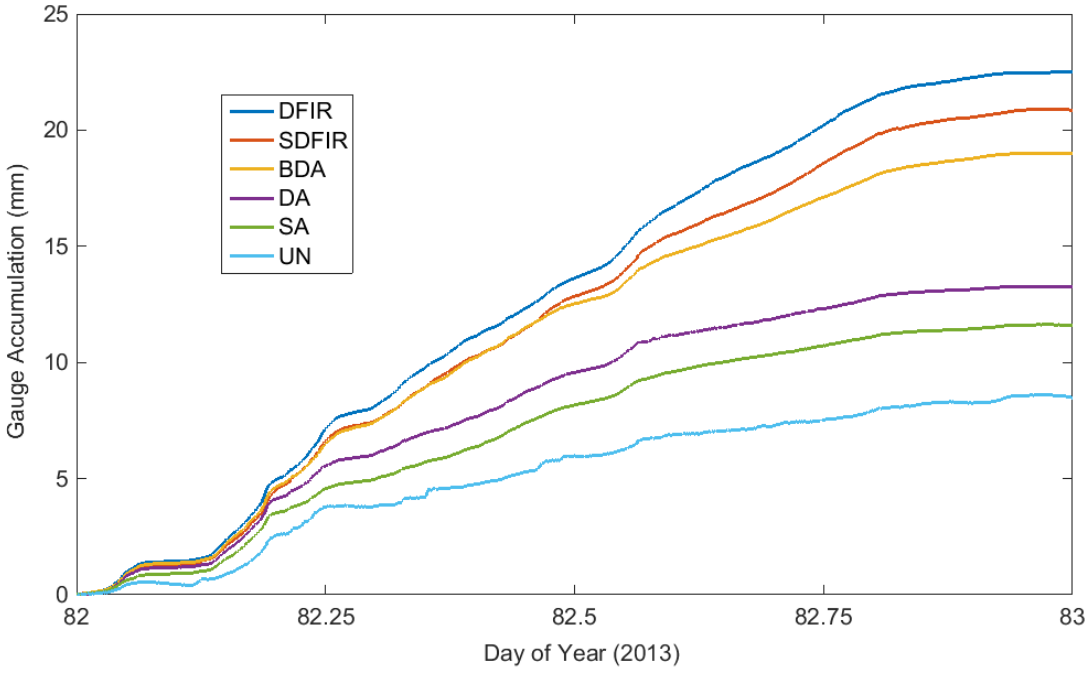

Figure 1: Example event from the US site with accumulated precipitation measured using the Double Fence Intercomparison Reference (DFIR, dark blue), small DFIR (SDFIR, red), Belfort double Alter (BDA, yellow), standard double Alter (DA, purple), single Alter (SA, green), and unshielded (UN, light blue) weighing precipitation gauges. The 24 hr mean T_{air} was -6.6 °C, and the mean gauge height wind speed was 3.6 m s^{-1}.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Hydrology and Earth System Sciences
Discussions

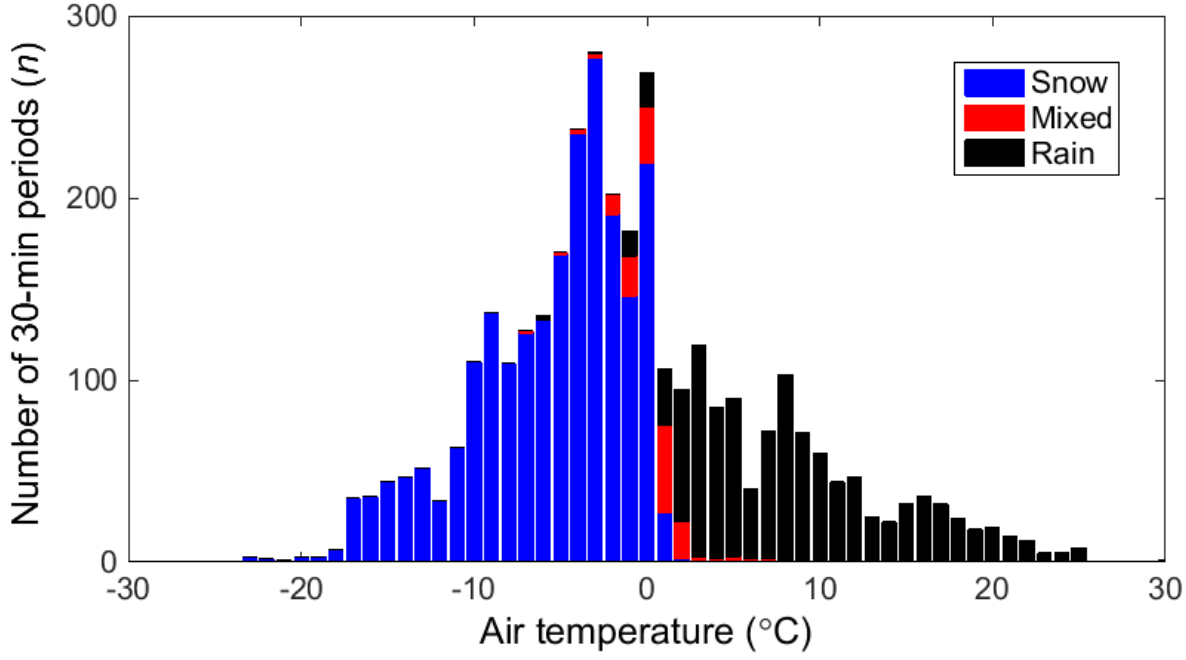

Figure 2: The temperature distribution of 30-min periods classified as snow, mixed, and rain from the US site.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

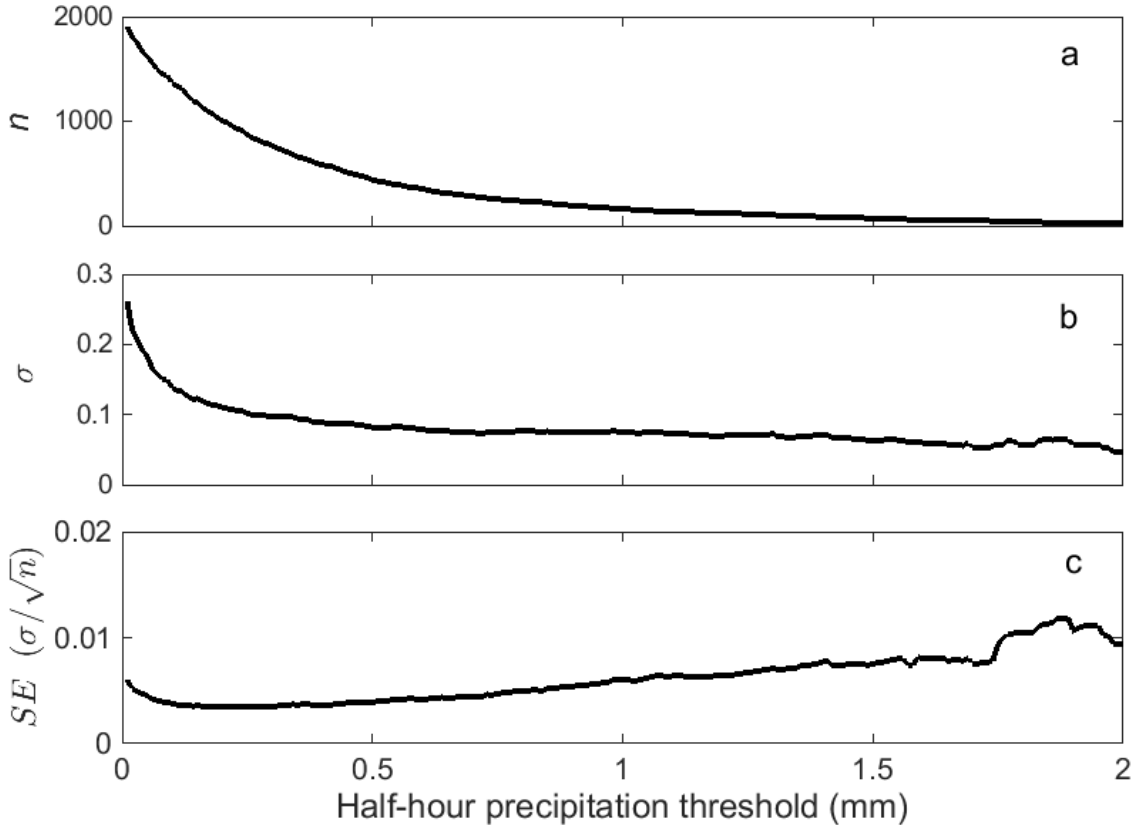

Figure 3: The effects of minimum threshold on transfer function development for 30-min periods from the US site. The number of 30-min periods (n) above the threshold (a), the standard deviation (σ) of the linear transfer function error (b), and the standard error ($SE = \sigma/\sqrt{n}$) of the transfer function (c) are shown. Only snow results are included here, with snow in this case identified by the present weather detector with more than 15 min of snow and less than 5 min of other precipitation types within each 30-min period.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Hydrology and Earth System Sciences
Discussions

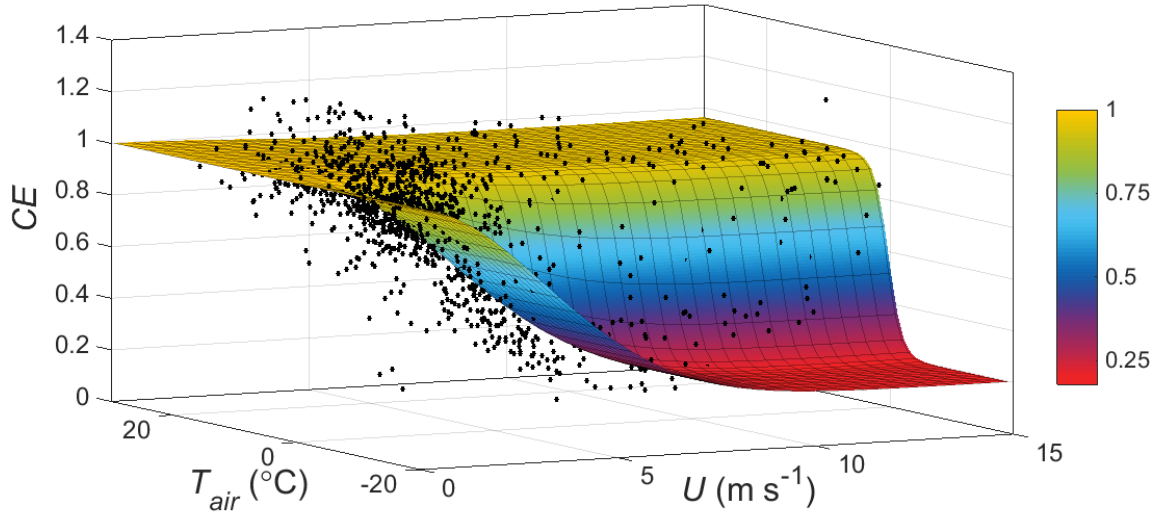

Figure 4: Example transfer function using the sigmoid function to describe the combined SA catch efficiency (*CE*) measurements from both the US and NOR sites as a function of air temperature (T_{air}) and gauge-height wind speed (*U*). Individual 30-min *CE* measurements are shown (black circles) along with the sigmoid function fit to them (coloured surface), with the colour of the surface indicating *CE* magnitude.

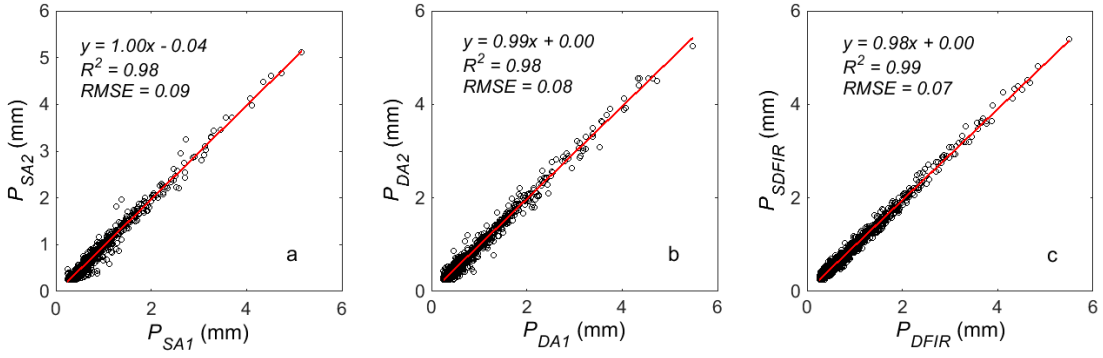

Figure 5: Comparison of 30-min precipitation (*P*) measured from three pairs of replicate or near-replicate gauge-shield combinations. Two single Alter (SA) gauges (a), two double Alter (DA) gauges (b), and a double fence intercomparison reference (DFIR) and a small DFIR (SDFIR) are compared. Only snow data are included here, with snow in this case identified by the Present Weather Detector.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-415, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 9 September 2016

Hydrology and Earth System Sciences
Discussions

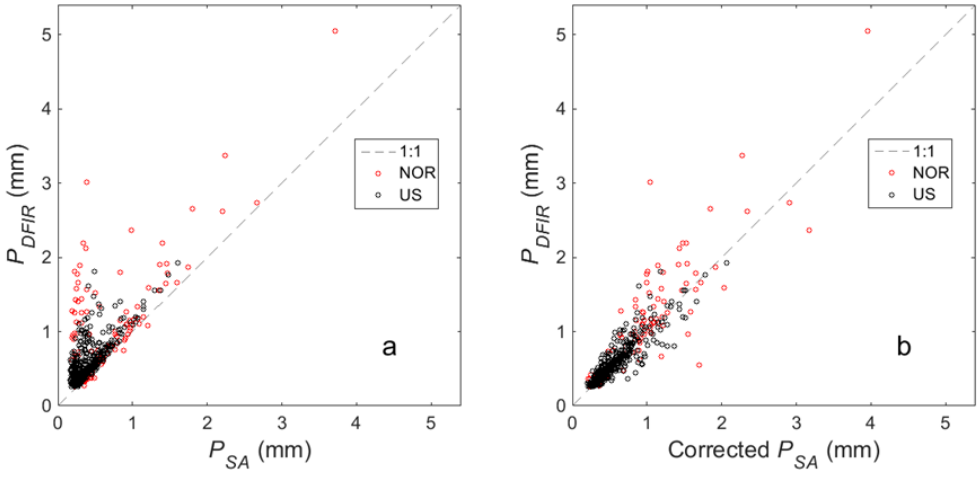

Figure 6: – Uncorrected (a) and corrected (b) SA precipitation (P_{SA}) vs. DFIR precipitation (P_{DFIR}) for snow only, where snow is here defined as T_{air} < -2.5 °C.

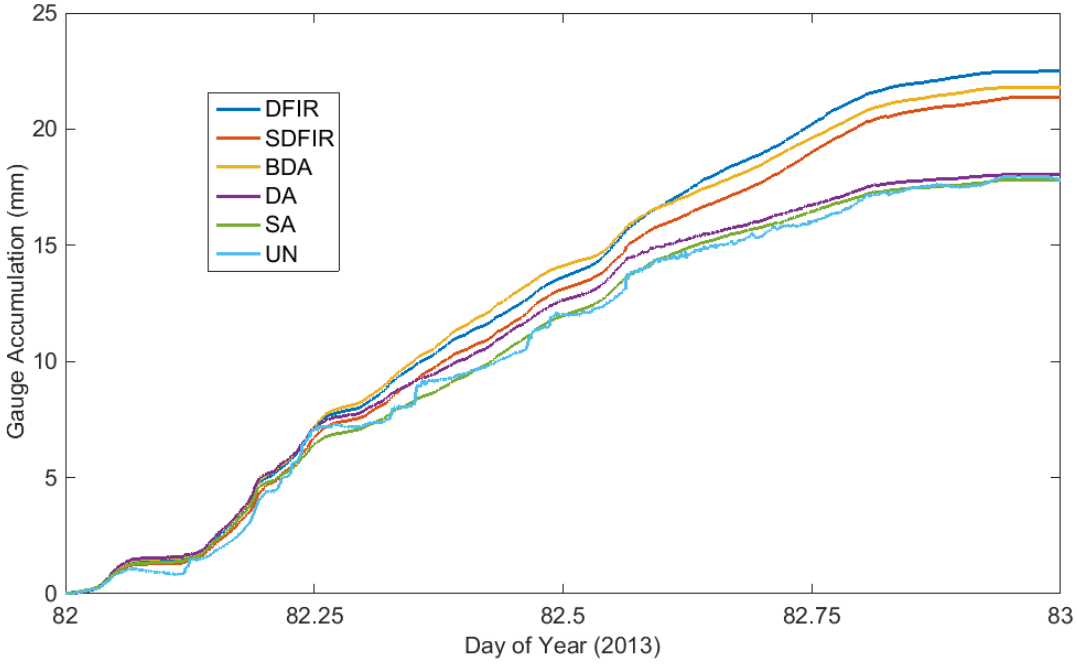

Figure 7: Example of improvement to the one day event from the US site shown in Figure 1 after applying the appropriate Exp corrections to the 1 min accumulations.